2019.1 Addendum: Paxata Technical Release Notes

The following are updates to your base 2019.1 Release Notes.

2019.1 SP5

Date: December 5, 2019

Paxata version: 2019.1.5.0.14263

<table>
<thead>
<tr>
<th>Introduced and Resolved in SP5</th>
</tr>
</thead>
<tbody>
<tr>
<td>FEATURE ENHANCEMENTS</td>
</tr>
<tr>
<td>• APF: graph now displays the Project dataset versions used in a Flow:</td>
</tr>
</tbody>
</table>

![Diagram of APF graph displaying Project dataset versions](image)

• APF: Project Flows List page is no longer case-sensitive.

• DataRobot connector enhancements: import of snapshotted/materialized datasets from the AI Catalog now supported; export to AI catalog now also supported. See the Data Source Support Matrix for details.

• Project Script Editor is now read-only. Previously all users, by default, could open and edit the Project through the script editor. The editor is now read-only by default.

Note: a group-level guardrail—#38. Enable Script editing via Raw Editor—has been introduced for customers who want to enable the edit mode.

• Disabled the import of XML External Entity (XXE) references to mitigate XXE attacks. Important: as a result of disabling XXE references, it is possible that some files containing XXE references will not be supported for import into Paxata.

• Mongo update to allow SCRAM-SHA-256 authentication.
Introduced and Resolved in SP5

• Pipeline enhancement: clicking on the "About" panel in the application inadvertently started a Pipeline while getting the Pipeline build info (build number details.) Starting a Pipeline is now no longer required to gather build info.

• SAML timeout token value is now configurable through a new PX property: px.sso.saml.maxAuthenticationAge
  The default value is 28800 seconds (8 hours).

RESOLVED

• APF automatic reimport for datasets with tables: selecting the "configure reimport options" on the Inputs tab for table-based datasets results in import failure, and the dataset must be manually reimported. Now Fixed. Original issue reported under Paxata JIRA ticket: AUTO-1321

• APF: when Japanese locale is enabled, clicking "View Project" in the Inputs tab for a dataset does not open associated Project. Now fixed. Original issue reported under Paxata JIRA ticket: AUTO-1324

• When importing a dataset into a Project, date parse errors incorrectly returned a cache error to user. Now fixed—an EvalError with remediation details are returned, and the dataset is successfully imported into Project. Original issue reported under Paxata JIRA ticket: PIPE-3002

• Pipeline error: "Could not instantiate class" error with jobs failing. Now fixed through optimizing deserialization/serialization in cache that improves compatibility of class binaries. Original issue reported under Paxata JIRA ticket: PIPE-3013

2019.1 SP4 Update 03
Date: November 07, 2019
Paxata version: 2019.1.4.3.13910

Resolved in SP4 Update 03

RESOLVED

• Updated S3 AWS Hadoop Library to version 3.2.1, which enables SSE-KMS support. This enables customer to manage the Master Keys (in addition to AWS). Original issue reported under Paxata JIRA ticket: LIB-3799

2019.1 SP4 Update 02
Date: October 29, 2019
Paxata version: 2019.1.4.2.13760

Resolved in SP4 Update 02

RESOLVED

• APF: "Datasource config not found" error displays when a lens is exported to any external data source. The error prevents users from opening the Outputs tab for the Flow. Now fixed. Original issue reported under Paxata JIRA ticket: AUTO-1326
**2019.1 SP4 Update 01**  
**Date:** October 22, 2019  
**Paxata version:** 2019.1.4.1.13334

### Resolved in SP4 Update 01

<table>
<thead>
<tr>
<th>RESOLVED</th>
</tr>
</thead>
</table>
| - Double byte character issue: garbled text for new column name that was generated through a Shape Step with double byte characters. Now fixed.  
  *Original issue reported under Paxata JIRA ticket: PROJ-4509*
| - Incorrect column data when a column is renamed to reuse a name previously used by a different column that was removed from the Project. This issue is now resolved but remediation steps may be required for your Projects. See below for details.  
  *Original issue reported under Paxata JIRA ticket: PIPE-3011* |

### Am I affected?

- **Issue:** if you have a column that reuses the name of a previously removed column in the Project, you may be affected.

- **Who is impacted:** any Project in which you remove more than one column and then reuse at least one of the removed column names for another column. Reuse may occur in one of two ways: (1) a column was renamed to reuse name of removed column OR (2) a Lookup or Append Step introduced a new column of the same name as the removed column.

  **Example one:** your Project has four columns with the names X1, X2, X3, X4. Columns X3 and X4 are then removed and, in a subsequent Step, column X2 is renamed to X4 (leaving only the columns X1 and X4 in your Project.)  
  *Expected behavior:* The newly renamed column X4 should contain data from X2.  
  *Actual behavior:* after the Step that renamed X2 to X4, column X4 contains data from the old column X4, which was previously removed.

  **Example two:** your Project has three columns Y1, Y2 and Y3. Columns Y1 and Y2 are then removed from the Project, leaving only Y3. A subsequent Lookup is then performed with a dataset that has a column named Y1, which is the name of a column that existed in the Project prior to the Lookup Step.

  *Expected behavior:* Y1 contains data from column Y1 in the Lookup dataset.  
  *Actual behavior:* Y1 contains data from the column that was removed from the original dataset, prior to the Lookup Step.

- **Remediation:** Please contact Paxata Customer Success for a tool you can use to identify potentially impacted Projects.
4
4

Paxata Confidential ©2019 Paxata, Inc.

Introducted and Resolved in SP4

CONNECTORS
• New connectors:
  o ThoughtSpot (export)
  o Hubspot (import)
  o Google Sheets (import)
  o CDH6—HDFS and Hive

See Data Source Support Matrix for details.

ENHANCEMENTS
• APF: when creating a new Flow, the APF engine now detects when there are multiple versions of a Project in the Flow and automatically selects the newest version of the Project for the Flow.
• Export format options: new ability to set the Unicode (UTF-8) "Byte Order Mark" (U+FEFF) for export formats that support this option. Original issue reported under Paxata JIRA ticket: LIB-3765.

RESOLVED
• Double-byte character issues resolved:
  o When publishing an AnswerSet with a filter, the double-byte character column name in the publish description is garbled and the Publish button is disabled. Now fixed. Original issue reported under Paxata JIRA ticket: PROJ-4500.
  o When publishing an AnswerSet in which the description exceeds the max chars, the Publish button is disabled with no error to indicate how to resolve. Now fixed. Original issue reported under Paxata JIRA ticket: PROJ-4488
  o Column names are garbled in Computed Column expression errors. Now fixed. Original issue reported under Paxata JIRA ticket: PROJ-4493.
  o When publishing an AnswerSet that has a sorted column, the column name in the description for publish is garbled. Now fixed. Original issue reported under Paxata JIRA ticket: PROJ-4500
  o When GroupBy (Pivot) is selected in the Shape menu, the initial display column name is garbled. Now fixed. Original issue reported under Paxata JIRA ticket: PROJ-4488.
• BigQuery connector: OAuth appears to be unverified. Now fixed. Original issue reported under Paxata JIRA ticket: CONN-1779
• Zendesk connector: ExternalId is interpreted on import as BigInt instead of String, resulting in errors for Users and Tickets Zendesk collection. Now fixed. Original issue reported under Paxata JIRA ticket: CONN-1762.
• The number of columns generated in Pivot and Transpose are limited to 1000 even when the Project column limit is set to enable more than 1000 columns. Now fixed. Original issue reported under Paxata JIRA ticket: CONN-4457.
• When datasets with long column names are appended together, the append feature trims the column names to ellipses. Two columns that start and end with the same characters can then no longer be distinguished from one another. Now fixed. Original issue reported under Paxata JIRA ticket: PROJ-4482.
• Library Versions page: when scrolling down through the versions of a dataset, the page occasionally resets to the top of the list. Now fixed. Original issue reported under Paxata JIRA ticket: UI-199.
## Introduced and Resolved in SP4

- **APF feature:** in some cases, the UI does not display a link to view errors for a Flow and, instead, displays the static message `{{RD.translatedState}}`. Now fixed. *Original issue reported under Paxata JIRA ticket: AUTO-1319.*

- **Pipeline:** long-running, orphaned TCP connections were capable of bringing down the system when available ports ran out. The root cause for orphaning the connections has been identified and this issue is now resolved. *Original issue reported under Paxata JIRA ticket: PIPE-3003.*

- **Pipeline:** cache failures due to canceled jobs during Lookup operation is now fixed. *Original issue reported under Paxata JIRA ticket: PIPE-2952.*

- **REST API:** encoding and `includeBOM` export options cannot be overridden in the Datasource Exports REST API. Now fixed. *Original issue reported under Paxata JIRA ticket: LIB-3759.*

- **Log files:** `dataFileExported` log reports incorrect ID for datafile. Now fixed. *Original issue reported under Paxata JIRA ticket: LIB-3768.*

## PLATFORM UPDATES

- New support for Mongo versions 4.0 and 4.2.

## 2019.1 SP3 Updates 1-4

**Date:** October 03, 2019  
**Paxata version:** 2019.1.3.4.13001

### Resolved in SP3 updates 1-4

- **JP customers:** updates to fix UI and connector issues with double byte characters in Paxata Projects, columns, AnswerSet descriptions and tenant names.  
  *Original issues reported under Paxata JIRA tickets: PROJ-4491, PROJ-4488, CONN-1685*

- **Pipeline OOM exception and subsequent server failure.** This issue was due to a resource leak when an internal resource connection was not properly closed during times of heavy loads on the system. Now fixed. *Original issue reported under Paxata JIRA ticket: PIPE-2988*

## 2019.1 SP3

**Date:** September 17, 2019  
**Paxata version:** 2019.1.3.2.12615

### Introduced and Resolved in SP3

#### CONNECTORS

- New connectors:  
  - Google BigQuery  
  - DataRobot  
  
  See the [Data Source Support Matrix](https://example.com) for details.

- **UI enhancement:** the "Test Data Source" button, on the Connector and Data Source configuration forms, has been moved to the top of the forms for better visibility.
## FEATURES AND ENHANCEMENTS

- **APF:** ownership of a Flow can now be transitioned from the person who initially setup the Flow. This feature can currently only be enabled through the REST API. See the APF API guide and refer to the "Change Owner Interface" section.

- **Security enhancements:**
  - Introduced expiration times for Web Authentication tokens. Default expiration time is now 14 days for these tokens. The default is configurable through the PX property: `px.auth.cookie.duration.secs`.
  - Files with specific extensions that are blocked by browsers are now also blocked on export from Paxata. By default, Paxata now blocks the following file extensions on export: `action`, `apk`, `app`, `bat`, `bin`, `cmd`, `com`, `command`, `cpl`, `csh`, `exe`, `gadget`, `infl`, `ins`, `inx`, `ipa`, `isu`, `job`, `jse`, `ksh`, `lnk`, `msc`, `msi`, `msp`, `mst`, `osx`, `out`, `paf`, `pif`, `prg`, `ps1`, `reg`, `rgs`, `run`, `scr`, `sct`, `shb`, `sht`, `vb`, `vbe`, `vbs`, `vbscript`, `workflow`, `ws`, `wsf`, `wsh`  
    
    This enhancement has been implemented through the introduction of a new guardrail in the application: "Library Export Blocked Extensions". This is a global guardrail setting but can be set per tenant level.
    
    This guardrail can also be set through the REST API. See `guardrail.library.export.blacklist` in the Guardrails API for details. Note that when a REST call is made to export file types that are blocked, a 400 error is returned.

- **Japanese translations of the UI, UI help and Help Shelf articles in the application are now available in this release. Original issue reported under Paxata JIRA ticket: DOC-1073**

- Paxata Administrators can now configure a default encoding value for both import and export. A new PX Property, `px.encoding.fallback`, defines a default encoding for connectors—for both parsers (import) and formatters (export). The default setting is UTF-8. All supported encodings are defined here: [https://docs.oracle.com/javase/8/docs/technotes/guides/intl/encoding.doc.html](https://docs.oracle.com/javase/8/docs/technotes/guides/intl/encoding.doc.html).  
    
    Note that individual import and export requests can override the encoding.
    
    IMPORTANT: if your Paxata installation is already using a specific `px.parser.encoding.fallback` setting, you must rename that property to `px.encoding.fallback` so that your existing encodings remain enforced.

## RESOLVED

- **Japanese customers: broken image link in the "Getting Started" Help Shelf article. Now fixed. Original issue reported under Paxata JIRA ticket: DOC-1073**

- **REST endpoint /cache/entries issued but times out with no response. Now fixed. Pipeline now recognizes when executors and the API respond with expected cache information. Original issue reported under Paxata JIRA ticket: PIPE-2933**
RESOLVED

- When publishing datasets containing date or timestamp values and those datasets are published with either the Direct Data Load or Parallel Ingest for Hive features enabled, the publish to Library may fail—with or without an indication that publish was unsuccessful. Now fixed.

*Original issue reported under Paxata JIRA ticket: PIPE-2965.*

Additionally, in cases where the dataset is successfully published, the resulting date/timestamp values in the dataset are incorrect and the dataset must be republished now that the root cause of the issue has been corrected. See below to determine if you have datasets that must be republished.

- **Who is impacted?** This issue impacts only production clusters that have the Direct Data Load (DDL) feature or the Parallel Ingest for Hive feature enabled, or had it enabled during the time AnswerSets were produced from Paxata Projects. Contact Paxata Customer Success if you are uncertain regarding the enabled state of these features.

- **Which Projects are affected?** Only Paxata Projects that have columns with date or timestamp values are affected.

- **What is the impact of this issue?** If the previous two conditions apply, then any datasets that were successfully published to the Library will contain erroneous date/timestamp data.

- **What are the remediation steps to take for redressing this issue?** If your Paxata Projects have been affected by this issue, you will need to republish the output for those Projects.

Note: Updates 01 and 02 were initially made to the 2019.1 SP2 Cloud software and then automatically included in the base GA base release for on premise customers.