

# Oracle® Cloud

## Extract Data from Unstructured Files Using OCI Vision



F74331-02  
August 2023



Oracle Cloud Extract Data from Unstructured Files Using OCI Vision,  
F74331-02

Copyright © 2022, 2023, Oracle and/or its affiliates.

Primary Author: Oracle Corporation

# Contents

## 1 About This Recipe

---

Overview	1-1
System and Access Requirements	1-1

## 2 Before You Install the Recipe

---

## 3 Install and Configure the Recipe

---

Configure the Oracle REST Vision API Connection	3-1
Configure the Oracle FTP Vision Invoice Rename Connection	3-2
Update Integration Properties	3-2

## 4 Activate and Run the Recipe

---

# Preface

This document describes how to install, configure, and run this recipe in Oracle Integration 3.

**Topics:**

- [Documentation Accessibility](#)
- [Diversity and Inclusion](#)
- [Related Resources](#)
- [Conventions](#)

## Documentation Accessibility

For information about Oracle's commitment to accessibility, visit the Oracle Accessibility Program website at <http://www.oracle.com/pls/topic/lookup?ctx=acc&id=docacc>.

**Access to Oracle Support**

Oracle customers that have purchased support have access to electronic support through My Oracle Support. For information, visit <http://www.oracle.com/pls/topic/lookup?ctx=acc&id=info> or visit <http://www.oracle.com/pls/topic/lookup?ctx=acc&id=trs> if you are hearing impaired.

## Diversity and Inclusion

Oracle is fully committed to diversity and inclusion. Oracle respects and values having a diverse workforce that increases thought leadership and innovation. As part of our initiative to build a more inclusive culture that positively impacts our employees, customers, and partners, we are working to remove insensitive terms from our products and documentation. We are also mindful of the necessity to maintain compatibility with our customers' existing technologies and the need to ensure continuity of service as Oracle's offerings and industry standards evolve. Because of these technical constraints, our effort to remove insensitive terms is ongoing and will take time and external cooperation.

## Related Resources

For more information, see these Oracle resources:

- Oracle Integration documentation in the Oracle Cloud Library on the Oracle Help Center.
- Oracle Cloud at <http://cloud.oracle.com>.

## Conventions

The following text conventions are used in this document.

<b>Convention</b>	<b>Meaning</b>
<b>boldface</b>	Boldface type indicates graphical user interface elements associated with an action, or terms defined in text or the glossary.
<i>italic</i>	Italic type indicates book titles, emphasis, or placeholder variables for which you supply particular values.
monospace	Monospace type indicates commands within a paragraph, URLs, code in examples, text that appears on the screen, or text that you enter.

# 1

## About This Recipe

Use this recipe to read and extract data from unstructured files such as PDF, JPG, PNG or TIFF using Oracle Cloud Infrastructure (OCI) Vision.



### Note:

This recipe is available as **FTP Server — OCI Vision | Extract Data** in the Integration Store. Oracle provides this recipe as a sample only. The recipe is meant only for guidance, and is not warranted to be error-free. No support is provided for this recipe.

## Overview

This recipe extracts data from unstructured files such as image files and PDFs stored in an FTP server using OCI Vision, and then rewrites the processed files to the FTP server.

To use the recipe, you must install the recipe and configure the connections and other resources within the recipe. Subsequently, you can activate and run the integration flow of the recipe manually or specify an execution schedule for it. When triggered the integration flow reads the files from the FTP server, extracts relevant data using OCI Vision, and then renames and writes the processed files back to the FTP server. Note that the recipe can run a maximum of 30 files at a time.

## System and Access Requirements

- Oracle Integration, Version 21.2.1 or higher
- Oracle Cloud Infrastructure (OCI)
- OCI Vision
- A federated user account in OCI
- A secure FTP (sFTP) server
- An FTP client to access the FTP server

# 2

## Before You Install the Recipe

You must perform the following configuration tasks on your FTP server and OCI instance in order to successfully connect to these external systems using Oracle Integration and extract data from unstructured files.


1. Access your FTP Server and upload files that are to be processed.
  - a. Log in to the server using your user name and password through an FTP client; for example, FileZilla.
  - b. Create a directory on the server from which to read the files. Note down the path of this directory.
  - c. Upload sample files into the directory you've created. Note down the format of the files.
2. Log in to your OCI instance as an **Administrator** and perform the following tasks. For more information, refer to the Oracle Cloud Infrastructure documentation.
  - a. Create a compartment.
  - b. Create a federated user.
  - c. Assign a role to the federated user.
  - d. Create a local user.
  - e. Create a user group.
  - f. Create a policy.
  - g. Obtain the following for creating a connection to OCI Vision from Oracle Integration. See section **OCI Signature Version 1 Security Policy Use** in Prerequisites for Creating a Connection.
    - i. Get your Tenancy and User OCIDs.
    - ii. Create your own private key or generate it from the OCI console.
    - iii. Get the key's fingerprint.

# 3

## Install and Configure the Recipe

On your Oracle Integration instance, install the recipe to deploy and configure the integration and associated resources.

1. On the Oracle Integration Home page, in the **Use a recipe/accelerator** section, click **View all**.
2. Find the recipe you want to install, then click **Get**.  
A message confirms that the recipe was successfully installed, and the recipe card shows **In use**.

3. Click **Configure**  on the recipe to configure its resources.

The project workspace opens, displaying all the resources of the recipe. Configure the following resources before you activate and run the recipe.

### Configure the Oracle REST Vision API Connection

1. In the Connections section, click the connection name.
2. In the Properties section, enter the following details:

Field	Information to Enter
Connection Type	Select <b>REST API Base URL</b> .
Connection URL	Enter the endpoint URL of your OCI Vision instance. For example: <code>https://vision.aiservice.&lt;tenancy location&gt;.oci.oraclecloud.com/20220125</code>

3. In the Security section, enter the following details:

Field	Information to Enter
Tenancy OCID	Enter the tenancy OCID that you obtained. See <a href="#">Before You Install the Recipe</a> .
User OCID	Enter the user OCID that you obtained.
Private Key	Click <b>Upload</b> and select the API key generated earlier.


 **Note:**

Before you upload the private key, you must convert it into the PKCS1 format.

Finger Print	Enter the finger print that was generated for the key.
--------------	--

4. Click **Save**. If prompted, click **Save** again.



5. Click **Test** to ensure that your connection is successfully configured. In the resulting dialog, click **Test** again.  
A message confirms if your test is successful.
6. To return to the project workspace, click **Go back** .


## Configure the Oracle FTP Vision Invoice Rename Connection

1. In the Connections section, click the connection name.
2. In the Properties section, enter the following details:

Field	Information to Enter
<b>FTP Server Host Address</b>	Enter the host address of your sFTP server.
<b>FTP Server Port</b>	Enter 22.
<b>SFTP Connection</b>	Select <b>Yes</b> from the list.


3. In the Security section, enter the following details:

Field	Information to Enter
<b>Security Policy</b>	Select <b>FTP Server Access Policy</b> .
<b>User Name</b>	Enter the user name to connect to your sFTP server.
<b>Password</b>	Enter the password to connect to your sFTP server.

4. Click **Save**. If prompted, click **Save** again.
5. Click **Test** to ensure that your connection is successfully configured. In the resulting dialog, click **Test** again.  
A message confirms if your test is successful.
6. To return to the project workspace, click **Go back** .

## Update Integration Properties

Update the integration properties with appropriate values.

1. In the Integrations section, click **Actions**  on the integration flow, then select **Update property values**.
2. In the Update property values panel, update the integration properties with appropriate values.
  - a. **filePattern**: This integration property holds the format of the files to be read from the FTP server. For example: \*.pdf
  - b. **directory**: This integration property holds the path of the directory on the FTP server from where the files must be read. For example: /user2/InvoiceDemo.
3. Click **Submit**.

A message confirms that the integration properties have been updated successfully.

# 4


## Activate and Run the Recipe

After you've configured the connections and other resources, you can activate and run the recipe.

1. In the project workspace, click **Activate**. In the Activate project panel, with the default project deployment selected, choose an appropriate tracing option, then click **Activate**.

A message confirms that the integration has been activated. Refresh the page to view the updated status of the integration.

2. Run the recipe.

- a. In the Integrations section of the project workspace, click **Actions**  on the integration flow, then select **Run**.

- b. On the Configure and run page, click **Run**.

You've now successfully submitted the integration for execution.

### Note:

You can also schedule this integration to run at a date, time, and frequency of your choosing. See [Define the Integration Schedule](#).

3. Monitor the running of the integration flow in Oracle Integration.
  - a. In the project workspace, click **Observe**. You'll see the integration flow being triggered and running successfully.
  - b. To manage errors in your project, see [Manage Errors in a Project](#).
4. Log in to your FTP server, and check if the processed files have been updated.

### Related Documentation

- [Using the REST Adapter with Oracle Integration 3](#)
- [Oracle Infrastructure Cloud Documentation](#)