

Course Exercises Guide

# Developing Applications in IBM Datacap

## V9.1.7

Course code WF318G ERC 1.0



## **June 2020 edition**

### **Notices**

This information was developed for products and services offered in the US.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:

*IBM Director of Licensing  
IBM Corporation  
North Castle Drive, MD-NC119  
Armonk, NY 10504-1785  
United States of America*

INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some jurisdictions do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.

IBM may use or distribute any of the information you provide in any way it believes appropriate without incurring any obligation to you.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to actual people or business enterprises is entirely coincidental.

### **Trademarks**

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the web at "Copyright and trademark information" at [www.ibm.com/legal/copytrade.shtml](http://www.ibm.com/legal/copytrade.shtml).

**© Copyright International Business Machines Corporation 2020.**

**This document may not be reproduced in whole or in part without the prior written permission of IBM.**

US Government Users Restricted Rights - Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

# Contents

<b>Trademarks . . . . .</b>	<b>vii</b>
<b>Exercises description . . . . .</b>	<b>viii</b>
Exercise objectives . . . . .	viii
User IDs and passwords . . . . .	ix
How to use the course exercise instructions . . . . .	ix
Checking for course corrections . . . . .	xi
<b>Exercise 1. Processing Datacap batches . . . . .</b>	<b>1-1</b>
Section 1. Prepare your system: Start the components . . . . .	1-3
1.1. Verify WebSphere Application Server deployment manager . . . . .	1-3
1.2. Start WebSphere Application Server components . . . . .	1-3
1.3. Check IBM FileNet P8 Platform and IBM Content Navigator . . . . .	1-4
1.4. Start Datacap Server . . . . .	1-6
1.5. Open Datacap Desktop . . . . .	1-6
Section 2. Process a batch in Datacap Desktop . . . . .	1-8
2.1. Explore the Datacap Desktop user interface . . . . .	1-8
2.2. Scan a batch . . . . .	1-9
2.3. Run the PagID task . . . . .	1-11
2.4. Run the Batch Profiler task . . . . .	1-12
2.5. Run the Verify task . . . . .	1-13
2.6. Run the Export task . . . . .	1-16
Section 3. Explore Datacap Navigator . . . . .	1-18
3.1. Start Datacap Navigator for business users . . . . .	1-18
3.2. View Job Monitor . . . . .	1-18
3.3. Check batch information . . . . .	1-19
3.4. Explore the User Settings . . . . .	1-21
Section 4. Process batches in Datacap Navigator . . . . .	1-22
4.1. Scan batches . . . . .	1-22
4.2. Review the Scan task output . . . . .	1-23
4.3. Start Datacap Rulerunner Service . . . . .	1-24
4.4. Run the PagID and Profiler tasks . . . . .	1-24
4.5. Run the Verify task . . . . .	1-25
4.6. Review the Verify task output . . . . .	1-27
4.7. Check your batch status . . . . .	1-27
4.8. Review the Export task output . . . . .	1-27
4.9. Stop Datacap Rulerunner Service . . . . .	1-28
<b>Exercise 2. Building a Datacap application with Forms Template . . . . .</b>	<b>2-1</b>
Section 1. Create a Datacap application . . . . .	2-2
1.1. Create a Datacap application . . . . .	2-2
1.2. Check the Datacap.xml file and application folder . . . . .	2-5
Section 2. Configure the application in IBM Content Navigator . . . . .	2-7
2.1. Add the new Datacap application as a repository . . . . .	2-7
2.2. Associate the new repository with the Datacap desktops . . . . .	2-9
2.3. Edit the workflow for the new Datacap application . . . . .	2-10
Section 3. Edit the application settings in Datacap Application Manager . . . . .	2-16
3.1. Check the Main tab . . . . .	2-16
3.2. Check the Rulerunner tab . . . . .	2-17
Section 4. Edit the task profiles and rulesets . . . . .	2-18

4.1. Configure task profiles for the new application .....	2-18
4.2. Check and clean the rulesets .....	2-20
Section 5. Define document hierarchy .....	2-23
5.1. Copy the images to be scanned to the new application folder .....	2-23
5.2. Create document hierarchy for the new application .....	2-23
Section 6. Configure fingerprints and zones .....	2-26
6.1. Configure the FingerprintAdd ruleset .....	2-26
6.2. Add the FingerprintAdd ruleset to a task profile .....	2-29
6.3. Configure Fingerprinting in Zones tab .....	2-30
6.4. Configure zones .....	2-32
6.5. Set OCR filters for the fields .....	2-33
<b>Exercise 3. Configuring rulesets for a Datacap application .....</b>	<b>3-1</b>
Section 1. Create ruleset for the VScan task .....	3-2
1.1. Create the VScan ruleset .....	3-2
1.2. Add the ruleset to the task profile .....	3-6
1.3. Start a batch to test the VScan ruleset .....	3-6
Section 2. Configure PagelD for the application .....	3-9
2.1. Configure the PagelD ruleset for barcode recognition .....	3-9
2.2. Add a page level rule to the PagelD ruleset .....	3-11
2.3. Configure the CreateDocs&Fields ruleset .....	3-13
2.4. Add the two rulesets to the PagelD task profile .....	3-15
2.5. Test the Page ID task for the application .....	3-15
Section 3. Add rulesets for the Profiler, Verify, and Export tasks .....	3-18
3.1. Configure the Recognize ruleset .....	3-18
3.2. Configure a rule for reading zones .....	3-19
3.3. Configure a rule for field extraction by using keyword search .....	3-20
3.4. Configure the Validate ruleset .....	3-22
3.5. Add field level rules to the Validate ruleset for checking the fields .....	3-22
3.6. Add rulesets to the Profiler, Validate, and Export task profiles .....	3-26
3.7. Configure the Export task .....	3-27
Section 4. Test the application .....	3-29
4.1. Run the VScan and PagelD tasks .....	3-29
4.2. Run the Profiler task .....	3-30
4.3. Advance the batch to Verify task .....	3-31
4.4. Process the Export task and verify the export file .....	3-31
4.5. Test the application with scanning multiple images in a batch .....	3-32
<b>Exercise 4. Troubleshooting a Datacap application .....</b>	<b>4-1</b>
Section 1. Examine the logs files for debugging .....	4-2
1.1. Examine the logs for a batch .....	4-2
1.2. Empty the images folder and run a batch .....	4-4
1.3. Examine the logs .....	4-6
1.4. Copy the files back to the images folder .....	4-7
Section 2. Troubleshoot the application in Datacap Studio .....	4-8
2.1. Edit a ruleset to introduce an error .....	4-8
2.2. Start a batch to test the application .....	4-8
2.3. Step the batch through PagelD .....	4-9
2.4. Cancel the batch .....	4-11
Section 3. Correct the error in the application and test .....	4-12
3.1. Edit a ruleset to correct the error .....	4-12
3.2. Test the application .....	4-12
<b>Exercise 5. Building an application to process multiple page types .....</b>	<b>5-1</b>
Section 1. Copy a Datacap application .....	5-2
1.1. Copy a Datacap application .....	5-2

Section 2. Configure the new application in Datacap Navigator . . . . .	5-5
2.1. Add the new Datacap application as a repository . . . . .	5-5
2.2. Associate the new repository with the Datacap user desktop . . . . .	5-8
2.3. Edit the workflow for the new Datacap application . . . . .	5-9
Section 3. Verify the application in Datacap Application Manager . . . . .	5-11
Section 4. Define Document hierarchy . . . . .	5-12
4.1. Copy the images to be scanned to the new application folder . . . . .	5-12
4.2. Create a document hierarchy for the new application . . . . .	5-13
Section 5. Configure rulesets and task profiles . . . . .	5-15
5.1. Configure the PageID ruleset for page recognition without barcode . . . . .	5-15
5.2. Configure the CreateDocs&Fields ruleset . . . . .	5-17
5.3. Configure the Recognize ruleset . . . . .	5-18
5.4. Configure the Validate ruleset . . . . .	5-18
5.5. Configure the Export ruleset . . . . .	5-20
Section 6. Configure Fingerprints and Zones . . . . .	5-21
6.1. Configure Fingerprinting . . . . .	5-21
6.2. Configure zones . . . . .	5-23
Section 7. Test the application . . . . .	5-24
7.1. Start a batch to test the application . . . . .	5-24
7.2. Run the Page ID task . . . . .	5-25
7.3. Run the Profiler task . . . . .	5-26
7.4. Advance the batch to Verify task and process the batch . . . . .	5-26
7.5. Process the Export task and verify the export file . . . . .	5-27
<b>Exercise 6. Configuring an application for recognition with OMR . . . . .</b>	<b>6-1</b>
Section 1. Configure OMR fields . . . . .	6-2
1.1. Configure document hierarchy for the new page . . . . .	6-2
1.2. Create a dictionary for the Options field . . . . .	6-3
1.3. Configure variables and child fields for the Options field . . . . .	6-4
Section 2. Configure rulesets and zones for OMR . . . . .	6-7
2.1. Start the Datacap Studio . . . . .	6-7
2.2. Configure the CreateDocs&Fields ruleset . . . . .	6-7
2.3. Configure the Recognize ruleset . . . . .	6-9
2.4. Configure zones . . . . .	6-10
Section 3. Test the application for OMR . . . . .	6-11
3.1. Run the VScan and PageID tasks . . . . .	6-11
3.2. Run the Profiler and Verify tasks . . . . .	6-12
3.3. Run the Export task and check the export file . . . . .	6-13
<b>Exercise 7. Exporting data to a FileNet Content Manager repository . . . . .</b>	<b>7-1</b>
Section 1. Export data to a FileNet Content Manager repository . . . . .	7-2
1.1. Start the Datacap Studio . . . . .	7-2
1.2. Add the Export to FileNet P8 ruleset . . . . .	7-2
1.3. Add a document level rule to set document parameters . . . . .	7-3
1.4. Add a page level rule to map the field values . . . . .	7-5
1.5. Add a document level rule to upload the scan document . . . . .	7-6
1.6. Add a ruleset to prepare the document for the upload . . . . .	7-8
1.7. Add a rule to reformat the currency value . . . . .	7-9
1.8. Add the rulesets to the Export task profile . . . . .	7-11
Section 2. Test the application . . . . .	7-12
2.1. Run the VScan and PageID tasks . . . . .	7-12
2.2. Run the Profiler and Verify tasks . . . . .	7-13
2.3. Run the Export task . . . . .	7-14
2.4. Check the export documents . . . . .	7-14
<b>Exercise 8. Creating page layouts . . . . .</b>	<b>8-1</b>

Section 1. Configure an application to create page layouts .....	8-2
1.1. Add the Convert PDF ruleset .....	8-2
1.2. Configure the Analyze Layout ruleset .....	8-5
1.3. Configure the Populate Layout field in DCO function .....	8-7
1.4. Add the AnalyzeLayout ruleset to the PageID task profile .....	8-8
Section 2. Test the application and verify the layout file .....	8-10
2.1. Start a batch to test the application .....	8-10
2.2. Run the Page ID task .....	8-11
2.3. Verify the layout file .....	8-12
2.4. Disable the Populate Layout field in DCO function .....	8-13
<b>Exercise 9. Creating a virtual page block and extracting values from tables .....</b>	<b>9-1</b>
Section 1. Create a Virtual Page Block .....	9-2
1.1. Check the file for virtual page block .....	9-2
1.2. Start Datacap Studio .....	9-3
1.3. Configure the Create a VPB ruleset .....	9-3
1.4. Add both VPB rulesets to the PageID task profile .....	9-6
Section 2. Test the VPB Configuration .....	9-7
2.1. Start a batch to test the application .....	9-7
2.2. Run the Page ID task .....	9-8
2.3. Check the newly created page (VPB) .....	9-9
Section 3. Configure data extraction from tables .....	9-11
3.1. Configure the Build VPB CCO ruleset .....	9-11
3.2. Configure the Table Extraction ruleset .....	9-13
3.3. Add the Build VPB CCO and Table Extraction ruleset to task profile .....	9-19
Section 4. Test the application .....	9-20
4.1. Start a batch to test the application .....	9-20
4.2. Run the Page ID task .....	9-20
4.3. Run the Profiler task .....	9-21
4.4. Advance the batch to Verify task and process the batch .....	9-21
4.5. Process the Export task and verify the export file .....	9-22
<b>Exercise 10. Extracting data from label-value pairs .....</b>	<b>10-1</b>
Section 1. Configure the Find Label Value Pairs function .....	10-2
1.1. Check the scanning document for label-value pairs .....	10-2
1.2. Start Datacap Studio .....	10-3
1.3. Configure the Label Value extraction ruleset .....	10-4
1.4. Add the Label Value extraction ruleset to the Profiler task profile .....	10-6
Section 2. Test the Find Label Value Pairs function .....	10-7
2.1. Start a batch to test the application .....	10-7
2.2. Run the Page ID task .....	10-7
2.3. Run the Profiler task .....	10-8
Section 3. Configure data extraction from label-value pairs .....	10-9
3.1. Disable the CopyLabelValuePairs () action .....	10-9
3.2. Configure the Field level rule for PolicyNumber .....	10-9
3.3. Configure the Field level rule for State .....	10-11
3.4. Configure the Field level rule for Vehicle Year .....	10-12
3.5. Configure the Field level rule for Make .....	10-13
3.6. Configure the Field level rule for Model .....	10-14
3.7. Configure the Field level rule for VIN .....	10-15
Section 4. Test the application .....	10-18
4.1. Run the VScan and Page ID tasks .....	10-18
4.2. Run the Profiler task .....	10-18
4.3. Advance the batch to Verify task and process the batch .....	10-19
4.4. Process the Export task and verify the export file .....	10-19

---

# Trademarks

The reader should recognize that the following terms, which appear in the content of this training document, are official trademarks of IBM or other companies:

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corp., registered in many jurisdictions worldwide.

The following are trademarks of International Business Machines Corporation, registered in many jurisdictions worldwide:

Db2®

Insight®

FileNet®

WebSphere®

IBM Cloud™

Windows is a trademark of Microsoft Corporation in the United States, other countries, or both.

Other product and service names might be trademarks of IBM or other companies.

# Exercises description

## Exercise objectives

After completing the exercises, students should be able to:

- Process batches in the Datacap clients for an existing application
- Build a Datacap application with Forms Template
- Examine the logs files for debugging
- Troubleshoot the applications in the Datacap Studio Test tab
- Build an application for processing multiple page types in a batch
- Configure an application for recognition with OMR
- Export data and scanned documents to a FileNet Content Manager repository
- Configure a Datacap application to create page layouts
- Create a virtual page block
- Extract data from a table
- Extract data from label-value pairs

## Exercise list

This course includes the following exercises:

- [Exercise 1, "Processing Datacap batches"](#)
- [Exercise 2, "Building a Datacap application with Forms Template"](#)
- [Exercise 3, "Configuring rulesets for a Datacap application"](#)
- [Exercise 4, "Troubleshooting a Datacap application"](#)
- [Exercise 5, "Building an application to process multiple page types"](#)
- [Exercise 6, "Configuring an application for recognition with OMR"](#)
- [Exercise 7, "Exporting data to a FileNet Content Manager repository"](#)
- [Exercise 8, "Creating page layouts"](#)
- [Exercise 9, "Creating a virtual page block and extracting values from tables"](#)
- [Exercise 10, "Extracting data from label-value pairs"](#)

# User IDs and passwords

The following table contains a list of user ID and password information for this course.

Entry point	User ID	Password
Windows Server 2016 operating system	Administrator	FileNet1
WebSphere Application Server (WAS) administration console	wasadmin	FileNet1
FileNet Content Manager repositories administrative user	p8admin	FileNet1
IBM Content Navigator administrative user	p8admin	FileNet1
IBM Datacap user (for the Datacap Studio, Datacap Desktop, and Datacap Navigator clients)	admin	admin

## How to use the course exercise instructions

### Exercise structure

Many exercises build on the previous exercises. It is strongly recommended that you follow the order as presented.

Each exercise is divided into sections with a series of numbered steps and lettered substeps:

- The numbered steps (1, 2, 3) represent actions to be done.
- The lettered substeps (a, b, c) provide detailed guidance on how to complete the action.



### Example

*Excerpt from Exercise 1*

— 1. Log in to Datacap Desktop.

    — a. Click **Start > IBM Datacap Clients > Datacap Desktop**

        You can also use the **Datacap Desktop** shortcut on the Windows desktop.

    — b. Enter the following values:

        ○ **User:** admin

        ○ **Password:** admin

        ○ **Station:** 1

\_\_\_ c. Click **Start**.

---

In this example, **Log in to Datacap Desktop** is the action to be completed. The substeps provide specific guidance on how to complete the login.

## Text highlighting in exercises

Different text styles indicate various elements in the exercises.

Words that are highlighted in **bold** represent GUI items that you interact with, such as:

- Menu items
- Field names
- Icons

Words that are highlighted with a `fixed font` include the following items:

- Text that you type or enter as a value
- System messages
- Directory paths
- Code

## Values for the data fields

The lab instructions provide values to enter for the data fields. Some of the values are not crucial (for example: Application name) and you might be able to enter different text. Keep in mind that the values are referenced in the subsequent steps. If you opt to enter a different value, you need to replace with the value that you entered. For this reason, it is convenient to use the values that are provided.

## Tracking your progress

As shown in the example step (in the **Exercise structure** section), you can see that an underscore precedes each numbered step and lettered substep.

You are encouraged to use these markers to track your progress by checking off each step as you complete it. Tracking your progress in this manner might be useful if you are interrupted while working on an exercise.

## Browser use and other tips

- This course is tested and the screen captures that are included in the Exercises are taken in **Mozilla Firefox**. However, the **Google Chrome** browser is available. The behavior may be different.
- If Firefox prompts to install updates, decline the updates.
- The exercises in this course use a set of lab files that can be found in the `C:\lafiles` folder on the Windows platform. The exercises point you to the lab files as you need them.

- The student system contains solutions for the exercises and the exercises point you to the solution as you need them.
- After completing the exercise, you are encouraged to go back over all the steps to make sure that you understand why you did each step and how you did it.

## Checking for course corrections

---



### Important

Online course material updates might exist for this course. To check for updates and course abstracts, see the IBM Cloud Education Course Information home page:

<https://ibm-learning-skills-dev.github.io/education/courseinfo.html>

---

# Exercise 1. Processing Datacap batches

## Estimated time

01:45

## Overview

In this exercise, you learn how to process batches in the Datacap clients.

## Objectives

After completing this exercise, you should be able to:

- Process batches in Datacap Desktop
- Process batches in Datacap Navigator

## Introduction

Before you start building your own application, you test an existing application by processing a batch and reviewing the output. You learn about the various tasks in a workflow for a Datacap application by running the workflow in two Datacap clients: Datacap Desktop (thick client) and Datacap Navigator (Thin client).

This exercise includes the following sections:

- [Section 1, "Prepare your system: Start the components"](#)
- [Section 2, "Process a batch in Datacap Desktop"](#)
- [Section 3, "Explore Datacap Navigator"](#)
- [Section 4, "Process batches in Datacap Navigator"](#)

## Requirements

None

## User accounts

Type	User ID	Password
Operating system	Administrator	FileNet1
FileNet Content Manager repositories administrative user	p8admin	FileNet1
IBM Content Navigator administrator	p8admin	FileNet1
IBM Datacap user (for the Datacap Studio, Datacap Desktop, and Datacap Navigator clients)	admin	admin



### Note

- Passwords are case-sensitive.
- The exercises in this course use a set of lab files that are stored in the `C:\labfiles` directory.
- The exercises point you to the lab files as you need them.

## Section 1. Prepare your system: Start the components

For the lab environment, you must start the IBM WebSphere Application Server that hosts the IBM FileNet Content Manager and IBM Content Navigator components. IBM FileNet Content Manager is used for exporting Datacap scanned documents and IBM Content Navigator is required for Datacap Navigator.

The **WebSphere Admin** folder on the desktop includes the scripts that you use to start the components.

### 1.1. Verify WebSphere Application Server deployment manager

WebSphere Application Server hosts the following applications:

- Content Platform Engine
- IBM Content Navigator
- Administration Console for Content Platform Engine (ACCE)

- \_\_\_ 1. Log in to the operating system by using the following credentials:
  - **User name:** Administrator
  - **Password:** FileNet1
- \_\_\_ 2. Click the Windows **Services** shortcut on the taskbar.
- \_\_\_ 3. Verify that the **IBM WebSphere Application Server V9.0 - EDUCellManager01** service shows the **Running** status.
- \_\_\_ 4. If the service is not already started, start the WebSphere Application Server by right-clicking the service and selecting **Start** and wait for the status to show running.
- \_\_\_ 5. Close the Windows **Services** window.

### 1.2. Start WebSphere Application Server components

- \_\_\_ 1. In Windows Explorer, from the Windows desktop, open the **WebSphere Admin** folder.
- \_\_\_ 2. Right-click the **\_1 Start node agent.bat** file, select **Run as administrator**, and wait for the command prompt to close.
- \_\_\_ 3. Right-click the **\_2 Start server1.bat** file, select **Run as administrator**, and wait for the command prompt to close.
- \_\_\_ 4. Right-click the **\_3 Start ICNserver1.bat** file, select **Run as administrator**, and wait for the command prompt to close.
- \_\_\_ 5. Close Windows Explorer.

### 1.3. Check IBM FileNet P8 Platform and IBM Content Navigator

In this section, you verify that all Content Platform Engine, IBM Content Navigator, and IBM Datacap are running. Because the applications rely on more software, testing the applications also ensures that the underlying software is also functioning properly within your system.

- \_\_\_ 1. Check the Ping page for Content Platform Engine.
  - \_\_\_ a. Open the Firefox browser, click the **Bookmarks** menu, and click **FileNet P8 System Check > CE Ping Page**.

You can also enter the following **URL**: <http://ecmedu01:9080/FileNet/Engine>

- \_\_\_ b. Verify that the **Content Engine Startup Context (Ping Page)** opens to indicate that Content Platform Engine content services are functioning properly.

This page contains information about the IBM FileNet P8 system such as the product name and version, and log file locations.

 Content Engine Startup Context (Ping Page)	
Key	
Local Host	ecmedu01
Start Time	Tue Jun 02 14:42:46 EDT 2020
Product Name	P8 Content Platform Engine - 5.5.3.0
Build Version	dap553.1500
Startup Message	P8 Content Platform Engine Startup: 5.5.3.0 dap553.1500
Operating System	Windows Server 2016 Standard 10.0
Available Processors	2

- \_\_\_ 2. Check the FileNet P8 CE System Health page.
  - \_\_\_ a. Click Firefox **Bookmarks > FileNet P8 System Check**, right-click **FileNet P8 CE System Health**, and select **Open in a New Tab**.
 

You can also enter the following **URL**: <http://ecmedu01:9080/P8CE/Health>
  - \_\_\_ b. Verify that the **IBM FileNet Content Manager - CE System Health** page opens.

The page includes information about FileNet P8 domain (EDU\_P8), object stores, and other resources. Each item has a link to view more details. The green circle shows that these resources are available.

  - \_\_\_ c. In the **Resources** section, click the **Object Stores** link to see a list of available object stores.

You access object stores as repositories in IBM Content Navigator.
- \_\_\_ 3. Check the Content Navigator Ping page.
  - \_\_\_ a. Click Firefox **Bookmarks > FileNet P8 System Check**, right-click **ICN Ping Page**, and select **Open in a New Tab**.

- You can also enter the following **URL**: <http://ecmedu01:9081/navigator/ping.jsp>
- \_\_\_ b. When you are prompted to log on, enter the following credentials and click **Log in**.
    - **User name:** p8admin
    - **Password:** FileNet1
  - \_\_\_ c. Verify that the IBM Content Navigator Ping Page is displayed to indicate that the IBM Content Navigator application is functioning properly.

IBM Content Navigator Ping Page	
Key	Value
<b>Product Name</b>	IBM Content Navigator
<b>Build Level</b>	icn306.000.5405 (201906101232)
<b>Version</b>	3.0.6

The page contains information about IBM Content Navigator including the product name and version.

- \_\_\_ 4. Open an IBM Content Navigator desktop page.
  - \_\_\_ a. From the Firefox **Bookmarks Toolbar**, right-click **Sample Desktop**, and select **Open in a New Tab**.
  - You can also enter the following **URL**: <http://ecmedu01:9081/navigator/>
  - \_\_\_ b. If you are prompted to log in, enter the following credentials and click **Log in**.
    - **User name:** p8admin
    - **Password:** FileNet1
- \_\_\_ 5. Verify that the IBM Content Navigator Desktop (**Sample Desktop**), as indicated on the upper left, opens to the **Browse** view.

As indicated on the upper right, this desktop is configured to browse the **Sales** repository by default.

The Browse view opens successfully when the following components are running and communicating within your student system:

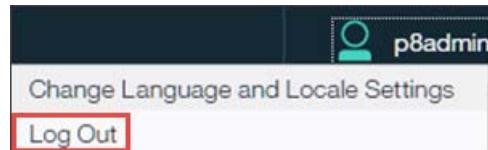
- A database system

Your student system uses the IBM Db2 database software. Every time a user logs in to the desktop, the desktop configuration is loaded from the Db2 database. This step demonstrates that the database is functional.

- An LDAP directory service to handle user authentication.

Your student system uses Active Directory.

- \_\_\_ 6. Log out of IBM Content Navigator.
  - \_\_\_ a. On the banner, click the user and select **Logout** from the list.



- \_\_\_ b. Click **Logout** and close all the browser tabs.

## 1.4. Start Datacap Server

- \_\_\_ 1. On the Desktop, double-click **Datacap Server Manager** the shortcut.

You can also open it from the Windows Start menu by clicking **IBM Datacap Services > Datacap Server Manager**.

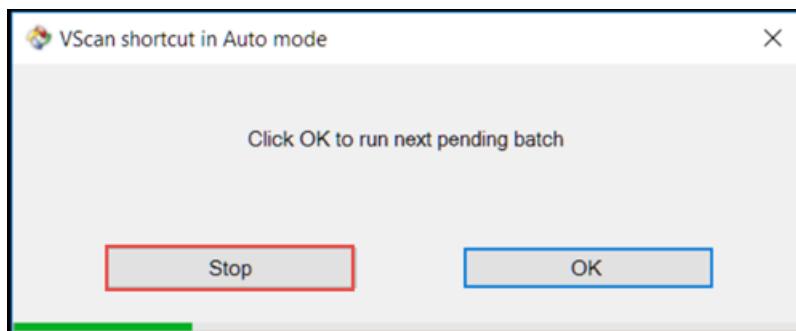
The Datacap Server Manager window opens.

- \_\_\_ 2. On the **Service** tab, verify that the service status shows **Running**.
- \_\_\_ 3. If it is not already started, click **Start** (green arrow icon) to start the service.
- \_\_\_ 4. Click **Close** to close the **Datacap Server Manager** window.

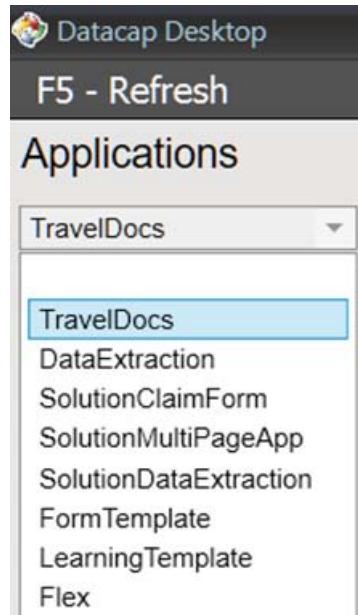
## 1.5. Open Datacap Desktop

In this section, you verify that you are able to access the Datacap Desktop client.

- \_\_\_ 1. Log in to Datacap Desktop.
  - \_\_\_ a. From the Windows Start, click **IBM Datacap Clients > Datacap Desktop**.  
On the Windows desktop, you can also use the **Datacap Desktop** shortcut.
  - \_\_\_ b. Enter the following values:
    - **User:** admin
    - **Password:** admin
    - **Station:** 1
  - \_\_\_ c. Click **Start**.
- \_\_\_ 2. If you are prompted with the **VScan shortcut in Auto mode** dialog box (Message: Click OK to run next pending page), click **Stop**.



- \_\_\_ 3. In the upper left of the navigation pane, ensure that **TravelDocs** is selected from the list of applications.



This section indicates that the server is active and connected.

- \_\_\_ 4. Leave the Datacap Desktop open for the following section.
- 



## Troubleshooting

If any of the clients are not coming up, do the following steps:

- \_\_\_ 1. In Windows **Services**, check that the **IBM WebSphere Application Server V9.0 - EDUCellManager01** service are running.  
Then, stop and start the components by using the following steps.
  - \_\_\_ 2. Open the **WebSphere Admin** folder on the Windows Desktop.
  - \_\_\_ 3. Right-click the **\_4 Stop ICNserver.bat** file, select **Run as administrator**, and wait for the command prompt to close.
  - \_\_\_ 4. Right-click the **\_5 Stop server1.bat** file, select **Run as administrator**, and wait for the command prompt to close.
  - \_\_\_ 5. Right-click the **\_6 Stop node agent.bat** file, select **Run as administrator**, and wait for the command prompt to close.
  - \_\_\_ 6. Right-click the **\_1 Start node agent.bat** file, select **Run as administrator**, and wait for the command prompt to close.
  - \_\_\_ 7. Right-click the **\_2 Start server1.bat** file, select **Run as administrator**, and wait for the command prompt to close.
  - \_\_\_ 8. Right-click the **\_3 Start ICNserver1.bat** file, select **Run as administrator**, and wait for the command prompt to close.
-

## Section 2. Process a batch in Datacap Desktop

In this section, you become familiar with the capabilities of the Datacap Desktop client and process a batch.

### 2.1. Explore the Datacap Desktop user interface

- 1. Before you begin the exercise, ensure that the components are running. If you did not start the components already, refer to [Section 1, "Prepare your system: Start the components"](#) to start them.

If you restart your system, you must start all the components.

The Datacap Desktop is opened from the previous section and the TravelDocs application is selected.

- 2. In Datacap Desktop, on the left pane, observe the available shortcuts in the **Task Shortcuts** list.
- 3. Check another application for its shortcuts.
  - a. From the list, select **Flex**.
  - b. Notice that the shortcuts vary based on what Tasks are defined for the application.
- 4. Select the **TravelDocs** Application again.
- 5. On the left pane, click the **All** shortcut and on the right pane, notice a list of batches.  
From this view, you can monitor all batches for a selected application. You can customize this view.
- 6. Hover your mouse over the status icon in the table to see the status name (Example: **pending** or **hold**) for each batch.

The **Task** column specifies the task at which the batch is pending.

Applications	Select Columns	Filter	By		
TravelDocs					
Task Shortcuts	Status	Queue ID	Batch ID	Job	Task
All		1	20200226.000000	Main Job	Verify
		2	20200228.000000	Navigator Job	Verify

- 7. In the Monitor view, select (single click) a batch to preview and In the rightmost pane, observe the details for the batch are shown.

**Note**

If you double-click the item, the item gets processed.

Status	Queue ID	Batch ID	Job	Details	Page View
	1	20200226.000000	Main		
	2	20200228.000000	Nav		

The preview has two tabs: **Details** and **Page View**.

- \_\_\_ 8. Adjust the column size and position.
  - \_\_\_ a. Drag the individual column by header to change its relative position in the table.
  - \_\_\_ b. Click a column header to sort by that column.
- \_\_\_ 9. Remove the columns from the Monitor view.
  - \_\_\_ a. On the upper left, click the **Select Columns** list.
  - \_\_\_ b. Scroll down and clear the following columns: **Docs** and **Pages**.
  - \_\_\_ c. Click outside the list and verify that the columns that you cleared are not shown on the monitor.
- \_\_\_ 10. Filter the list of batches.
  - \_\_\_ a. On the toolbar, set the **Filter** field to: **Queue ID**
  - \_\_\_ b. Set the **By** field to: **1** and press Enter.
  - Ensure the value (1) is available in **Queue ID** column.
  - \_\_\_ c. Check that only one item with the specified **Queue ID** is shown.

Select Columns		Filter	By
		Queue ID	1
Status	Queue ID	Batch ID	Job Start

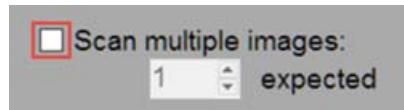
Queue ID  
Batch ID  
b 2/26/2020

## 2.2. Scan a batch

In this section, you run the Virtual Scan task to complete the page input.

- \_\_\_ 1. From the **Shortcut** list, select **Virtual Scan**.
- \_\_\_ 2. In the **Scan from** field, click the **Browse for files to import** icon.
- \_\_\_ 3. Browse to the C:\Datacap\TravelDocs\images folder, select the **car1.tif** image, and click **Open**.

- \_\_\_ d. Clear the **Scan multiple images** option.

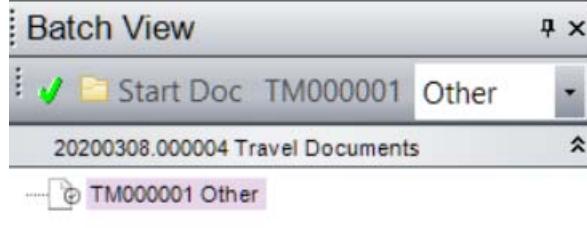


### Note

If you select only one image and select the **Scan multiple images** option (**expected=0**), it scans all of the images in the folder.

If you select multiple images, the **Scan multiple images** option is selected automatically. The expected field is set to the number of the images that you selected.

- \_\_\_ e. Click **Scan**.
- \_\_\_ 4. In the **Batch View** pane on the right, verify that the image details are listed.



The image is shown in the **Image View** on the left pane.

- \_\_\_ 5. Complete the batch scan.
- \_\_\_ a. At the end of the page, click **Submit**.
  - \_\_\_ b. Click **OK** to acknowledge the message that the scan is finished.
  - \_\_\_ c. Click **Stop** to stay on the main page.
- \_\_\_ 6. Leave **Datacap Desktop** open for the next section.
- \_\_\_ 7. Open the batch folder and review the Scan results.
- \_\_\_ a. Open Windows Explorer and browse to the batches folder:  
C:\Datacap\TravelDocs\batches
  - \_\_\_ b. Verify that a folder is created for your batch with a name that has today's date in this format: <date.00000N> where N is a number.
- Example: **20200308.000004**
- \_\_\_ 8. Check the VScan.xml file for the <date.00000N> batch.
- \_\_\_ a. Open the <date.00000N> batch folder.
  - \_\_\_ b. Right-click **VScan.xml** and select **Open with > Internet Explorer** to open it in a browser.
  - \_\_\_ c. If you are prompted to allow the scripts to run and about the security, click **Yes**.

- \_\_\_ d. Verify that the scanned page details.
- The batch **Type** is **TravelDocs**
  - The Document Type is not set yet
  - The page **type** is **Other**

<b>B</b>	<b>20200308.000004</b>
	TYPE : <b>TravelDocs</b>
	STATUS : <b>0</b>
	ScanOperator : <b>admin</b>
	ScanStation : <b>1</b>
<b>P</b>	<b>TM000001</b>
	TYPE : <b>Other</b>
	STATUS : <b>49</b>
	IMAGEFILE : <b>tm000001.tif</b>
	ScanSrcPath : <b>C:\Datacap\TravelDocs\images\Carl.tif</b>
	PageName : <b>Carl</b>

- \_\_\_ e. Close the browser tab.
- \_\_\_ 9. Leave the Windows Explorer folder open for the next task.

## 2.3. Run the PageID task

In this section, you run the PageID task to complete the page identification.

- \_\_\_ 1. In Datacap Desktop, click **All** from the **Task Shortcuts** list.
- \_\_\_ 2. Double-click the batch that you created with today's date and time and the Task name: **PageID**.
- \_\_\_ 3. Wait for the task to complete and if prompted, click **OK** to acknowledge the message that the batch is finished.
- \_\_\_ 4. Select the **All** shortcut and check that your batch is moved to the **Batch Profiler** task.



### Note

To run the **Page ID** task for all the pending batches, click the **PageID** shortcut from the Task shortcut list. It automatically processes all pending batches. The user has options to stop a particular batch or to continue to process each of the next pending batches.

- 
- \_\_\_ 5. Open your batch folder and review the PageID task output.
  - \_\_\_ a. If it is not already open, open Windows Explorer, open the batch folder:  
**C:\Datacap\TravelDocs\batches\<today's date.00000N>**

- \_\_\_ b. Right-click **PageID.xml** and select **Open with > Internet Explorer** to open it in a browser.
- \_\_\_ c. If you are prompted to allow the scripts to run and about the security, click **Yes**.
- \_\_\_ 6. Verify that the page type is now identified as **Rental\_Agreement**.

Recall that in the previous section, the page was identified as **Other** by default. After the **PageID** task, the specific page ID is assigned.

B	<b>20200308.000004</b>
	TYPE : TravelDocs
	STATUS : 0
	ScanOperator : admin
	ScanStation : 1
P	<b>TM000001</b>
	TYPE : <b>Rental_Agreement</b>
	STATUS : 49

- \_\_\_ 7. Close the browser tab and leave the Windows Explorer folder open for the next task.

## 2.4. Run the Batch Profiler task

In this section, you run the Profiler task to complete the document assembly and extract the values.

- \_\_\_ 1. In Datacap Desktop, click **All** from the **Task Shortcuts** list.
- \_\_\_ 2. Double-click the batch that you created with today's date and time and the task name, **Batch Profiler**.
- \_\_\_ 3. Wait for the task to complete and click **OK** to acknowledge the message that the batch is finished.
- \_\_\_ 4. On the job monitor, verify that your batch is moved to **Verify** task.
- \_\_\_ 5. In your batch folder, review the Batch Profiler task output.
  - \_\_\_ a. If the folder is already not open, open Windows Explorer and your batch folder:  
C:\Datacap\TravelDocs\batches\<today's date.00000N>
  - \_\_\_ b. Right-click **Profiler.xml** and select **Open with > Internet Explorer** to open it in a browser.
  - \_\_\_ c. If you are prompted to allow the scripts to run and about the security, click **Yes**.
- \_\_\_ 6. Verify that the document type is identified as **Car\_Rental** for the **Rental\_Agreement** page.

Recall that in the previous section, the page was identified as **Rental\_Agreement**.

B	20200308.000004
	TYPE : TravelDocs
	STATUS : 0
	ScanOperator : admin
	ScanStation : 1
	NeedsFixup : No
	MyProtectedPassword2 : [secured]
	018fcf898264a46d062833181a21e337419294659787cb29fe4a10b8302f4ae4b09
	[/secured]
D	20200308.000004.01
	TYPE : Car_Rental
	STATUS : 0
P	TM000001
	TYPE : Rental_Agreement
	STATUS : 1

- \_\_\_ 7. Close the browser tab and leave the Windows Explorer folder open for the next task.



#### Note

To run the **Batch Profiler** task for all the pending batches, click the **Batch Profiler** shortcut from the **Task Shortcuts** list. It automatically processes the pending batches. The user has options to stop a particular batch or to continue to process each of the next pending batches.

## 2.5. Run the Verify task

The task in this procedure completes the data validation and verification.

Datacap Desktop is already open from the previous section and TravelDocs is selected.

- \_\_\_ 1. In Datacap Desktop, process the batch for the verify step.
- \_\_\_ a. Click **All** from the **Task Shortcuts** list.
- \_\_\_ b. Double-click the batch that you created with today's date and time and the task name: **Verify**
- The Verify page opens with field names and values in the middle pane.
  - A green check mark is shown for the fields to indicate that the value is correct.

- If the confidence level for a value is low, a red X is shown for the fields.

	Name	Value
✓	Stick Text Field	
✓	Text_Field	
✓	Car pick-up date	Mon, Oct 4, 2010
✗	Car pick-up loc...	New York (JFK)
✓	Return date	Fri, Oct 8, 2010
✗	Car return locat...	New York (JFK)
✓	Car type	Full size

- \_\_\_ 2. Verify that the **Car pick-up location** and **Car return location** fields have the correct value **New York (JFK)** by comparing the values on the original scanned image on the leftmost pane.



### Hint

In the **Image View** toolbar, use the **Zoom In** icon to enlarge the image.

- \_\_\_ 3. Verify that the **Car options** field has the correct value: **101**

The value **1** indicates that the option is selected and **0** indicates that the option is not selected for the following items: **GPS Navigation**, **Child Seat**, and **Fuel Service**

- \_\_\_ 4. Complete the task.

- \_\_\_ a. On the middle pane, at the end of the page, click **Submit**.

- \_\_\_ b. Click **OK** to acknowledge the message: All documents are complete. Finish Batch?

- \_\_\_ c. Click **OK** to acknowledge the message that the batch is finished.

- \_\_\_ 5. Select the **All** shortcut and check that your batch is moved to **Export** task.

- \_\_\_ 6. In your batch folder, review the Verify task output.

- \_\_\_ a. If the folder is already not open, open Windows Explorer and your batch folder: C:\Datacap\TravelDocs\batches\<today's date.00000N>

- \_\_\_ b. Right-click **Verify.xml** and select **Open with > Internet Explorer** to open it in a browser.

- \_\_\_ c. If you are prompted to allow the scripts to run and about the security, click **Yes**.

- \_\_\_ d. Scroll down and in the **DATAFILE** field, click **tm000001.xml**

```
MatchType : Fingerprint
ZoneRead : FPXML
DATAFILE : tm000001.xml
```

- \_\_\_ e. If you are prompted to allow the scripts to run and about the security, click **Yes**.

- \_\_\_ f. Verify that the file contains values for the fields.

F	Pickup_Date
	Text value : Mon, Oct 4, 2010
	Char confi : 9999999999999999
	TYPE : Pickup_Date
	Position : 174,389,524,452
	STATUS : 0
F	Pickup_Location
	Text value : New York (JFK)
	Char confi : 9999999999999999
	TYPE : Pickup_Location
	Position : 174,515,527,575

- \_\_\_ 7. Close the browser and leave the Windows Explorer folder open for the next task.



### Information

Occasionally, after you run the Verify step, you want to rerun the PageID or Profiler task again on the same batch to make sure the page identification and extraction of data tasks were completed correctly.

In Datacap Desktop, you can configure the batch to run a task that was already run as described in the following steps:

- Select the batch that you want to edit.
- On the right pane, select the task that you want to rerun (Example: **Page ID**) from the list.

QueueID	8
BatchID	20200308.000003
Job	Main Job
Task	Verify
Status	Virtual Scan
JobStart	Page ID
JobTime	Batch Profiler AM Verify

- For the **Status** field, select **Pending** from the list.

The screenshot shows a configuration dialog with the following fields:

- QueueID: 8
- BatchID: 20200308.000003
- Job: Main Job
- Task: Page ID (highlighted with a red box)
- Status: pending

- Click **Apply** and verify that the **Task** column for your batch shows **Page ID** as you configured.

Status	Queue ID	Batch ID	Task	Job	Job Start
...	8	20200308.000003	Page ID	Main Job	3/8/2020 8:34 AM
■	2	20200228.000000	Verify	Navigator Job	2/28/2020 3:06 AM

You can now process your batch again from the **Page ID** task. It moves through **Profiler** and **Verify** steps.

## 2.6. Run the Export task

The task in this procedure completes the export of the data to a file on the local directory.

In a later exercise, you learn about exporting data to a repository, for example, FileNet Content Manager.

- 1. In Datacap Desktop, click **All** from the **Task Shortcuts** list.
- 2. Double-click the batch that you created with today's date and time and the task name: **Export**
- 3. Wait for the task to complete and in the **All** shortcut, verify that your batch is removed from the list.  
Because all the processes are completed for your batch.
4. Close Datacap Desktop.



### Note

To run the **Export** task for all the pending batches, click the **Export** shortcut from the Task shortcuts list. It automatically processes the pending batches. The user has options to stop a particular batch or to continue to process each of the next pending batches.

- \_\_\_ 5. In Windows Explorer, open the export folder: C:\Datacap\TravelDocs\export  
The export task outputs the data in a TXT file and an XML file formats.
- \_\_\_ 6. Open the <today's date.00000N>. txt file and verify that the file contains the extracted values from the image as comma-separated values.

```

1 *****
2 Export for batch #20200308.000004,03/08/2020,09:55:59
3 ,,,Mon, Oct 4, 2010,New York (JFK),Fri, Oct 8,
2010,New York (JFK),Full size,Navigation System
Fuel Service,582.77

```

- \_\_\_ 7. Close the file.
- \_\_\_ 8. Check the data export XML file.
  - \_\_\_ a. In the same folder, double-click the <today's date.00000N>. xml file to open it in Notepad++.
  - \_\_\_ b. Verify that the page displays the xml content: Document name (**Car\_Rentals**), page name (**Rental\_Agreements**), and the extracted field values

```

<?xml version='1.0' ?>
<BatchID_20200308.000004>
  <Flights/>
  <Hotels>
    <Other_Charges/>
  </Hotels>
  <Car_Rentals>
    <Rental Agreements>
      <TM000001>
        <Pickup_Date>Mon, Oct 4, 2010</Pickup_Date>
        <Pickup_Location>New York (JFK)</Pickup_Location>
        <Return_Date>Fri, Oct 8, 2010</Return_Date>
        <Return_Location>New York (JFK)</Return_Location>
        <Car_Type>Full size</Car_Type>
        <Options>Navigation System Fuel Service</Options>
        <Total_Cost>582.77</Total_Cost>
      </TM000001>
    </Rental Agreements>
  </Car_Rentals>
</BatchID_20200308.000004>

```

- \_\_\_ 9. Close the file and Notepad++.

## Section 3. Explore Datacap Navigator

In this section, you become familiar with the capabilities of the Datacap Navigator client.

### 3.1. Start Datacap Navigator for business users

- 1. Open the Mozilla Firefox browser and click the **Datacap Navigator** in the Bookmarks toolbar.

You can also enter the following URL:

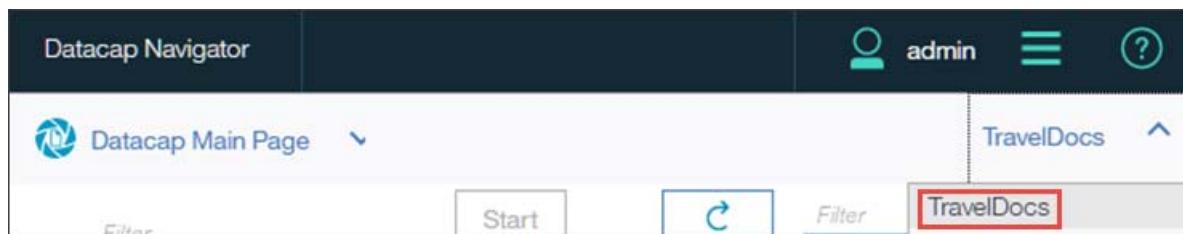
`http://ecmedu01:9081/navigator/?desktop=datacap`

- 2. Enter the following values:
  - **User name:** admin
  - **Password:** admin
- 3. Click **Log in**.

The Datacap view opens.

### 3.2. View Job Monitor

- 1. On the upper right of the page, verify that the **TravelDocs** application is selected from the list.



- 2. On the left pane, observe the task shortcuts for **TravelDocs**.

The middle pane is the Job Monitor. A list of batches for the selected application is shown.

Queue ID	Batch	Job	Task	Status
1	20200226.000000	Main Job	Verify	pending
2	20200228.000000	Navigator Job	Verify	hold
8	20200308.000003	Main Job	Verify	pending
9	20200308.000004	Main Job	Export	Job done

**Note**

The list of tasks on your student image might be different from the one shown in the screen capture. It depends on the number of batches run.

- 
- \_\_\_ 3. Notice the batch that you completed in Datacap Desktop is listed with today's date, **Task** name **Export**, and **Status** name **Job Done**.

### 3.3. Check batch information

- \_\_\_ 1. In the Job Monitor middle pane, select (single click) a batch in the list.
- 

**Note**

If you double-click an item, it opens for processing.

- 
- \_\_\_ 2. On the right pane, expand **System Properties** (Click the blue triangle) to see information for the selected batch.

You can collapse or expand the Properties pane by clicking the three vertical dots. You can also click and hold the edge of the panes, and drag to resize the width.

- \_\_\_ 3. Click a column header to sort by that column.
- 

**Note**

In IBM Content Navigator administration desktop, you configure to hide or show the columns and column order for each application.

- 
- \_\_\_ 4. View the batch history.

- \_\_\_ a. In the Job Monitor list, select (single click) a batch and click **View History** from the menu bar.
- \_\_\_ b. Batch history opens in a separate window.

The history details depend on the task step of the batch.

Batch History					
Task	Station	Operator	Start Time	Run Time	
Virtual Scan	1	admin	3/8/2020, 8:51:10 AM	85	
Page ID	1	admin	3/8/2020, 8:58:44 AM	9	
Batch Profiler	1	admin	3/8/2020, 9:11:09 AM	7	
Verify	1	admin	3/8/2020, 9:25:18 AM	1135	
Export	1	admin	3/8/2020, 9:55:59 AM	9	

- \_\_\_ c. Click **Close** to close the window.
- \_\_\_ 5. Filter the job list with the filter criteria.
  - \_\_\_ a. In the **Task** column, choose a value and on the upper right of the page, in the **Filter** field, type that value: **Export**



### Important

The value that you enter must be available in the **Task** column.

- \_\_\_ b. Check that only the items with the specified task are shown.

Start	Edit Job	Edit Batch	Delete	View History		Export	X	
3 of 8 items shown.		Clear filter						
Queue ID	Batch	Job	Task	Status	Job Start	Job Tid		
9	20200308.000004	Main Job	Export	Job done	3/8/2020, 8:51 AM	9		
11	20200310.000001	Navigator Job	Export	Job done	3/10/2020, 4:25 AM	1		
16	20200602.000002	Main Job	Export	Job done	6/2/2020, 3:52 PM	4		

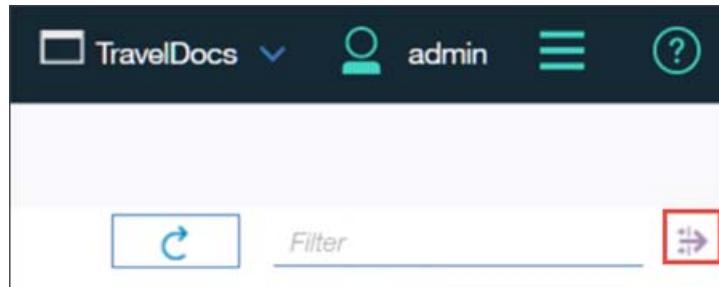
- \_\_\_ c. Optionally, test another criteria for example, a value from the **Status** column (Example: **Pending**)

- \_\_\_ d. On the line below the toolbar, click **Clear filter** to go back to the complete list.  
 You can also click X on the filter field.



### Note

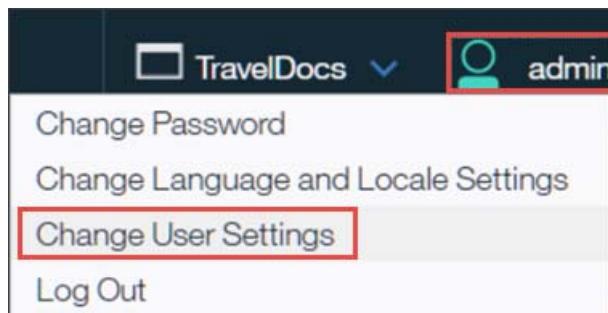
You can do an advanced filtering by clicking the arrow icon next to the Filter field.



## 3.4. Explore the User Settings

Business users can change some of the settings.

- \_\_\_ 1. On the banner, click **admin** and select **Change User Settings** from the list.



- \_\_\_ 2. In the **Settings > Global** tab, observe that you can configure to hide the following features.
- Title bar for all the widgets
  - Shortcut pane when you start the task
- \_\_\_ 3. Observe the tabs for each task such as **Scan**, **Upload**, **Classify**, and **Verify**.  
 You learn about the user settings in detail in a different course.
- \_\_\_ 4. Click **Cancel** to close the window.
- \_\_\_ 5. Leave Datacap Navigator open for the next section.

## Section 4. Process batches in Datacap Navigator

In this section, you process a batch in the Datacap Navigator client and check the output.

### 4.1. Scan batches

Datacap Navigator is open from the previous section and the **TravelDocs** application is selected.

- \_\_\_ 1. Scan the **Flight1.tif** image from the C:\Datacap\TravelDocs\images folder.
  - \_\_\_ a. On the left pane, click **Navigator scan**.
  - \_\_\_ b. On the **Scan with Datacap Navigator client** tab, click **Browse**.
  - \_\_\_ c. In the **Open** window, browse to the C:\Datacap\TravelDocs\images folder.
  - \_\_\_ d. Select **Flight1.tif** and click **Open**.

On the **Scan with Datacap Navigator client** tab, next to **Browse**, the file location is shown.

  - \_\_\_ e. On the upper right, click **Scan**.
2. On the right, in the **Batch Structure** pane, verify that the image ID (**TM000001**) is listed and on the middle pane, in the image viewer, the scanned image (**Airline #1**) is shown.
3. Scan another image.
  - \_\_\_ a. Click **Browse** and browse to the C:\Datacap\TravelDocs\images folder.
  - \_\_\_ b. Select **Car1.tif** and click **Open**.
  - \_\_\_ c. Click **Scan** and observe that you have different options in the list (**Append**, **Insert**, or **Replace**).

You can replace the image, or scan more images to add to the batch.

ID	Type
20200602.000003	Travel Documents
TM000001	Other

- \_\_\_ d. Click **Append** and verify that the image is added.

**Scanned Pages(2/2)** indicate that two pages are scanned.

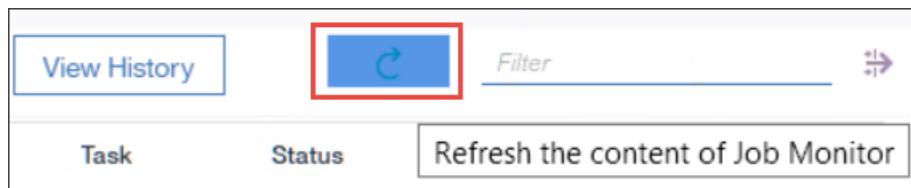


- \_\_\_ 4. At the top of the page, click **Submit**.

The scanned pages are uploaded automatically to the Datacap Server.

- \_\_\_ 5. Refresh the list.

- \_\_\_ a. On the main page, click the **Refresh** icon to refresh the contents of Job Monitor.



- \_\_\_ b. If the complete batch list is shown, on the toolbar, click **Clear filter**.
- \_\_\_ 6. Verify that your batch is listed with today's date, **Job name Navigator Job**, **Task name Page ID**, and the **Status is Pending**.

## 4.2. Review the Scan task output

In this section, you review the files. The output is similar to the ones that you processed in Datacap Desktop in an earlier section.

- \_\_\_ 1. Open Windows Explorer and browse to the batches folder:  
C:\Datacap\TravelDocs\batches
- \_\_\_ 2. Verify that the batch that you scanned created a folder with a name that has today's date (latest) in this format: <date.00000N> where N is a number.  
Example: **20200310.000001**  
Since you processed more than one batch today, look for the latest one.
- \_\_\_ 3. Check the nscan.xml file for the <date.00000N> batch.
  - \_\_\_ a. Open the <date.00000N> batch folder.
  - \_\_\_ b. Double-click the **nscan.xml** file to open it in Notepad++.
  - \_\_\_ c. Verify that the pages you scanned have the following information.
    - The batch **Type** is **TravelDocs**
    - The page **type** is **Other** for both entries (**PageName: Flight1** and **Car1**)
    - Values for the two **IMAGEFILE** fields: <Examples: tm000001.tif and tm000002.tif>
  - \_\_\_ d. Close the file and leave the Windows Explorer folder open for rest of this section.

## 4.3. Start Datacap Rulerunner Service

The rulerunner runs the background tasks and it must be running for the batch to progress to Verify task.

- \_\_\_ 1. From the Windows Start menu, click **IBM Datacap Services > Datacap Rulerunner Manager**.

You can also use the **Datacap Rulerunner Manager** shortcut on the Windows desktop.

- \_\_\_ 2. In the **Datacap Rulerunner Manager** window, select the **Rulerunner** tab.
- \_\_\_ 3. The **Status** shows that the service is **Stopped**.
- \_\_\_ 4. Click **Start** to start the server and wait for the status to show **Running**.
- \_\_\_ 5. Click **Close** at the bottom of the page.

## 4.4. Run the PageID and Profiler tasks

The **PageID** and **Profiler** tasks complete the page identification, document assembly, and data extraction. Rulerunner is configured to run the **PageID**, **Profiler**, and **Export** tasks automatically. When you started Rulerunner in the previous section, the **PageID** and **Profiler** tasks run automatically.

- \_\_\_ 1. On the Datacap Navigator main page, click the **Refresh** icon to refresh the contents of Job Monitor.
- \_\_\_ 2. Verify that your batch (latest **Queue ID**) is now at the **Verify** task.
- \_\_\_ 3. Leave the Datacap Navigator open for the next section.
- \_\_\_ 4. In Windows Explorer, in the same batch folder as in the previous section, verify that the following two files are created.
  - PageID.xml
  - Profiler.xml
- \_\_\_ 5. Check the PageID.xml file for the <date.00000N> batch.
  - \_\_\_ a. In Windows Explorer, double-click the **PageID.xml** file to open it in NotePad++.
  - \_\_\_ b. Verify that the page type is identified as **Air\_Ticket** for the **Flight1** page and **Rental\_Agreement** for the **Car1** page.
  - \_\_\_ c. Close the file.
- \_\_\_ 6. Check the Profiler.xml file for the <date.00000N> batch.
  - \_\_\_ a. In Windows Explorer, double-click the **Profiler.xml** file to open it in Notepad++.
  - \_\_\_ b. Verify that the document type is identified as **Flight** for the **Air\_Ticket** page type and **Car\_Rental** for the **Rental\_Agreement** page type.
  - \_\_\_ c. Close the file.

## 4.5. Run the Verify task

The task in this section completes the data verification.

The Datacap Navigator Monitor view is open from the previous section and your batch is at the **Verify** task in the **pending** state.

- 1. In the Datacap Navigator Monitor view, double-click your batch (latest **Queue ID**) to open it.
- 2. In the **Verify** tab, check the field values against the image and click **Submit** two times to complete the two pages.
- 3. Click **OK** on the message dialog box: All documents are complete. Do you want to finish batch?
- 4. Leave the Datacap Navigator open for the next section.



### Note

To run the **Verify** task for all the pending batches, click the **Verify** shortcut from the Task shortcut list. The user can configure the Datacap Navigator to automatically open the next pending batch in the queue when you submit a current batch.



### Troubleshooting

If no exceptions found, the batch advances to the **Verify** task after the **PageID** and **Profiler** tasks.

If any exceptions were found in the pages that are scanned, the batch item moves to **NFixUp** step.



### Information

Occasionally, after you run the Verify step, you want to rerun the PageID or Profiler task again on the same batch to ensure the page identification and data extraction tasks were completed correctly.

In Datacap Navigator, you can configure the batch to run a task that was already run as described in the following steps:

- Select the batch that you want to edit and in the toolbar, click **Edit Job**.

	Start	Edit Job	Edit Batch	Delete	View History
Queue ID	Batch	Job	Task	Status	
14	20200515.000000	Navigator Job	Verify	pending	
11	20200310.000001	Navigator Job	Export	Job done	

- In the **Edit Job Properties** dialog box, select the task that you want to rerun (Example: **Page ID**) from the list.

- For the **Status** field, select **pending** from the list.

**Edit Job Properties**

Batch ID:	20200515.000000
Batch directory:	C:\Datacap\TravelDocs\batches\20200515.000000
Page file:	Profiler.xml
Task ID:	Page ID
Status:	pending
Operator:	<none>
Station:	<none>
Priority:	5
Stored operator:	<none>
Stored station:	1
<input type="button" value="Apply"/> <input type="button" value="Cancel"/>	

- Click **Apply** and verify that the **Task** column for your batch now shows **Page ID** as you configured.

Start	Edit Job	Edit Batch	Delete	View History	↻
Queue ID	Batch	Job	Task	Status	
14	20200515.000000	Navigator Job	Page ID	pending	
11	20200310.000001	Navigator Job	Export	Job done	

You can now process your batch again from the **Page ID** task. It moves through **Profiler** and **Verify** steps.

## 4.6. Review the Verify task output

- 1. In Windows Explorer, in the same batch folder as in the previous section, right-click **Verify.xml** and select **Open with > Internet Explorer** to open it in a browser.
- 2. If you are prompted to allow the scripts to run, click **Yes** two times.
- 3. For each document (**Flight** and **Car\_Rental**), verify that the **DATAFILE** field has a value: **tm00000n.xml**  
Example: **tm000001.xml**
- 4. Check the DATAFILE file for each page.
  - a. For the **Flight** document, click the **DATAFILE** file link **tm00000n.xml** to open it in the browser.
  - b. If you are prompted to allow the scripts to run, click **Yes** two times.
  - c. Check that the fields (**Outbound\_To: San Francisco (SFO)**) have valid data.
  - d. Close the browser tab.
  - e. Leave the Windows Explorer open for the next section.

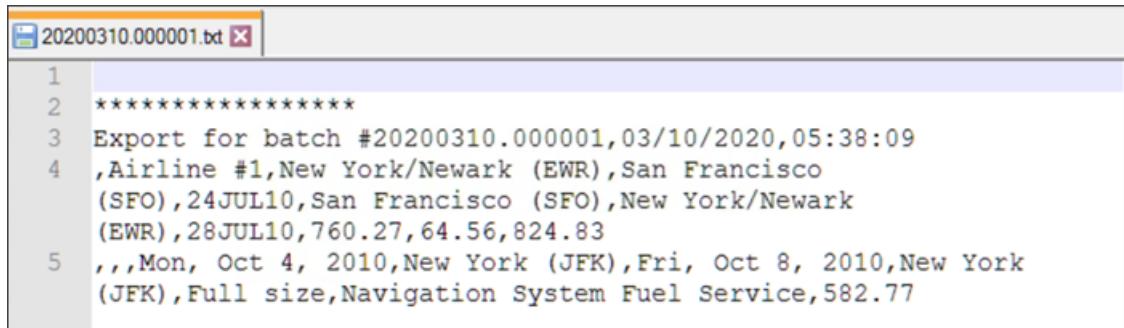
## 4.7. Check your batch status

After you complete Verify task, your batch moves to the Export task. The Export task runs automatically, and the batch is completed.

- 1. On the Datacap Navigator Job Monitor view, click the **Refresh** icon and verify that your batch (latest **Queue ID**) is now at the **Export** task with the following values.
  - The **Job** column: **Navigator Job**
  - The **Task** column: **Export**
  - The **Status** column: **Job done**
- 2. Log out of Datacap Navigator.
  - a. On the banner, click the username (**admin**) and select **Log Out** from the list.
  - b. Click **Log Out**.
  - c. Close the browser.

## 4.8. Review the Export task output

- 1. In Windows Explorer, in the same batch folder as in the previous section, verify that the **Export.xml** file is created.
- 2. Check the data export text file.
  - a. In Windows Explorer, open the export folder: **C:\Datacap\TravelDocs\export**
  - b. Open the **<today's date.00000N>.txt** file (latest) and verify that the file contains the extracted values from the images as comma-separated values.



The screenshot shows a Windows Notepad window with the title bar '20200310.000001.txt'. The content of the window is a log file with the following text:

```
1 ****  
2 *****  
3 Export for batch #20200310.000001,03/10/2020,05:38:09  
4 ,Airline #1,New York/Newark (EWR),San Francisco  
(SFO),24JUL10,San Francisco (SFO),New York/Newark  
(EWR),28JUL10,760.27,64.56,824.83  
5 ,,,Mon, Oct 4, 2010,New York (JFK),Fri, Oct 8, 2010,New York  
(JFK),Full size,Navigation System Fuel Service,582.77
```

- \_\_\_ 3. Close the file, Windows Explorer, and the browser.

#### 4.9. Stop Datacap Rulerunner Service

- \_\_\_ 1. Double-click the **Datacap Rulerunner Manager** shortcut on the Windows desktop.
- \_\_\_ 2. In the **Rulerunner** tab, verify that the status shows **Running** and click **Stop**.
- \_\_\_ 3. Wait for the service to stop and click **Close** at the bottom of the page.

**End of exercise**

## Exercise review and wrap-up

This exercise showed how to process the batches for the Datacap applications by running the workflow in the Datacap clients.

---

# Exercise 2. Building a Datacap application with Forms Template

## Estimated time

02:00

## Overview

In this exercise, you build a Datacap application with Forms Template in Datacap Studio.

## Objectives

After completing this exercise, you should be able to:

- Create an application with Forms Template
- Configure the application in IBM Content Navigator
- Edit the application settings in Datacap Application Manager
- Define document hierarchy
- Configure fingerprints and zones

## Introduction

Datacap provides templates for application development. For example, Forms and Learning templates.

In this exercise, you create a Datacap application with Forms Template in Datacap Studio.

This exercise includes the following sections:

- [Section 1, "Create a Datacap application"](#)
- [Section 2, "Configure the application in IBM Content Navigator"](#)
- [Section 3, "Edit the application settings in Datacap Application Manager"](#)
- [Section 4, "Edit the task profiles and rulesets"](#)
- [Section 5, "Define document hierarchy"](#)
- [Section 6, "Configure fingerprints and zones"](#)

## Requirements

None

## Section 1. Create a Datacap application

To make the Datacap application development faster, you can copy an existing application and customize it. In this exercise, you build a new application in Datacap Studio to understand different components that makes up an application.

In the following exercises, you reuse the application that you create in this exercise to build other applications.

In the previous unit, you learned about the application design. You use those design principles in this exercise to configure the application.



### Note

A solution of the Datacap application for this exercise is provided in C:\Datacap folder with the application name: **SolutionClaimForm**

You can access this application like any other Datacap application.

### 1.1. Create a Datacap application

- \_\_\_ 1. Before you begin the exercise, ensure that the components are running. If you did not start the components already, refer to [Section 1, "Prepare your system: Start the components"](#) to start them.

If you restart your system, you must start all the components.

- \_\_\_ 2. Log in to Datacap Studio.

- \_\_\_ a. Click **Start > IBM Datacap Developer Tools > Datacap Studio**

You can also use the **Datacap Studio** shortcut on the Windows desktop.

- \_\_\_ b. In the **Applications** dialog box, select **TravelDocs** and click **Next**.

- \_\_\_ c. On the **Taskmaster Login** page, enter the following values:

- **User ID:** admin

- **Password:** admin

- **Station ID:** 1

- \_\_\_ d. Click **Finish**.

The **TravelDocs** application is open in Datacap Studio.

**Hint**

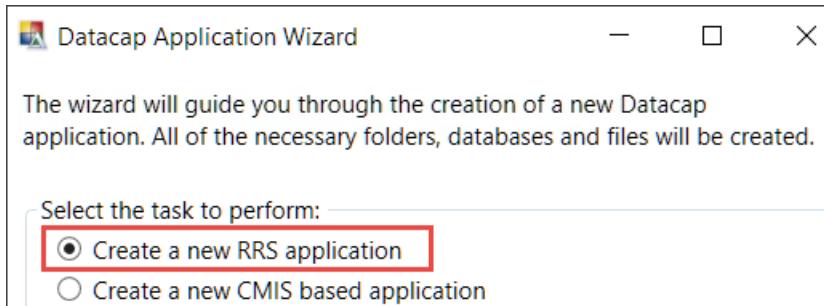
To create a new application, selection of any application or login is not required. But the login ensures that all the Datacap components are started and the system is in a running state.

To open Datacap Studio without opening an application:

- \_\_\_ a. Click **Start > IBM Datacap Developer Tools > Datacap Studio** or use the **Datacap Studio** shortcut on the Windows desktop.
- \_\_\_ b. In the **Applications** dialog box, click **Close**.

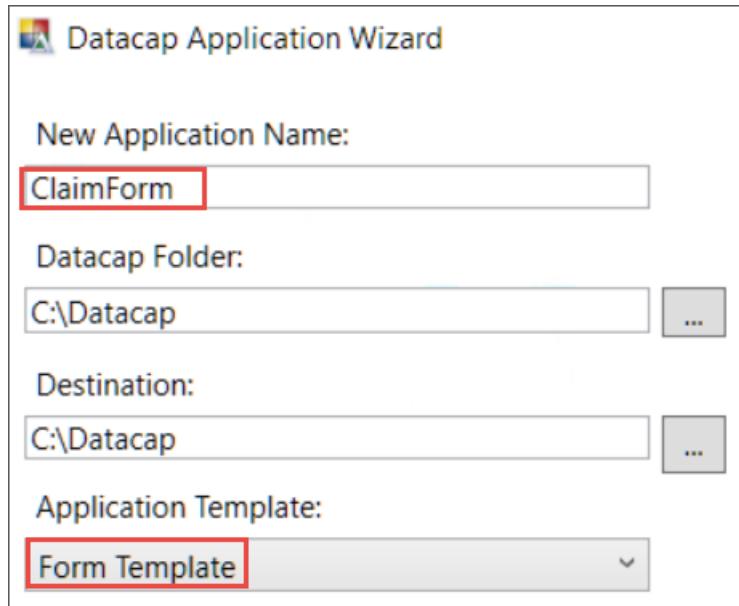
Datacap Studio opens without any application.

- \_\_\_ 3. On the upper right of the toolbar, click the **Datacap application wizard**  icon.
- \_\_\_ 4. In **Datacap Application Wizard**, select the **Create a new RRS application** option.

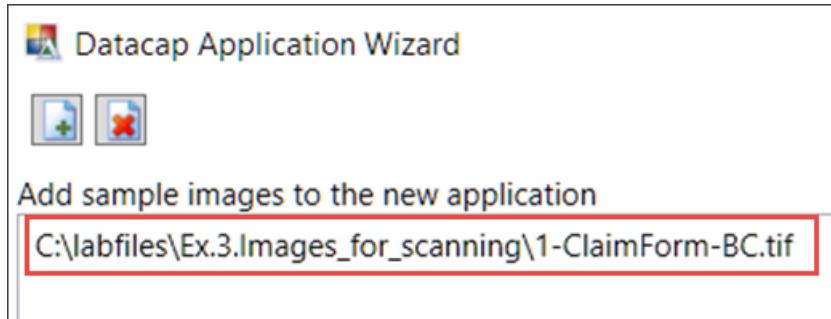


- \_\_\_ 5. Enter a name and select a template for the new application.
  - \_\_\_ a. Click **Next**.
  - \_\_\_ b. For the **New Application Name** field, enter: **ClaimForm**
  - \_\_\_ c. For the **Application Template** field, if it is not already selected, select **Form Template** from the list.

Leave the default folder values: C:\Datacap



- \_\_\_ 6. Click **Next** two times.
- \_\_\_ 7. Add sample images to the new application.
  - \_\_\_ a. Click the plus symbol.
  - \_\_\_ b. In the **Open** window, browse to the C:\labfiles\Ex.3.Images\_for\_scanning folder, select the 1-ClaimForm-BC.tif file, and click **Open**.



- \_\_\_ 8. Click **Next** and click **Finish**.
- \_\_\_ 9. Review the Application wizard summary for the list of folders, files, and databases and click **Close** to close the wizard.



- \_\_\_ 10. Click **Exit** at the upper right corner to close Datacap Studio.

## 1.2. Check the Datacap.xml file and application folder

- \_\_\_ 1. Open Windows Explorer, browse to the C:\Datacap folder and open the **Datacap.xml** file in NotePad++.
- \_\_\_ 2. Verify that an entry is added for the new application (**ClaimForm**).

```

1  <?xml version="1.0" encoding="UTF-8" standalone="yes"?>
2  <datacap ver="9.0">
3      <app name="TravelDocs" ref="TravelDocs"/>
4      <app name="DataExchange" ref="C:\Datacap\DataExchange"/>
5      <app name="SolutionClaimForm" ref="C:\Datacap\SolutionClaimForm"/>
6      <app name="SolutionMultiPageApp" ref="C:\Datacap\SolutionMultiPageApp"/>
7      <app name="SolutionDataExtraction" ref="C:\Datacap\SolutionDataExtraction"/>
8      <app name="FormTemplate" ref="Templates\FormTemplate"/>
9      <app name="LearningTemplate" ref="Templates\LearningTemplate"/>
10     <app name="Flex" ref="Flex"/>
11     <app name="APT" ref="APT"/>
12     <app name="Medical Claims" ref="Medical Claims"/>
13     <app name="ClaimForm" ref="C:\Datacap\ClaimForm"/>
14 </datacap>

```

For every new application, Datacap adds an entry with the new application name in the datacap.xml file. The XML file is initially created when Datacap is installed.

- \_\_\_ 3. Close the file and close NotePad++.
- \_\_\_ 4. In Windows Explorer, in the C:\Datacap folder, verify that a folder with the name of the new application (**ClaimForm**) is created, and open the folder.

The following subfolders and files are created for the new application:

### Folders:

- **batches** where information for the batches that you run are stored
- **dco\_ClaimForm** where configuration details for the application are stored
- **export** where the extracted data that is exported
- **fingerprint** where fingerprints are stored for the application
- **images** where images to be scanned are stored

**Files:**

- ClaimForm.app - application information
  - Files with the .MDB extensions (database files)
- 



**Note**

You learn about some of these the folders and files in the following sections.

---

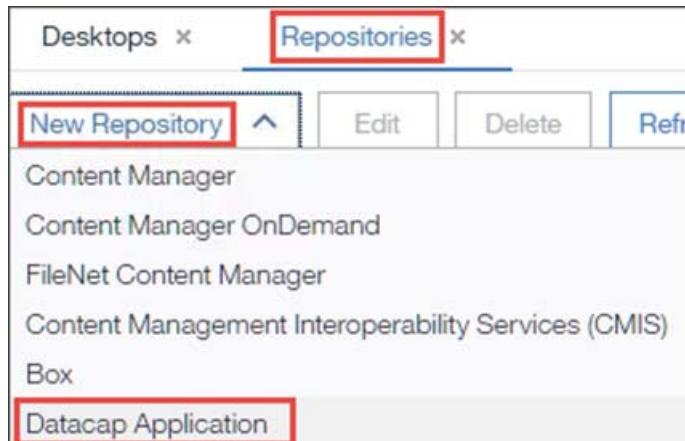
## Section 2. Configure the application in IBM Content Navigator

In this section, you configure the new application as a repository in Datacap Navigator administration tool. You associate the repository with the desktop so that the new application is accessible in the Datacap Navigator clients.

You also edit the workflow for the new application in Datacap Navigator.

### 2.1. Add the new Datacap application as a repository

- 1. Start the IBM Content Navigator administration tool.
  - a. Open the Firefox browser and click the **ICN Admin** shortcut.  
You can also enter the **URL**: `http://ecmedu01:9081/navigator/?desktop=admin`
  - b. Enter the following values:
    - **User name**: p8admin
    - **Password**: FileNet1
  - c. Click **Log In**.
- 2. Open the New Repository tab.
  - a. On the left pane, click **Repositories**.  
On the **Repositories** tab, a list of the repositories that are configured is shown.
  - b. Click **New Repository** and from the list, select **Datacap Application**.



- 3. On the **New Repository** tab > **General** subtab, set the values for the following properties:

Field	Value
Display name	ClaimForm
ID	<b>ClaimForm</b>
Datacap wTM URI	<code>http://ecmedu01:85/ServicewTM.svc</code>
Application	<b>ClaimForm</b>

Default Station	<b>1</b>
Use ActiveX in IE	<b>No</b>
Use Virtual Viewer	<b>Yes</b>

**Note**

Details for the fields:

- For **Display name**, you can provide any name, but use of the Datacap application name helps to track the application associated with the repository.
- The **ID** field is automatically populated based on the Display name field entry. The ID value can be edited.
- For **Datacap wTM URI**, The value `ecmedu01` is the server name where wTM is installed.
- For **Application**, select the value from the list. The list contains all the available applications as listed in Datacap.xml file.

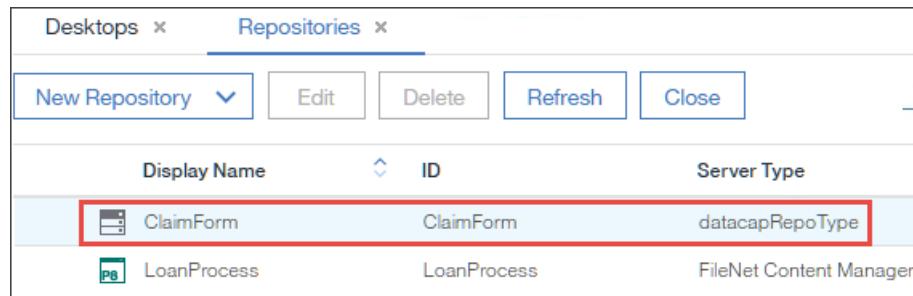
General
Configuration Parameters

---

You must connect to the repository before you configure parameters, system properties, folders, and so on.

* Display name: <span>(i)</span>	ClaimForm
* ID: <span>(i)</span>	ClaimForm
* Datacap wTM URI <span>(i)</span>	<a href="http://ecmedu01:85/ServicewTM.svc">http://ecmedu01:85/ServicewTM.svc</a>
* Application <span>(i)</span>	<input style="border: 1px solid #ccc; padding: 5px; width: 150px; height: 25px;" type="button" value="ClaimForm"/> <span style="font-size: small;">▼</span>
* Default Station <span>(i)</span>	1
Use ActiveX in IE <span>(i)</span>	<input type="radio"/> Yes <input checked="" type="radio"/> No
Use Virtual Viewer <span>(i)</span>	<input checked="" type="radio"/> Yes <input type="radio"/> No
<input style="border: 1px solid #0070C0; background-color: #0070C0; color: white; padding: 5px; width: 150px; height: 25px;" type="button" value="Connect..."/>	

- \_\_\_ 4. Test the repository connection.
  - \_\_\_ a. Click **Connect**.
  - \_\_\_ b. On the **Log In** page, enter the credentials for a user that has administration access to the repository.
    - **User name:** admin
    - **Password:** admin
  - \_\_\_ c. Click **Log in** and verify that the login is successful.
  - \_\_\_ d. Click the **Configuration Parameters** subtab which is now active.
- \_\_\_ 5. Click **Save and Close** to save the settings and close the new repository tab.  
If **Save and Close** is not enabled, click the **New Repository** tab first and then click **Save and Close**.
- \_\_\_ 6. On the **Repositories** tab, verify that the new repository (**ClaimForm**) is listed.



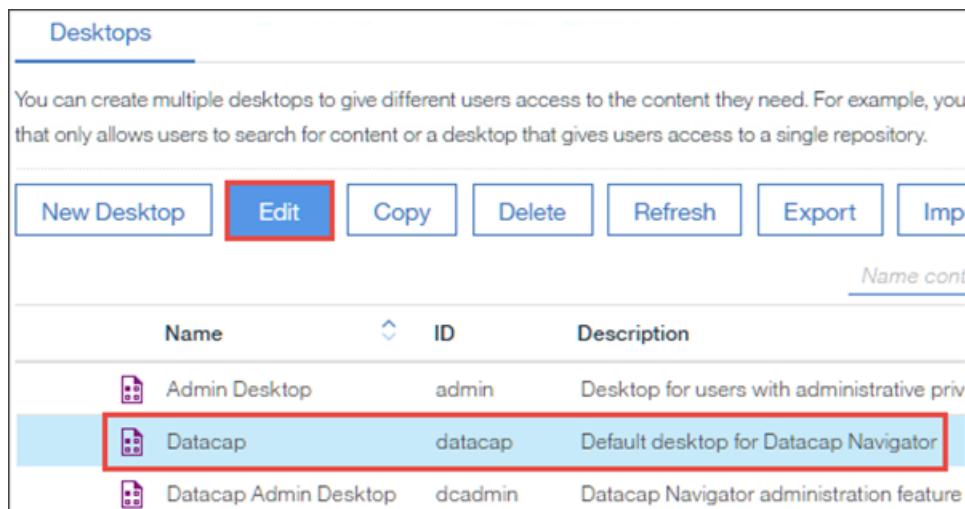
A screenshot of a software interface titled "Repositories". At the top, there are buttons for "New Repository", "Edit", "Delete", "Refresh", and "Close". Below this is a table with columns: "Display Name", "ID", and "Server Type". There are two rows of data. The first row has a thumbnail icon, the display name "ClaimForm", the ID "ClaimForm", and the server type "datacapRepoType". This row is highlighted with a red border. The second row has a thumbnail icon, the display name "LoanProcess", the ID "LoanProcess", and the server type "FileNet Content Manager".

Display Name	ID	Server Type
 ClaimForm	ClaimForm	datacapRepoType
 LoanProcess	LoanProcess	FileNet Content Manager

- \_\_\_ 7. Close the **Repositories** tab.

## 2.2. Associate the new repository with the Datacap desktops

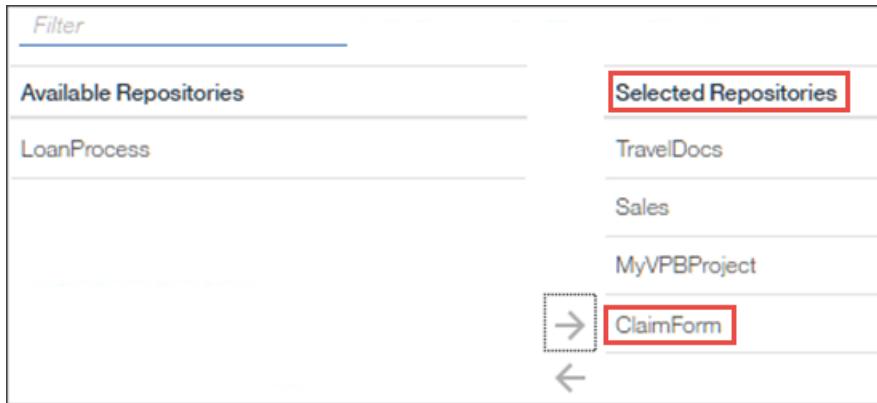
- \_\_\_ 1. In the **Desktops** tab, select the **Datacap** desktop from the list and click **Edit**.



A screenshot of a software interface titled "Desktops". At the top, there are buttons for "New Desktop", "Edit", "Copy", "Delete", "Refresh", "Export", and "Import". Below this is a table with columns: "Name", "ID", and "Description". There are three rows of data. The first row has a thumbnail icon, the name "Admin Desktop", the ID "admin", and the description "Desktop for users with administrative priv...". The third row has a thumbnail icon, the name "Datacap", the ID "datacap", and the description "Default desktop for Datacap Navigator". This row is highlighted with a red border. The second row has a thumbnail icon, the name "Datacap Admin Desktop", the ID "dadmin", and the description "Datacap Navigator administration feature".

Name	ID	Description
 Admin Desktop	admin	Desktop for users with administrative priv...
 Datacap	datacap	Default desktop for Datacap Navigator
 Datacap Admin Desktop	dadmin	Datacap Navigator administration feature

- \_\_\_ 2. On the **Datacap** tab, select the **Repositories** tab.
- \_\_\_ 3. From the **Available Repositories** pane, select **ClaimForm** and click forward arrow to move it to the **Selected Repositories** pane.



- \_\_\_ 4. Click **Save and Close** to save the settings and close the **Datacap** tab.
- \_\_\_ 5. Associate the new repository with the Datacap admin desktop.

The steps are similar to the **Datacap** desktop that you completed.

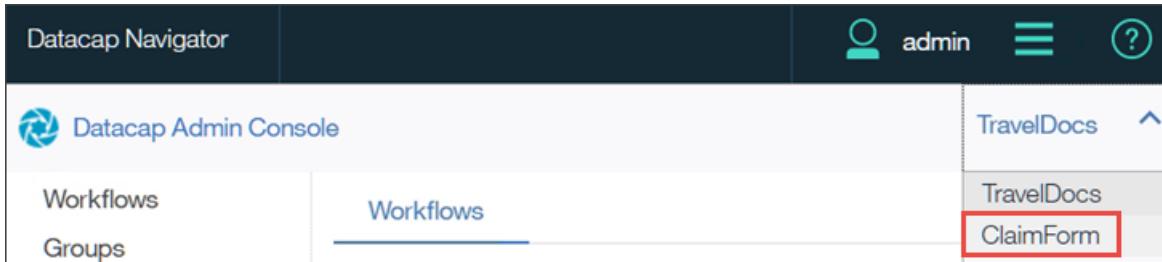
- \_\_\_ a. In the **Desktops** tab, select the **Datacap Admin Desktop** from the list and click **Edit**.
- \_\_\_ b. On the **Datacap Admin Desktop** tab, select the **Repositories** tab.
- \_\_\_ c. From the **Available Repositories** pane, select **ClaimForm** and click forward arrow to move it to **Selected Repositories** pane.
- \_\_\_ d. Click **Save and Close** to save the settings and close the tab.
- \_\_\_ 6. Log out of the IBM Content Navigator administration tool.

### 2.3. Edit the workflow for the new Datacap application

The Application wizard creates various types of jobs for the application. In this section, you review the jobs and delete the ones that you don't need. Deleting them is optional but helps keep the application only with the essential elements.

- \_\_\_ 1. Start the Datacap Administration console.
  - \_\_\_ a. In the Firefox browser, click the **DC Admin** shortcut.  
You can also use the **URL**: <http://ecmedu01:9081/navigator/?desktop=dcadmin>
  - \_\_\_ b. Enter the following values:
    - **User name**: admin
    - **Password**: admin
  - \_\_\_ c. Click **Log In**.

- \_\_ 2. In the Datacap Administration Console banner, open your application.
- \_\_ a. On the upper right of the page, click **TravelDocs** select **ClaimForm** from the applications list.
- The Datacap applications that are configured in Content Navigator admin tool are listed.



- \_\_ b. On the **Log In** page, enter the following values:
- **User name:** admin
  - **Password:** admin
- \_\_ c. Click **Log In**.
- \_\_ d. On the upper right of the page, verify that **ClaimForm** is selected.
- \_\_ 3. Check the Jobs that the wizard created and delete the ones that you don't need.
- \_\_ a. In the **Workflows** tab, select **ClaimForm** and click **Edit**.
- \_\_ b. In the **ClaimForm** tab, select **Jobs** subtab and view the jobs that are created by the application wizard.
- For this exercise, you keep the **Navigator Job** and **Demo\_SingleTIFFs** job.
- \_\_ 4. Delete the jobs that you don't need for your application.
- \_\_ a. On the **Jobs** subtab, select **Fixup Job**.
- \_\_ b. Click **Delete** and when you are prompted, click **Delete** again.
- \_\_ c. Repeat the steps to delete the jobs in the following list.
- **Fixup Job**
  - **Web Job**
  - **Demo\_MultiFormat**
  - **Verify\_Export**
  - **Manual Select**
- \_\_ 5. Rename the **Demo\_SingleTiFFs** Job as **Claim Form Job**.
- Renaming is optional. You can provide any name that is meaningful to your project.
- \_\_ a. Select **Demo\_SingleTiFFs** and click **Edit**.

- \_\_\_ b. On the **Demo\_SingleTIFFs** tab, **General** subtab, for the **Name** field, edit the value to: Claim Form Job

Job: **Demo\_SingleTIFFs**

General		Tasks
* Name <small>i</small>	Claim Form Job	
Description	Standard processing of single page images	
* Priority <small>i</small>	5	<small>v</small>

- \_\_\_ c. Click **Save** and verify that the tab is renamed as **Claim Form Job**.
- \_\_\_ 6. Edit the Profiler task.
- \_\_\_ a. On the **Claim Form Job** tab, select **Tasks** subtab and view the tasks that are available for this job.
- \_\_\_ b. Select **Profiler** and click **Edit**.
- \_\_\_ c. In the **Profiler** tab, **General** subtab, for the **Mode** field, select **Normal** from the list.

Task: **Profiler**

General		Advanced	Layout
* Name <small>i</small>	Profiler		
Description	Recognize/Validate w/Rules		
Mode <small>i</small>	Normal <small>Normal</small> Batch creation Router		
Queue by <small>i</small>			
Store <small>i</small>			

The step decides whether you want to route the workflow to Fixup step. For this application, router is not required.

- \_\_\_ d. Click **Save and Close** to save the settings and close the **Profiler** tab.

- \_\_ e. On the **Claim Form Job > Tasks** subtab, verify that the **Profiler** task shows the **Normal** mode.

Name	Description	Mode	Program	Queue By
Profiler	Recognize/Valid ate w/Rules	Normal	rulerunner.exe	None

- \_\_ 7. Add the Verify task.

- \_\_ a. On the **Claim Form Job > Tasks** subtab, click **New Task**.
- \_\_ b. In the **New Task > General** subtab, for the **Name** field, from the list, select **Verify**.
- \_\_ c. For the **Program** field, ensure that **Multiple** is selected.  
With this choice, you can run the Verify task on any available programs.
- \_\_ d. Click **Save** to save the new task.

The screenshot shows the Datacap application interface with the following details:

- Workflows**: Workflows x
- ClaimForm**: ClaimForm x
- \*Claim Form Job**: \*Claim Form Job x
- Verify**: Verify x

**Task: Verify**

**General** tab is selected.

**Name**: Verify

**Description**: Verify with Rule Validation

**Mode**: Normal

**Queue by**: None

**Store**: None

**Program**: Multiple

Buttons at the top: Save and Close, Save, Reset, Close.

- \_\_ e. Click **Close** to close the tab.
- \_\_ 8. Change the order the tasks.
- \_\_ a. In the **Claim Form Job > Tasks** subtab, select the **Verify** task that you created.

- \_\_\_ b. Click **Move up** to place **Verify** in between the **Profiler** and **Export** tasks.

Job: Claim Form Job					
General		Tasks			
New Task	Edit	Refresh	Delete	Move Up	Move Down
Name	Description	Mode	Program	Queue By	
➡ VScan	Run VScan Rules	Batch creation	rulerunner.exe	None	
➡ PageID	Page Identification Rules	Router	rulerunner.exe	None	
➡ Profiler	Recognize/Validate w/Rules	Normal	rulerunner.exe	None	
➡ Verify	Verify with Rule Validation	Normal	super.aspx	None	
➡ Export	Export via Rules	Normal	rulerunner.exe	None	

- \_\_\_ c. Click **Save** to save the changes to the Job.
- \_\_\_ 9. Edit the Verify task.
- \_\_\_ a. Select the **Verify** task and click **Edit**.
- \_\_\_ b. In the **Verify** tab, click the **Advanced** subtab.
- \_\_\_ c. Scroll down to the **Rulerunner settings** section and ensure that the **Main task profile** field is set to: **Verify**

▼ Rulerunner settings

Main task profile ⓘ	<input type="text" value="Verify"/>
Additional task profiles ⓘ	
Data preparation task profile ⓘ	

- \_\_\_ d. Scroll down to the **DCO tree view > Actions** section and clear the **Check Integrity** option.

Disassemble Documents ⓘ	<input checked="" type="checkbox"/>
Check Integrity ⓘ	<input type="checkbox"/>
Undo ⓘ	<input checked="" type="checkbox"/>

The Document Integrity setting is used when a document has multiple pages to arrange the pages in the correct order.

For this exercise, the document has only one page, so this setting is not required.

- \_\_\_ e. Scroll down to the **Display variables** section and for the **Props for var Type** field, change the value **0** to **1**.

The value **1** indicates that you can change the classification in **Verify** task.

The screenshot shows a configuration interface for a document. At the top, there is a checkbox labeled "Edit only for page" with an information icon. Below it is a section titled "Display variables" with a plus sign (+) and a minus sign (-) for adding or removing variables. Under this section, there are two entries: "Props for var TYPE" set to "1,setup" and "Props for var STATUS" set to "0". Both entries have their own information icons.

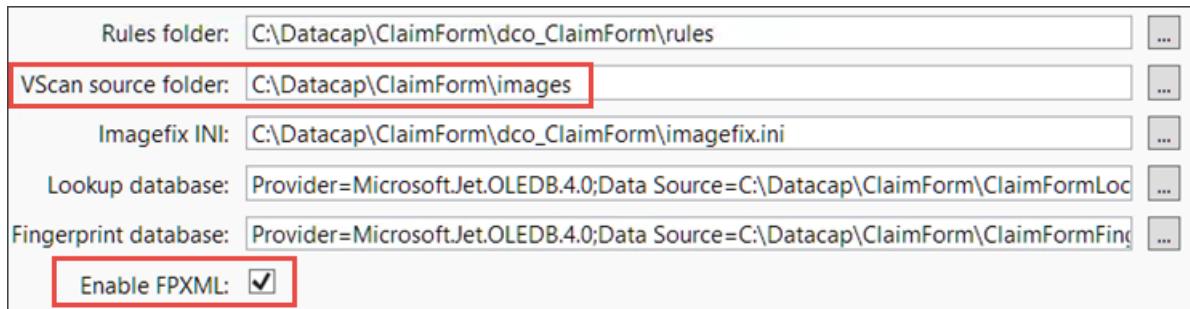
- \_\_\_ f. Click **Save and Close** to save the settings and close the **Verify** tab.  
\_\_\_ g. Click **Save and Close** on all other opened tabs.  
\_\_\_ 10. Log out of Datacap Admin Console.  
\_\_\_ a. On the banner, click **admin > Log Out**.  
\_\_\_ b. Click **Log Out** and close the browser.

## Section 3. Edit the application settings in Datacap Application Manager

When a new application is created, its settings are available in the Datacap Application Manager. In this section, you review and edit the initial application settings that the wizard created.

### 3.1. Check the Main tab

- 1. Click **Start > IBM Datacap Services > Datacap Application Manager**  
You can also use the **Datacap Application Manager** shortcut on the desktop.
- 2. On the **Datacap Application Manager** window, on the left pane, verify that **ClaimForm** is listed and click it.
- 3. On the **Main** tab on the right, verify that the following fields have the values and pointing to the C:\Datacap\ClaimForm folder:
  - **Administration, Engine, Lookup database**, and **Fingerprint database** fields
  - **Batch folder, Export folder, Fingerprint folder** fields
  - **Enable FPXML** (option selected)
- 4. For the **VScan source folder** field, change the value to: C:\Datacap\ClaimForm\images  
This change simplifies the folder structure.  
The image files that are used for scanning must be copied to this folder location. Datacap scans the images from this folder.

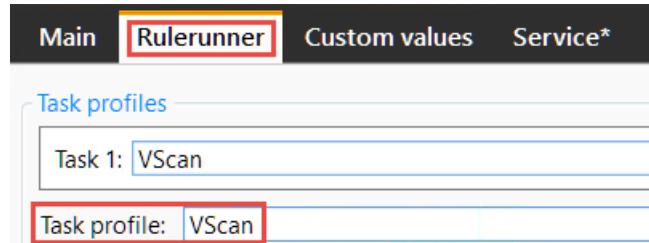


- 5. Click **Save changes** at the end of the page.

### 3.2. Check the Rulerunner tab

In this section, you check the Rulerunner tab and delete the tasks that are not required for the application. The tasks that are listed in this tab are run as background tasks by Rulerunner.

- 1. On the **Datacap Application Manager** window, click the **Rulerunner** tab.
- 2. For the **Task 1: VScan**, for the **Task profile**, edit the name to **VScan**.



- 3. For **Task 2: Convert**, click the **Remove this item** (red X) icon next to it and on the **Confirm deletion** dialog box, click **Yes**.
- 4. Delete **Task 6: NProfiler** and **Task 7: NPageID** and click **Save changes**.
- 5. Verify that the completed list has four tasks.

The Task numbering gets updated when you restart the Datacap Application Manager.

This screenshot shows the Datacap Application Manager with the Applications panel on the left and the Application settings panel on the right. The Applications panel lists several projects: TravelDocs, DataExtraction, MyVPBProjectSolution, ClaimForm, NewClaimForm, SolutionClaimForm, MultiPageApp, and MyInsightProjectSolution. The Application settings panel shows the Rulerunner tab selected. It displays a list of tasks under 'Task profiles': Task 1: VScan, Task 2: PageID, Task 3: Profiler, and Task 4: Export. Each task entry includes a 'Task profile:' field where the task name is repeated.

- 6. Close the Datacap Application Manager.



#### Note

In the Application Manager, for each task, a **Task** and **Task profile** (for example, two of VScan, two of PageID) are listed.

**Task 1: VScan** comes from database (shown in Datacap Navigator) and **Task profile: VScan** comes from collection.xml (shown in Datacap Studio). Application Manager maps both the fields.

## Section 4. Edit the task profiles and rulesets

Each task is linked to a task profile. The task profile defines the order of one or more rulesets to be run for that task. In this section, you configure task profiles for your application in Datacap Studio.

### 4.1. Configure task profiles for the new application

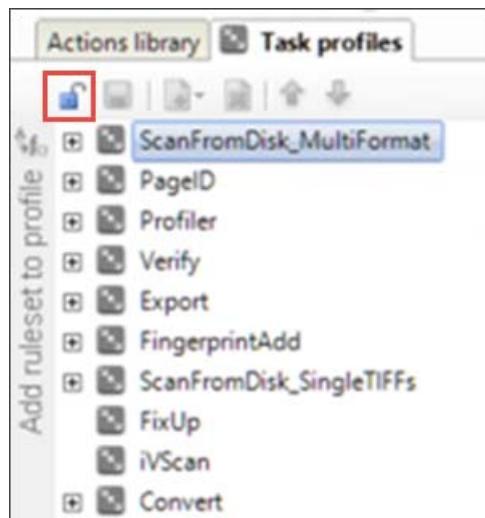
- \_\_\_ 1. Log in to Datacap Studio.
  - \_\_\_ a. Double-click the **Datacap Studio** shortcut on the Windows desktop.
  - \_\_\_ b. In the **Applications** dialog box, select **ClaimForm** and click **Next**.
  - \_\_\_ c. On the **Taskmaster Login** page, enter the following values:
    - **User ID:** admin
    - **Password:** admin
    - **Station ID:** 1
  - \_\_\_ d. Click **Finish**.
- \_\_\_ Datacap Studio opens with the **Rulemanager** tab selected.
- \_\_\_ 2. Delete the default task profiles that the application wizard created.



#### Note

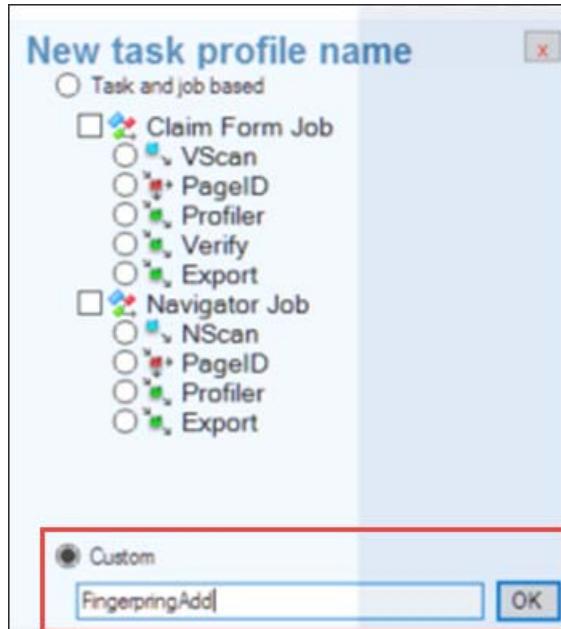
The application wizard creates several default task profiles. To clean up, you can delete the ones that you don't need. For this exercise, you delete all the default ones and add the ones that are required for this application.

- \_\_\_ a. On the upper right pane, click the **Task profiles** tab and lock it by clicking the lock icon.

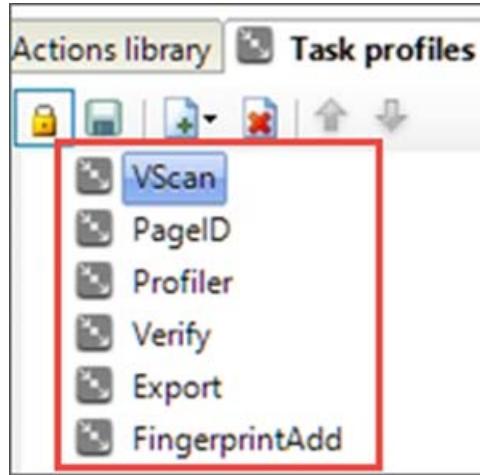


- \_\_\_ b. Delete all the existing task profiles by selecting a profile and clicking the Remove icon from the toolbar.

- \_\_\_ 3. Add VScan task.
- \_\_\_ a. In the toolbar for the **Task profiles**, click the **Add a new task profile**  icon.
  - \_\_\_ b. In the **New task profile name** dialog box, in the **Claim Form Job** node, select **VScan**.
- 
- \_\_\_ c. Click **OK**.
- \_\_\_ d. In the **Task profiles** tab, verify that **VScan** is listed.
- \_\_\_ 4. Repeat step 3 to add the following tasks from the **Claim Form Job** in the order, they are listed:
- **Page ID**
  - **Profiler**
  - **Verify**
  - **Export**
- \_\_\_ 5. Create the FingerprintAdd task profile.
- \_\_\_ a. Click the **Add a new task profile**  icon.
  - \_\_\_ b. In the **New task profile name** dialog box, select **Custom**, enter **FingerprintAdd**.



- \_\_\_ c. Click **OK**.
- \_\_\_ d. In the **Task profiles** tab, verify that six tasks are listed in the order as shown in the following screen capture:



- \_\_\_ 6. Click the **Save changes** icon on the **Task profiles** tab and unlock the **Task profiles** by clicking the lock icon.
- \_\_\_ 7. Leave the Datacap Studio open for the next section.

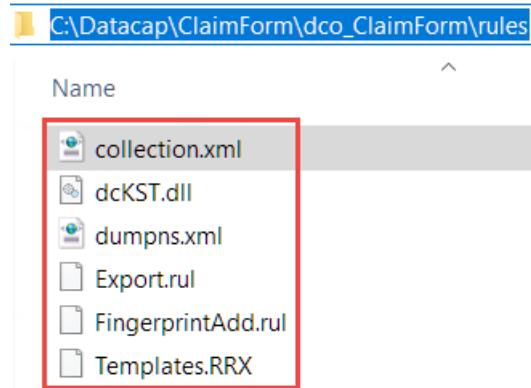
## 4.2. Check and clean the rulesets

The Application wizard creates several rulesets. In this section, you delete the rulesets that you don't need for your application and also delete the references. Deleting is optional, but it makes the application clean with only the required rulesets.

- \_\_\_ 1. In the Datacap Studio > **Rulemanager** tab > **Rulesets** pane, for the **ClaimForm** node, notice that some of the rulesets have an icon with a darker blue border.  
They are precompiled rulesets and are not [needed for the ClaimForm application. You delete the rulesets in the following steps.]
- \_\_\_ 2. Click **Exit** on the upper right corner and exit the Datacap Studio.  
A restart of Datacap Studio is required to refresh the changes after deletion.
- \_\_\_ 3. In Windows Explorer, delete the compiled rulesets that you don't need for the application.
  - \_\_\_ a. Open Windows Explorer and browse to the C:\Datacap\ClaimForm\dco\_ClaimForm\rules folder.  
Each compiled ruleset has a set of two files. For example, **CreateDocuments.Rul.dll** and **CreateDocuments.Rul.dll.config**
  - \_\_\_ b. Delete both the files.
  - \_\_\_ c. Delete the following files for the rulesets:
    - Document Integrity.rul
    - Export Statistics.rul

- IdentityPages.Rul.dll
- IdentityPages.Rul.dll.config
- Image\_Convert.Rul.dll
- Image\_Convert.Rul.dll.config
- ImageEnhancement.Rul.dll
- ImageEnhancement.Rul.dll.config
- Import files - all.rul
- ImportFiles.Rul.dll
- ImportFiles.Rul.dll.config
- Profile Statistics.rul
- RecognizePagesAndFields.Rul.dll
- RecognizePagesAndFields.Rul.dll.config
- Routing.rul
- ValidateFields.Rul.dll
- ValidateFields.Rul.dll.config

\_\_\_ d. Verify that the **rules** folder now contains a few key files as shown in the screen capture.



### Troubleshooting

If you accidentally deleted any of the required files, you can retrieve them from the `C:\labfiles\rules_Solution` folder.

- 
- \_\_\_ 4. Edit the `collection.xml` file to remove the reference for the compiled rulesets that you deleted.
- \_\_\_ a. In Windows Explorer, in the `C:\Datacap\ClaimForm\dco_ClaimForm\rules` folder, open the `collection.xml` file in Notepad++.

- \_\_\_ b. Observe the lines that refer to the rulesets that you deleted and notice that the task profiles are listed below the rulesets.
  - \_\_\_ c. Close the file.  
To avoid any XML format errors to the `collection.xml`, you copy and paste the solution file as described in the following steps:
  - \_\_\_ d. In Windows Explorer, in the `C:\Datacap\ClaimForm\dco_ClaimForm\rules` folder, delete the `collection.xml` file.
  - \_\_\_ e. Open another Windows Explorer window, browse to the `C:\labfiles` folder, copy the `collection_solution.xml` file, and paste it to the `C:\Datacap\ClaimForm\dco_ClaimForm\rules` folder.
  - \_\_\_ f. Rename the `collection_solution.xml` file to `collection.xml`.
- \_\_\_ 5. Open the file and verify that it contains the entries as shown in the screen capture.

```

<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<rsapp name="ClaimForm">
  <rsc>
    <ruleset name="Export" id="7"/>
    <ruleset name="FingerprintAdd" id="8"/>
  </rsc>
  <tps>
    <tprofile name="VScan"/>
    <tprofile name="PageID"/>
    <tprofile name="Profiler"/>
    <tprofile name="Verify"/>
    <tprofile name="Export"/>
    <tprofile name="FingerprintAdd"/>
  </tps>
</rsapp>

```

- \_\_\_ 6. Close the file.

## Section 5. Define document hierarchy

In this section, you configure Datacap objects (DCO) for your application in Datacap Studio.

### 5.1. Copy the images to be scanned to the new application folder

- 1. In Windows Explorer, browse to the C:\Datacap\ClaimForm\images folder and delete the two subfolders: `Input_MultiFormat` and `Input_SingleTIFFs`.
  - 2. From the C:\labfiles\Ex. 3.Images\_for\_scanning folder, copy the **1-ClaimForm-BC.tif** file and paste it to the C:\Datacap\ClaimForm\images folder.
- Recall that you simplified the folder structure and the defined the images folder in the Datacap Application Manager.
- 3. Open the **1-ClaimForm-BC.tif** file and notice the fields.

You use the information on this file and create the following Datacap objects (DCO) in the next section:

- Document type: **Insurance Docs**
- Page type: **Claim page**
- Fields:
  - o **Name**
  - o **Policy**
  - o **License**
  - o **Incident**

You can assign any text that is easy to identify to the names of the Datacap objects (DCO).

- 4. Open other images in the C:\labfiles\Ex. 3.Images\_for\_scanning folder and notice that the field layout is similar in all the files.

They all have the same layout with different set of values. You use these images for testing your application.

### 5.2. Create document hierarchy for the new application

- 1. Log in to Datacap Studio.
  - a. Click the **Datacap Studio** shortcut on the Windows desktop.
  - b. In the **Applications** dialog box, select **ClaimForm** and click **Next**.
  - c. On the login page, enter the following values:
    - o **User ID**: admin
    - o **Password**: admin
    - o **Station ID**: 1
  - d. Click **Finish**.

Datacap Studio opens with the **Rulemanager** tab selected.

- \_\_\_ 2. On the **Rulemanager** tab, on the upper left, in the **Document hierarchy** pane, click the **Lock DCO for editing**  icon to lock it.

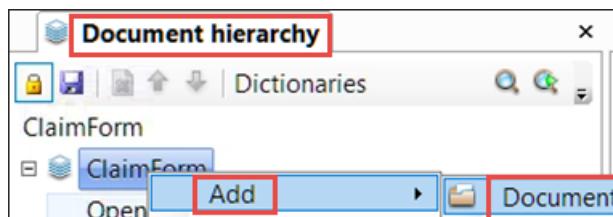
- \_\_\_ 3. Expand **ClaimForm** and check the nodes.

The **Other** page is the default page type that is assigned to each read page until the page is correctly identified.

The **Document** is the default document type.

- \_\_\_ 4. Create a Document type.

- \_\_\_ a. Right-click **ClaimForm** and select **Add > Document**.



- \_\_\_ b. Select **Document1**, click, and rename it as: **Insurance Docs**

- \_\_\_ 5. Create a Page type.

- \_\_\_ a. Right-click **Insurance Docs** and click **Add > Page**.

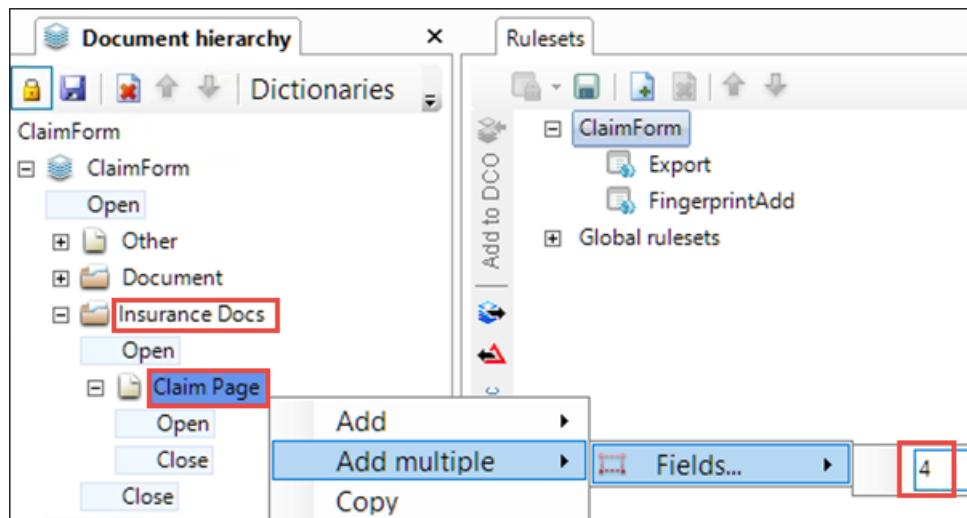
- \_\_\_ b. Select **Insurance Docs > Page1**, click, and rename it as: **Claim Page**

You can use any string. You cannot use the name of the application or another object.

- \_\_\_ 6. Add fields to the page.

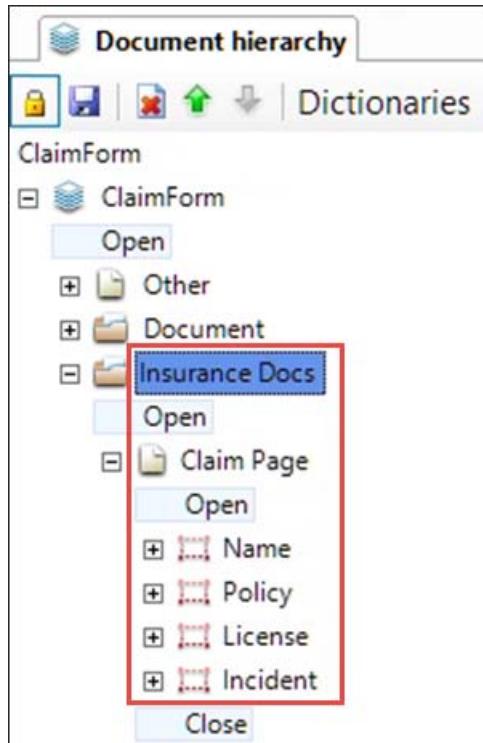
- \_\_\_ a. Right-click **Claim Page** and select **Add multiple > Fields**.

- \_\_\_ b. In the text box, enter **4** for the number of fields you are going to create and press Enter.



- \_\_\_ c. Verify that four fields are listed for the **Claim Page**.

- \_\_\_ d. Rename the fields with the following names in the order it is shown.
- **Field1:** Name
  - **Field2:** Policy
  - **Field3:** License
  - **Field4:** Incident
- \_\_\_ 7. Verify that the completed DCO has a document type, a page type, and four fields as shown in the following screen capture.



- \_\_\_ 8. Click the **Save changes** icon and click the **Unlock DCO lock** icon to unlock the DCO. Notice that for each Datacap object, Open and Close nodes are listed. They indicate the order of operations.
- \_\_\_ 9. Leave the Datacap Studio open for the next section.

## Section 6. Configure fingerprints and zones

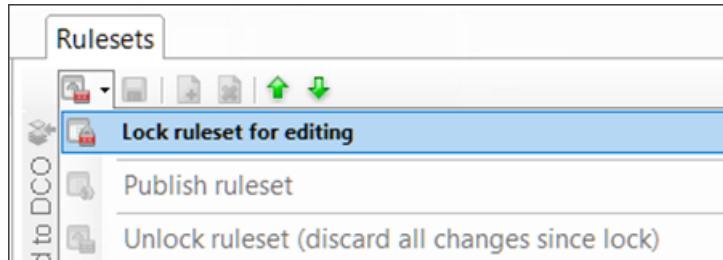
In this section, you configure FingerprintAdd ruleset and configure Fingerprints and zones in the Zones tab.

### 6.1. Configure the FingerprintAdd ruleset

In the following sections, you are going to use an OCR engine for the PageID ruleset to create the Fingerprint (CCO file). The OCR engine must match with the one that is used in the FingerprintAdd ruleset.

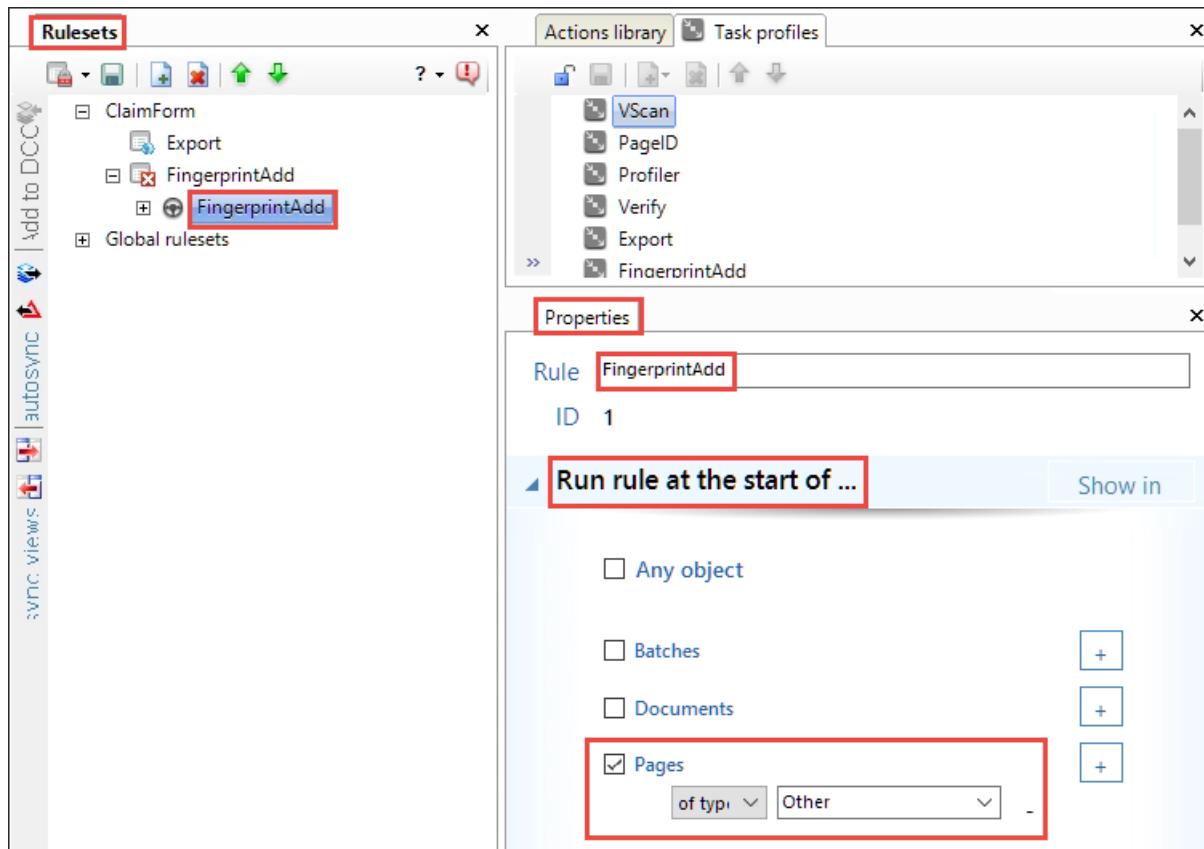
The application wizard creates a default FingerprintAdd ruleset. In this section, you review and edit the FingerprintAdd ruleset to match the OCR engine.

- \_\_\_ 1. In the **Rulesets** pane, expand **ClaimForm** and notice that the following rulesets are listed.
  - Export
  - FingerprintAdd
- \_\_\_ 2. Select the **FingerprintAdd** ruleset, click the lock icon on the toolbar, and from the list, click **Lock ruleset for editing**.



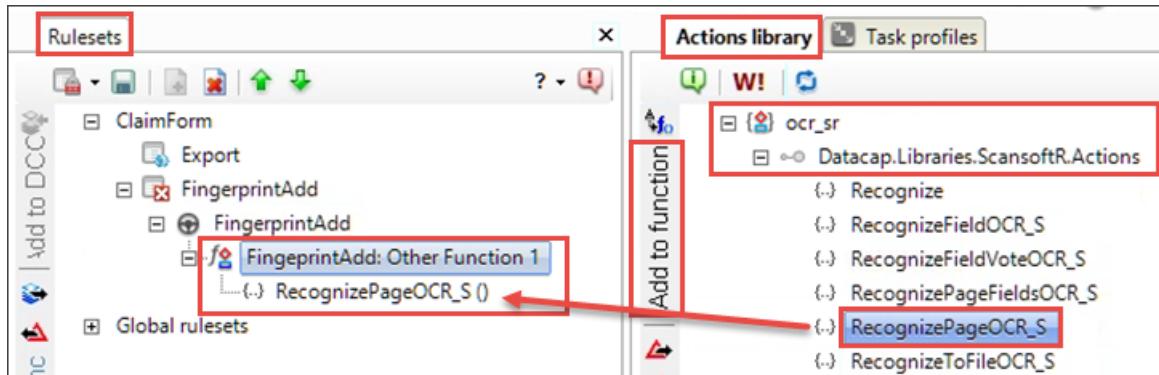
- \_\_\_ 3. Check that the rule runs at the page level.
  - \_\_\_ a. In the **Rulesets** pane, expand the **FingerprintAdd** ruleset and select the **FingerprintAdd** rule.
  - \_\_\_ b. On the **Properties** tab in the lower-right, expand the **Run rule at start of** node.

- \_\_\_ c. Verify that **Pages** is selected and for the **of type** field, **Other** is selected.



- \_\_\_ 4. Delete the action for FingerprintAdd: Other Function 1.
  - \_\_\_ a. Expand **FingerprintAdd > FingerprintAdd > FingerprintAdd: Other Function 1**.
  - \_\_\_ b. Right-click the **AnalyzeImage ()** action and click **Remove** to delete it.
- \_\_\_ 5. Add the RecognizePageOCR\_S action to the FingerprintAdd: Other Function 1 function.
  - \_\_\_ a. Select the **FingerprintAdd: Other Function 1** function.
  - \_\_\_ b. On the upper right pane, in the **Actions library** tab, scroll to the **ocr\_sr** library, expand the **ocr\_sr > Datacap.Libraries.ScansoftR.Actions** node, and select the **RecognizePageOCR\_S** action.

- \_\_\_ c. Click the **Add to function** vertical bar on the left side of the **Action library** tab (between the **Rulesets** and the **Actions library** panes).



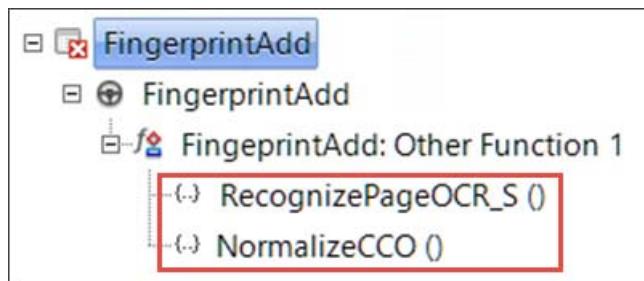
- \_\_\_ d. In the **Rulesets** pane, verify that the **RecognizePageOCR\_S()** action is added to **FingerprintAdd: Other Function 1**.

This action does not require any parameters. This action creates the Fingerprint file.

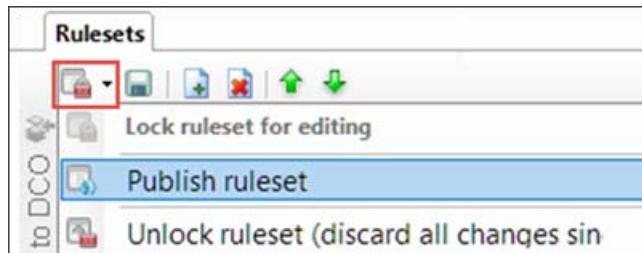
- \_\_\_ 6. Add the NormalizeCCO action to FingerprintAdd: Other Function 1.

The step is similar to the one you completed.

- Select **FingerprintAdd: Other Function 1**.
  - On the **Actions library** tab, from the **SharedRecognitionTools > SharedRecognitionTools.Actions** action library, select the **NormalizeCCO** action.
  - Click the **Add to function** vertical bar on the left side of the **Action library** tab.
- \_\_\_ 7. In the **Rulesets** pane toolbar, click the **Save changes** icon.
- \_\_\_ 8. Verify that the function contains the two actions in the order as shown in the following screen capture.



- 9. Select the **FingerprintAdd** ruleset, on the **Rulesets** pane toolbar, click the lock icon, and click **Publish ruleset** from the list.



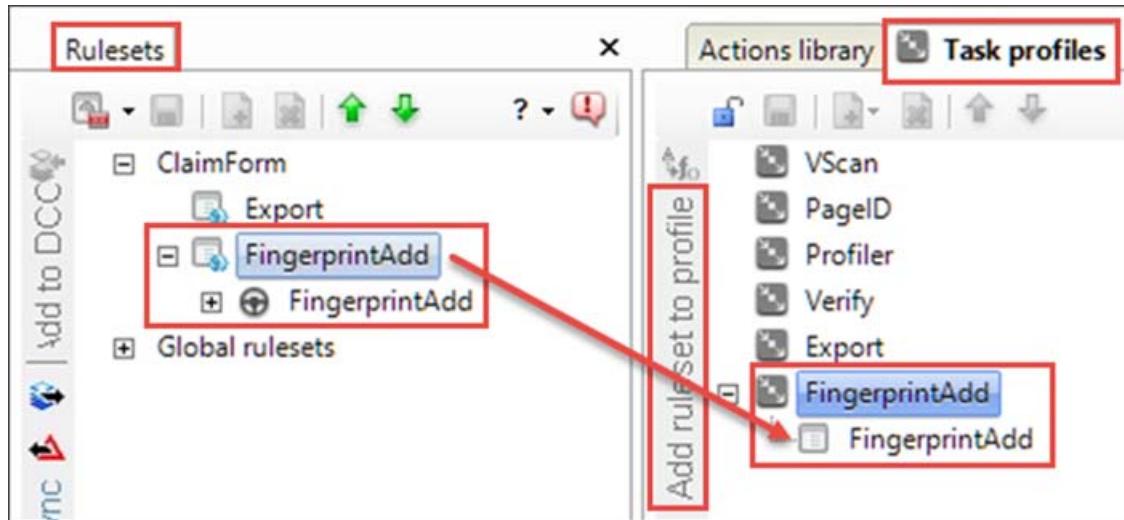
## 6.2. Add the FingerprintAdd ruleset to a task profile

A ruleset is recognized for running when it is associated with a task profile.

- 1. On the upper right pane, click the **Task profiles** tab, lock it by clicking the lock icon.
- 2. Select the **FingerprintAdd** task profile.

This is a special task profile and notice that it is outside of the five workflow steps that configured. Datacap recognizes this name and runs this task when you configure zones in the next section.

- 3. On the **Rulesets** tab, select the **FingerprintAdd** ruleset and on the right of the **Rulesets** pane, click the **Add ruleset to profile** vertical bar (between the **Rulesets** and the **Task profiles** panes).

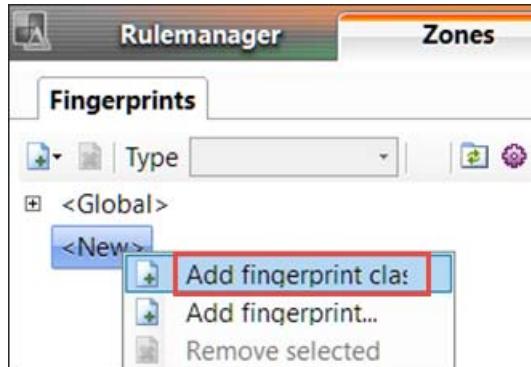


- 4. Click the **Save changes** icon on the **Task profiles** tab and unlock the **Task profiles** by clicking the lock icon.
- 5. Expand the completed **FingerprintAdd** task profile and verify that it contains the **FingerprintAdd** ruleset as shown in the following screen capture.

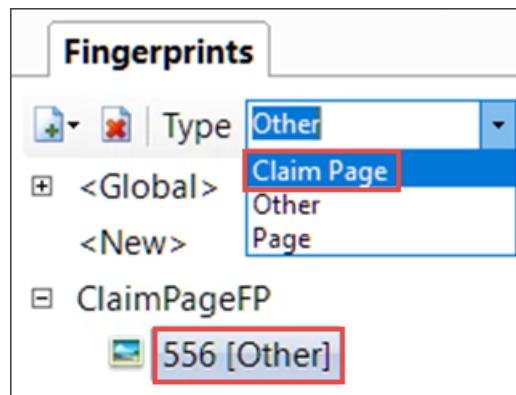
### 6.3. Configure Fingerprinting in Zones tab

In this section, you create fingerprinting to be used for recognizing the zones and extract values.

- \_\_\_ 1. In Datacap Studio, click the **Zones** tab.
- \_\_\_ 2. Create a fingerprint class.
  - \_\_\_ a. On the upper left pane, in the **Fingerprints** tab, right-click **New** and select **Add fingerprint class** from the list.



- \_\_\_ b. In the **New fingerprint class** dialog box, for the **Class name** field, enter: **ClaimPageFP** and click **OK**.
- \_\_\_ 3. Create a fingerprint.
  - \_\_\_ a. In the **Fingerprints** pane, right-click **ClaimPageFP** and from the list, select **Add fingerprint**.
  - \_\_\_ b. In the **Open** window, browse to the **C:\Datacap\ClaimForm\images** folder, select the **1-ClaimForm-BC.tif** file, and click **Open**.  
It takes a few moments to complete the cleanup of the image and create a fingerprint.
  - \_\_\_ c. In the **Fingerprints** pane, expand **ClaimPageFP** and verify that a fingerprint is created (for example, **556[Other]**).  
You might get a different number.
  - \_\_\_ d. Select **556[Other]** and on the toolbar, for the **Type** field, from the list, select **Claim Page**.



- \_\_\_ e. Verify that the value is changed to **556[Claim Page]**.



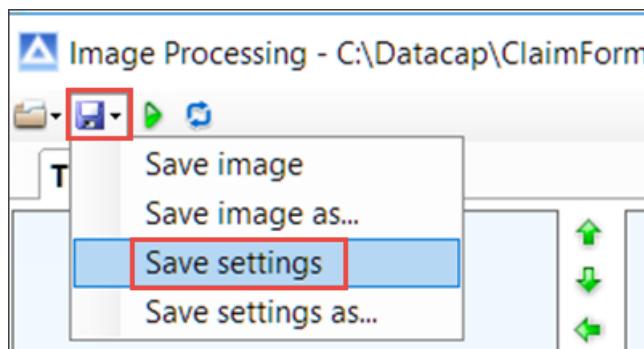
## Information

The items in the Type list are stored in the CLAIMFORM.XML file in the application folder (C:\Datacap\ClaimForm\dco\_ClaimForm).

- \_\_\_ 4. Configure the image settings.
  - \_\_\_ a. Select **556[Claim Page]** and verify that the image is opened in the **Image** pane on the right.
  - \_\_\_ b. On the **Image** tab, on the upper right of the page, click the gear icon.
  - \_\_\_ c. On the **Image Processing** page, in the **Properties** tab on the right, for the following fields, set the values.
    - Deskew: **True**
    - Despeckle: **True**
    - Line Removal: **False**



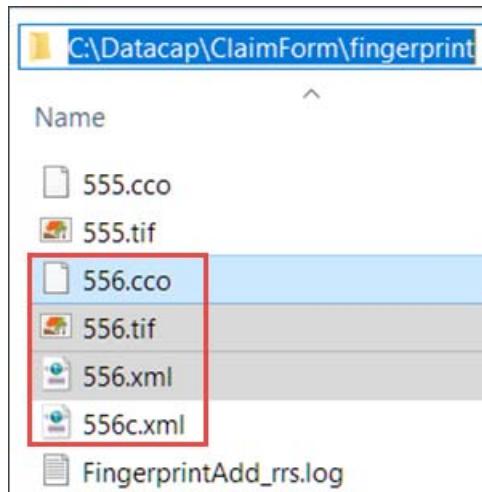
- \_\_\_ d. Click the **Save** icon and select **Save settings** from the list.



- \_\_\_ e. On the Setting saved dialog box, click **OK** and close the **Image Processing** page.
- \_\_\_ f. On the **Zones > Image** tab, for the image, verify that the lower barcode (**CLAIM-29A**) is intact.

Compare it with the original image that you used to create fingerprint  
(C:\Datacap\ClaimForm\images\1-ClaimForm-BC.tif).

- \_\_\_ 5. Check the fingerprint files that are created.
  - \_\_\_ a. In Windows Explorer, browse to the C:\Datacap\ClaimForm\fingerprint folder.
  - \_\_\_ b. Verify that the following files that start with **556** are created:  
You might have a different number on your system. The application wizard created <555>, which is the default fingerprint.
    - 556.tif**
    - 556.xml**
    - 556c.xml**
    - 556.cco**

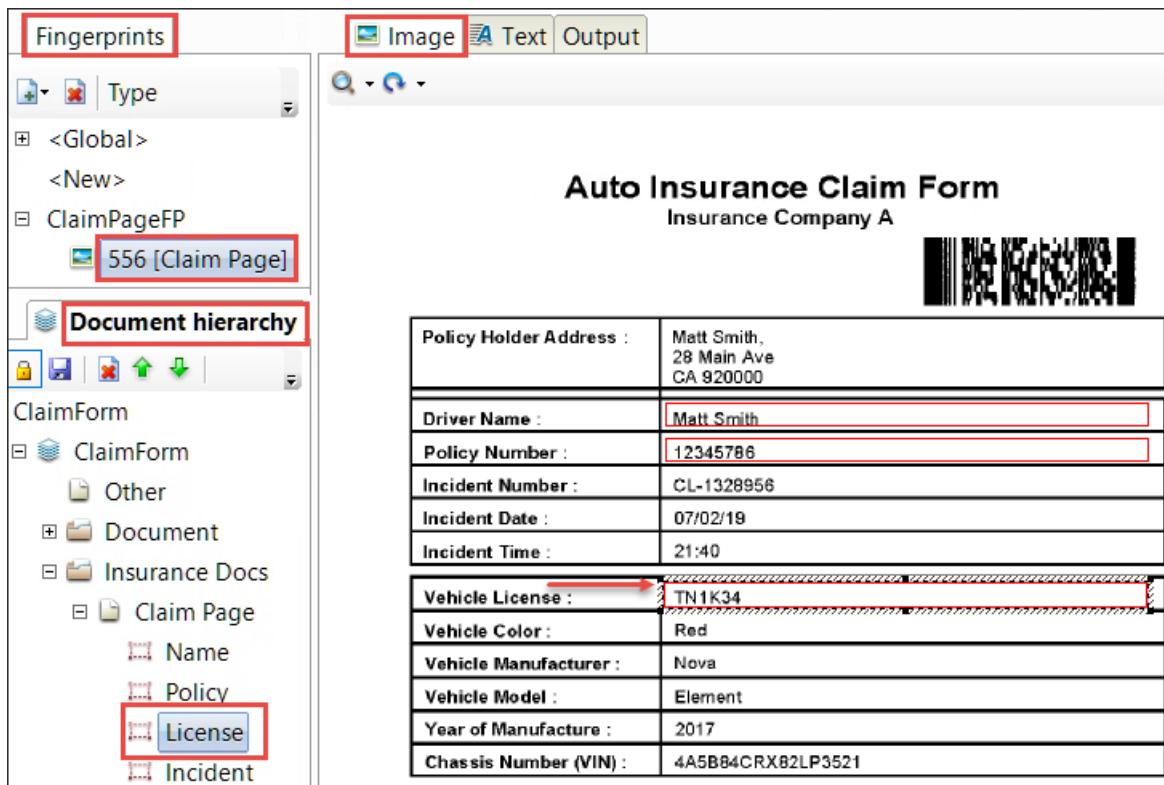


## 6.4. Configure zones

- \_\_\_ 1. On Datacap Studio **Zones** tab, in the **Fingerprints** tab, select **ClaimPageFP > 556 [Claim Page]** and verify that the image is shown on the **Image** tab.
- \_\_\_ 2. In the **Document hierarchy** tab on the left, click the lock icon to enable the DCO editing.
- \_\_\_ 3. Expand **ClaimForm > Insurance Docs > Claim Page** and select the **Name** field.
- \_\_\_ 4. On the **Image** tab, draw the zone around the value for the **Driver Name** field.

Draw from upper left corner to the lower right corner inside the box and avoid the border line added to the drawing. Refer to the following screen capture.

- 5. Repeat the steps to draw zones to other two fields: **Policy** (with **Policy number**) and **License** (with **Vehicle License**).



- 6. On the **Document hierarchy** pane toolbar, click the **Save** icon.



#### Note

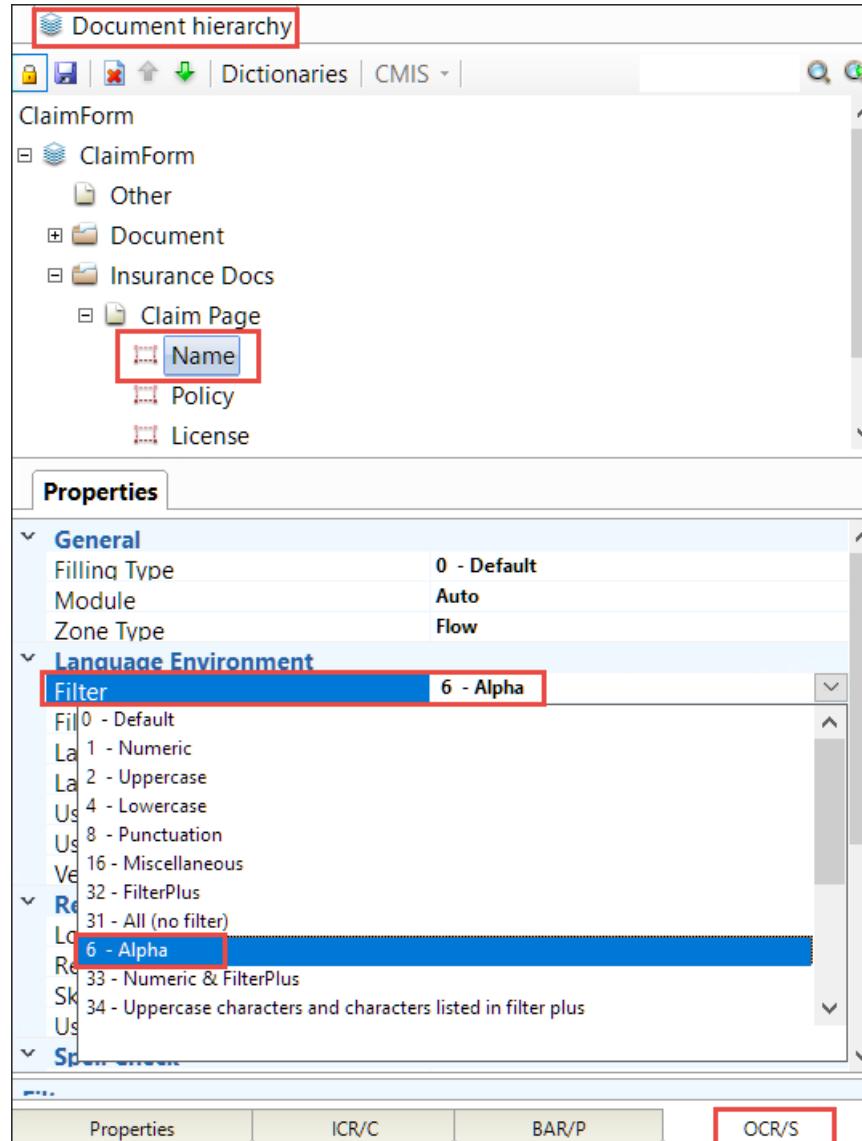
The **Incident** field does not require zones. To extract the value, you implement the keyword search method and use actions from the **Locate** action library.

## 6.5. Set OCR filters for the fields

In the **Zones** tab, in the **Document hierarchy** pane, the **ClaimForm > Insurance Docs > Claim Page** nodes are already expanded. It is also locked for editing from the previous section.

- 1. In the **Document hierarchy** pane, select the **Name** field.
- 2. In the **Properties** pane on the lower left, select the **OCR\_S** tab.

- \_\_\_ 3. For the **Filter** field, select **6 - Alpha** to limit the value to only alphabets for this field.



- \_\_\_ 4. Repeat the steps for the **Policy** field and select **1 - Numeric** for filter.

For the **License** field, no change is needed. Leave **0 - Default** for filter. Because License contains both letters and numbers.

- \_\_\_ 5. In the **Document hierarchy** pane, click **Save** and click the lock icon to unlock the DCO.  
 \_\_\_ 6. Notice that the icon color is changed for the fields after you complete the zones configuration.

## End of exercise

## Exercise review and wrap-up

This exercise showed how to build a Datacap application that uses Form Template in Datacap Studio.

# Exercise 3. Configuring rulesets for a Datacap application

## Estimated time

02:00

## Overview

This exercise teaches how to configure rulesets for a Datacap application in Datacap Studio, associate them with task profiles, and test the application.

## Objectives

After completing this exercise, you should be able to:

- Create rulesets
- Assign the rulesets to task profiles
- Test the application

## Introduction

Rulesets define what happens, Document hierarchy (DCO) define where does it happen, and Task profiles define when does it happen.

In a Datacap application, a ruleset consists of one or more rules. Each rule contains functions and actions. In the previous exercise, you created a Datacap application, defined DCO, and configured fingerprints and zones. In this exercise, you configure rulesets, rules, functions, and actions for the tasks in your application. You associate the rules to process specific objects in the document hierarchy. You assign the rulesets to specific task profiles and test the configuration.

This exercise includes the following sections:

- [Section 1, "Create ruleset for the VScan task"](#)
- [Section 2, "Configure PageID for the application"](#)
- [Section 3, "Add rulesets for the Profiler, Verify, and Export tasks"](#)
- [Section 4, "Test the application"](#)

## Requirements

This exercise builds on [Exercise 2, "Building a Datacap application with Forms Template"](#). Before starting this exercise, you must complete [Exercise 2, "Building a Datacap application with Forms Template"](#).

## Section 1. Create ruleset for the VScan task

In this section, you configure rules, functions, and actions for the VScan task for your application. You associate the rules to process specific objects in the document hierarchy. You assign the rulesets to specific task profiles and test the configuration.



### Note

A Datacap application solution for this exercise is provided in C:\Datacap folder with the application name, **SolutionClaimForm**. You can access this application like any other Datacap application.

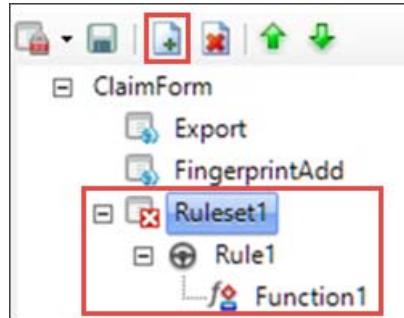
### 1.1. Create the VScan ruleset

This ruleset defines the rules, functions, and actions for scanning images.

- 1. Before you begin the exercise, ensure that the components are running. If you did not start the components already, refer to [Section 1, "Prepare your system: Start the components"](#) to start them.  
If you restart your system, you must start all the components.
- 2. If it is not already open, log in to Datacap Studio.
  - a. Double-click the **Datacap Studio** shortcut on the Windows desktop.
  - b. In the **Applications** dialog box, select **ClaimForm** and click **Next**.
  - c. On the **Taskmaster Login** page, enter the following values:
    - **User ID:** admin
    - **Password:** admin
    - **Station ID:** 1
  - d. Click **Finish**.

Datacap Studio opens with the **Rulemanager** tab selected.
- 3. In the **Rulesets** tab, select **ClaimForm** and in the toolbar, click the **Add child object**  icon.

A new ruleset node is added to **ClaimForm** with the name **Ruleset1** and it has a red x flag to indicate that it is locked for editing.



**Ruleset1** has **Rule1** and **Rule1** contains **Function1**.

- \_\_\_ 4. Select **Ruleset1**, click, and rename it to: vScan

You can also select **Ruleset1** and in the **Properties** tab on the lower right, for the **Ruleset** field, type the new name.

- \_\_\_ 5. Rename the following objects:

- **Rule1**: (B)Open



### Hint

The name **(B) Open** indicates that this rule runs at the batch level and at the opening of the batch.

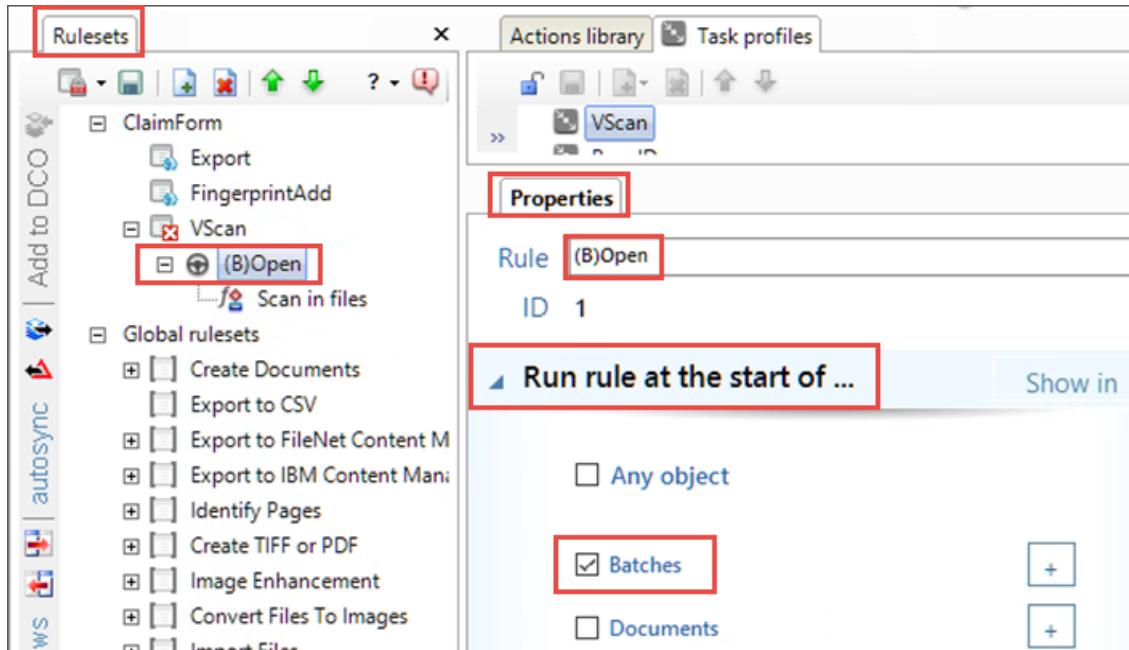
As you create more rules, arrange them in the following order (B) Open > (P) Open > (F) Open to indicate the object level. (P) for page level and (F) for field level.

You can enter any name for a rule, but this naming convention helps identify the order of operations. The naming convention also helps to avoid adding multiple Batch open rules. Because in a ruleset, you are allowed to have only one Batch open level operation.

- **Function1**: Scan in files
- \_\_\_ 6. Assign the (B) Open rule to batches.

This assignment determines that the actions that are defined in the rule are run on this object.

- \_\_\_ a. In the **Rulesets** pane, select the **VScan > (B) Open** rule.
- \_\_\_ b. On the lower-right pane, in the **Properties** tab, expand the **Run rule at start of** node.
- \_\_\_ c. Select **Batches**.



- \_\_\_ 7. Add an action to the Scan in files function.
  - \_\_\_ a. Select the **Scan in files** function.
  - \_\_\_ b. On the upper right pane, click the **Actions library** tab, scroll to the **mvscan** library, expand the **mvscan > Datacap.Libraries.mvScan.Actions** nodes, and select the **set\_folder** action.

This action specifies the top-level folder to search for the files to be scanned.

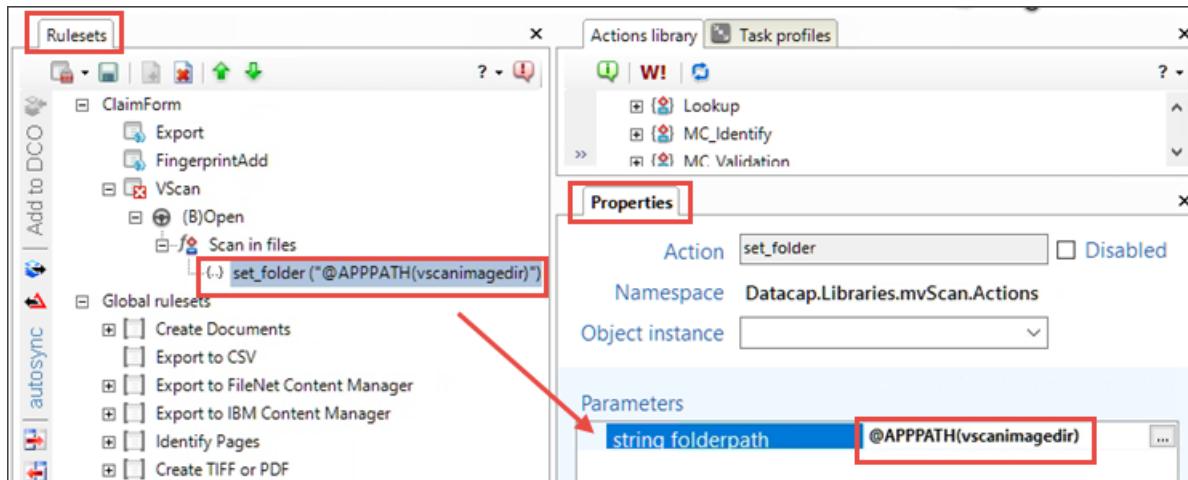


### Hint

Select an action in the Action library tab, click the information icon to get more information. Example: Parameters that are needed for the action, at what level the action is run (batch, document, pages, or fields), and what other actions are required to achieve the results.

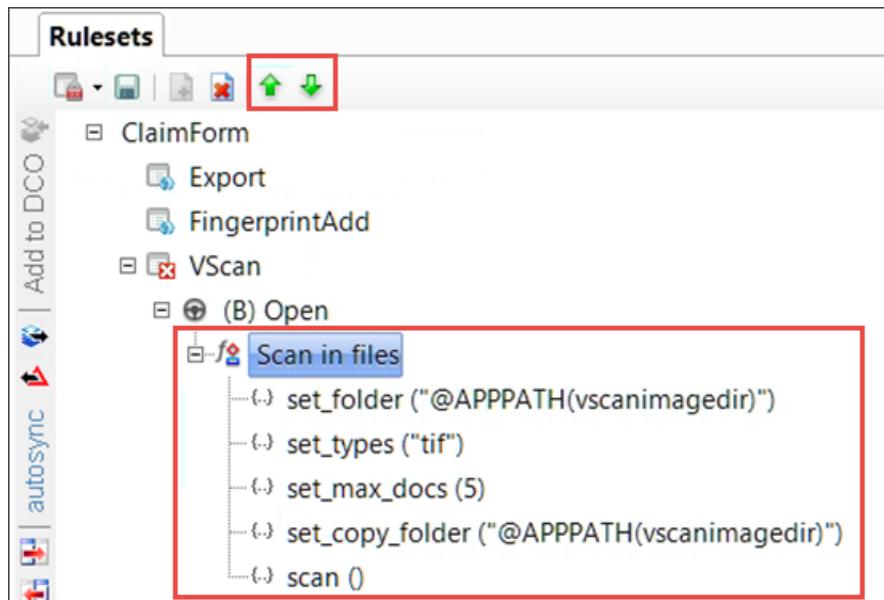
- \_\_\_ c. On the left side of the **Action library** tab, click the **Add to function** vertical bar between the **Rulesets** and the **Actions library** panes).
- \_\_\_ d. In the **Rulesets** pane, for the **(B) Open > Scan in files** function, verify that the **set\_folder ("")** action is added and select the action.
- \_\_\_ e. In the **Properties > Parameters** section, set the **string folderpath** field to: **@APPATH(vscanimagedir)**

Recall that the images folder is defined in Datacap Application Manager. The image folder is the parameter value.

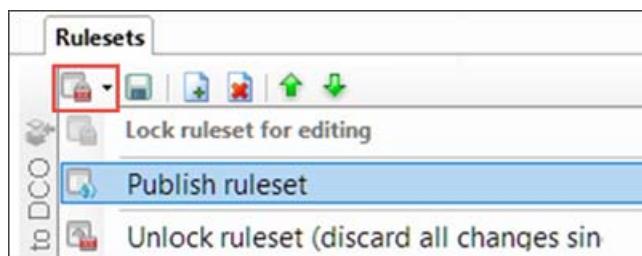


- \_\_\_ 8. Add other actions to the Scan in files function.
  - \_\_\_ a. Select the **Scan in files** function and from the **mvscan** action library, add the **set\_types** action.
  - \_\_\_ b. In the **Rulesets** pane, select the **set\_types ("")** action and in the **Properties** tab > **Parameters** section, for the **string extensions** field, enter: **tif**  
This value specifies the type of files that the task scans.
  - \_\_\_ c. Select the **Scan in files** function and from the **mvscan** action library, add the **set\_max\_docs** action.

- \_\_\_ d. In the **Rulesets** pane, select the **set\_max\_docs ("")** action and in the **Properties** tab > **Parameters** section, for the **int nDocs** field, enter: 5  
 This value determines the maximum number of the documents that are scanned in one batch.
- \_\_\_ e. Select the **Scan in files** function and from the **mvscan** action library, add the **set\_copy\_folder** action.
- \_\_\_ f. In the **Rulesets** pane, select the **set\_copy\_folder ("")** action and in the **Properties** tab > **Parameters** section, for the **string folderpath** field, enter:  
`@APPATH(vscanimagedir)`
- \_\_\_ g. Select the **Scan in files** function and from the **mvscan** action library, add the **scan** action.
- \_\_\_ 9. In the toolbar, use the arrows to arrange the actions in the order as shown in the following screen capture.



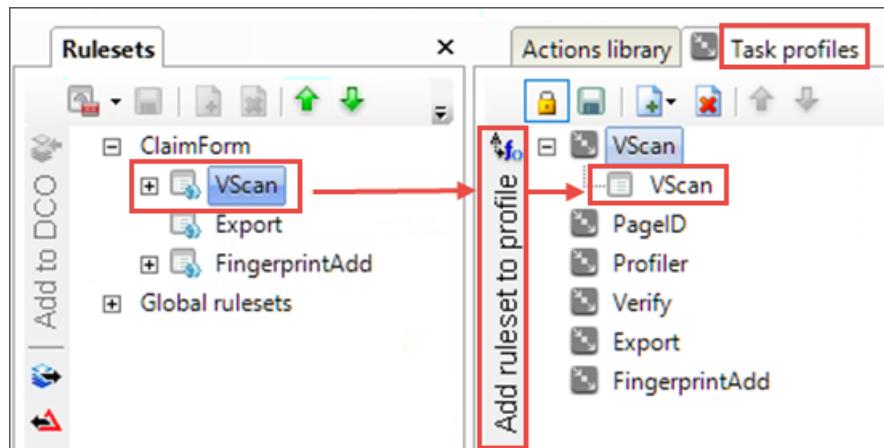
- \_\_\_ 10. Save and publish the ruleset.
- \_\_\_ a. Click the **Save changes** icon in the toolbar.
- \_\_\_ b. Select the **VScan** ruleset, on the toolbar of the **Rulesets** tab, click the lock icon, and from the list, click **Publish ruleset**.



- \_\_\_ c. Select the **VScan** ruleset, on the toolbar, click the **Move up** icon to place the ruleset as the first one in the list (above the **Export** ruleset).

## 1.2. Add the ruleset to the task profile

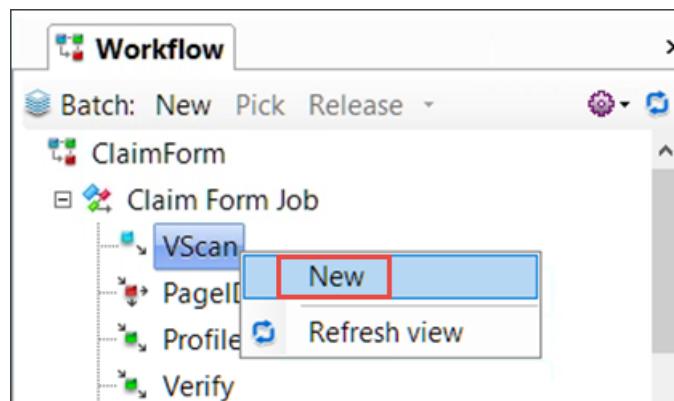
- \_\_\_ 1. On the upper right pane, click the **Task profiles** tab, lock it by clicking the lock  icon.
- \_\_\_ 2. Select the **VScan** task profile.
- \_\_\_ 3. On the **Rulesets** tab, select the **VScan** ruleset and on the right of the **Rulesets** pane, click **Add ruleset to profile** (between the **Rulesets** and the **Task profiles** panes).
- \_\_\_ 4. On the **Task profiles** tab, expand the **VScan** task profile and verify that the **VScan** ruleset is added.



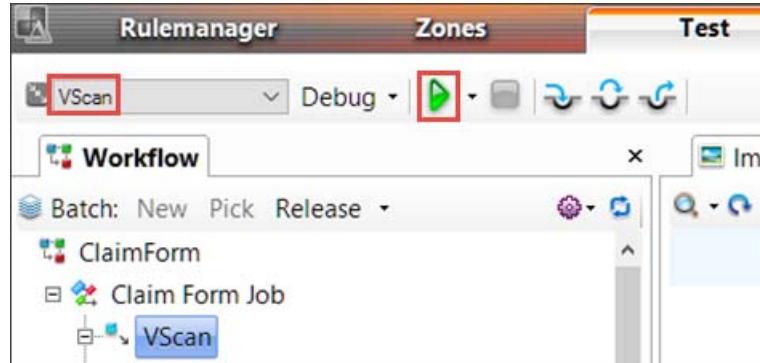
- \_\_\_ 5. On the **Task profiles** tab, click the **Save changes**  icon and unlock the **Task profiles** by clicking the lock icon.

## 1.3. Start a batch to test the VScan ruleset

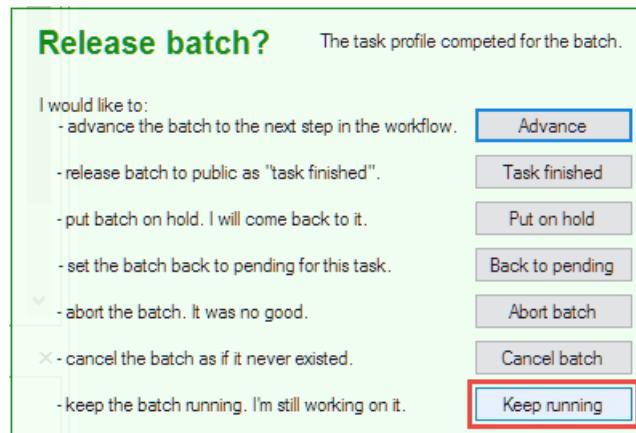
- \_\_\_ 1. In Datacap Studio, click the **Test** tab on the top toolbar.
- \_\_\_ 2. On the left, notice that the **Runtime batch hierarchy** tab has no items.  
When you run a test batch, the batch has items.
- \_\_\_ 3. Start a batch.
  - \_\_\_ a. In the upper left pane, on the **Workflow** tab, expand **ClaimForm > ClaimForm Job** right-click **VScan**, and select **New**.



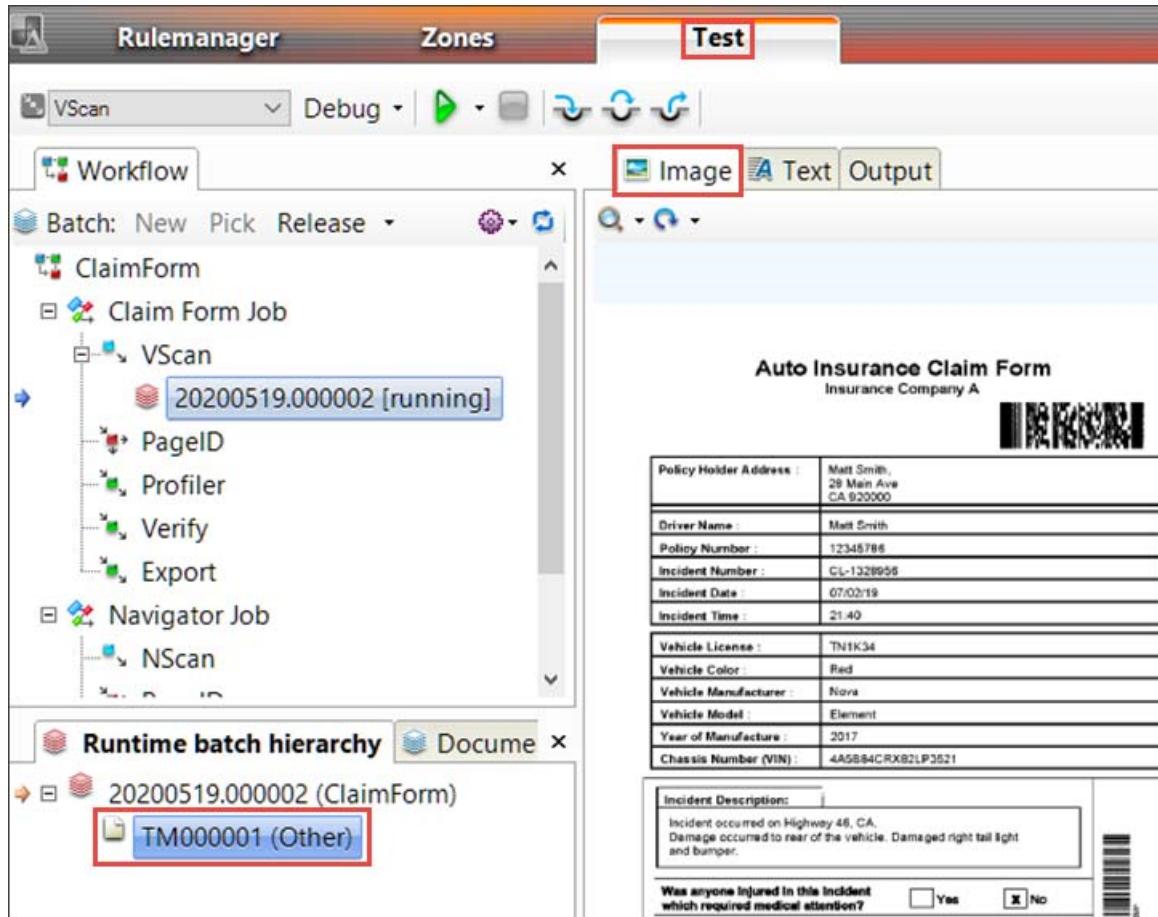
- \_\_\_ b. In the top toolbar, ensure that **VScan** is selected and click the **Process rules for target object** (green arrow icon) next to it.



- \_\_\_ 4. In the **Release batch** dialog box, click **Keep running**.

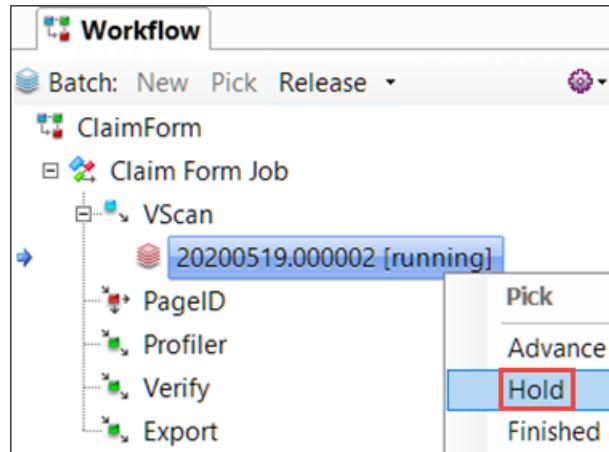


- \_\_\_ 5. In the **Runtime batch hierarchy tab**, notice that it has a batch item now, expand the batch node and verify that a page is listed as **Other**.
- \_\_\_ 6. Click **TM000001 (Other)** and in the **Image** tab, verify that the scanned image is shown.  
This test confirms that the VScan ruleset is configured correctly.



- 7. In the upper-left pane, on the **Workflow** tab, expand **ClaimForm > ClaimForm Job > VScan**, and right-click the running batch and select **Hold**.

You can configure more rulesets and continue the test with this batch.



## Section 2. Configure PagID for the application

In this section, you define rulesets for identifying the page by using the barcode method. You associate the rules to process specific objects in the document hierarchy. You assign the rulesets to specific task profiles and test the configuration.

### 2.1. Configure the PagID ruleset for barcode recognition

The PagID ruleset contains rules, functions, and actions for classifying the document and identifying the page.

You build the **PagID** and **CreateDocs&Fields** rulesets for the PagID task.

- \_\_\_ 1. In Datacap Studio, click the **Rulemanager** tab from the top toolbar.
- \_\_\_ 2. Add a ruleset.
  - \_\_\_ a. In the **Rulesets** tab, select **ClaimForm** and click **Add child object** (plus) icon.  
A new ruleset node is added to the **ClaimForm** node with the name **Ruleset1** and it has a red x to indicate that it is locked for editing.
  - \_\_\_ b. Select and click each object and rename it with the following values:
    - **Ruleset1:** PageID
    - **Rule1:** (B) Open  
The values that you are going to set in the actions for this rule does not change per page and they need to be set only one time. Therefore, it is efficient to run the rule at the batch level.
    - **Function1** to Set Fingerprint Params
- \_\_\_ 3. Assign the (B) Open rule to batches.
  - \_\_\_ a. In the **Rulesets** pane, select the **PagID > (B) Open** rule.
  - \_\_\_ b. In the **Properties** tab, expand the **Run rule at start of** node.
  - \_\_\_ c. Select **Batches**.



- \_\_\_ 4. Add an action to the Set Fingerprint Params function.
  - \_\_\_ a. Select the **Set Fingerprint Params** function.
  - \_\_\_ b. On the upper right pane, in the **Actions library** tab, expand the **autodoc** action library and select the **SetFingerprintDir** action.
  - \_\_\_ c. On the left side of the **Action library** tab, click the **Add to function** vertical bar between the **Rulesets** and the **Actions library** panes.
  - \_\_\_ d. In the **Rulesets** pane, for the **Set Fingerprint Params** function, verify that the **SetFingerprintDirectory ("")** action is added and select the action.
  - \_\_\_ e. In the **Properties** tab > **Parameters** section, for the **string StrParam** field, enter: `@APPATH(fingerprint)`

Recall that the Fingerprint folder is defined in the Datacap Application Manager.

- \_\_\_ 5. Select the **Set Fingerprint Params** function and from the **AutomaticDocumentFingerprinting > Datacap.Libraries.AutomaticDocumentFingerprinting.Actions** action library, add the **SetFingerprintSearchArea** action.
- \_\_\_ 6. In the **Rulesets** pane, select the **SetFingerprintSearchArea ("", "")** action and in the **Properties** tab > **Parameters** section, for the **string matchStart** field, enter 0.0 and for the **string matchEnd** field, enter: 0.5

The action controls the portion of the current page that is used to find a matching fingerprint.

- \_\_\_ 7. Select the **Set Fingerprint Params** function and from the **AutomaticDocumentFingerprinting > Datacap.Libraries.AutomaticDocumentFingerprinting.Actions** action library, add the **SetMinimumMatchingTolerance** action.
- \_\_\_ 8. In the **Rulesets** pane, select the **SetMinimumMatchingTolerance ("")** action and in the **Properties** tab > **Parameters** section, for the **string Minimum** field, enter: 0.7

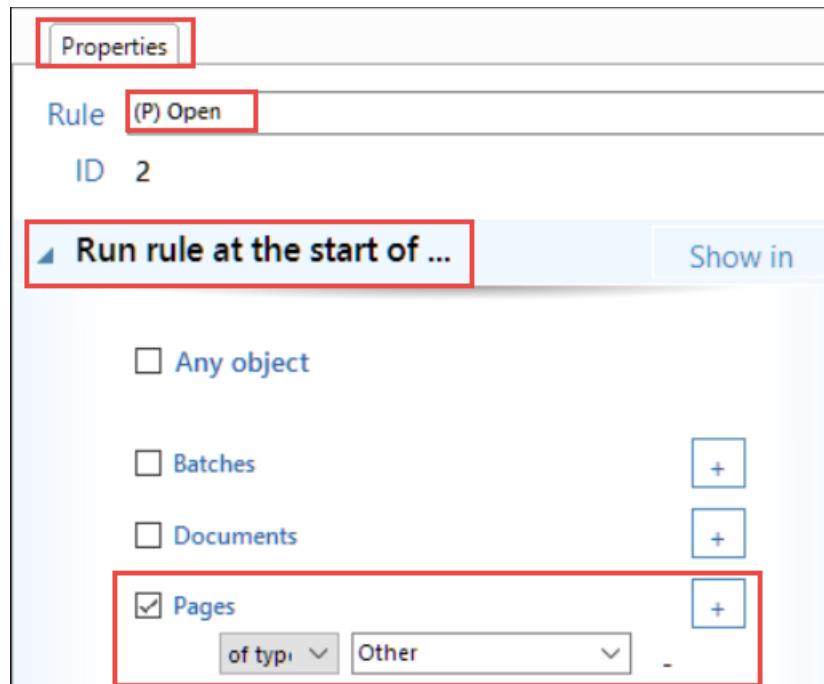
This action sets the minimum confidence matching tolerance rating to determine whether a fingerprint is a match.

- \_\_\_ 9. In the **Rulesets** pane toolbar, click the **Save changes**  icon.
- \_\_\_ 10. Verify that the completed **PageID > (B) Open** rule contains all the actions that you configured as shown in the screen capture.



## 2.2. Add a page level rule to the PagelD ruleset

- \_\_\_ 1. Select **PagelD** and click the **Add child object** (plus) icon.
- \_\_\_ 2. Select and click each object and rename it with the following values:
  - **Rule1:** (P)Open
  - **Function1:** Classify
- \_\_\_ 3. Assign the (P) Open rule to batches.
  - \_\_\_ a. In the **Rulesets** pane, select the **PagelD > (P) Open** rule.
  - \_\_\_ b. In the **Properties** tab, expand the **Run rule at start of** node.
  - \_\_\_ c. Select **Pages** and click the plus sign next to it to select the specific page.
  - \_\_\_ d. For the **of type** field, select **Other** from the list.



- \_\_\_ 4. Add an action to the Classify function.
  - \_\_\_ a. Select the **Classify** function.
  - \_\_\_ b. On the **Actions library** tab, expand **Barcode > Datacap.Libraries.Barcode.Actions** and select the **MatchAnyBarcode** action.
  - \_\_\_ c. On the left side of the **Action library** tab, click the **Add to function** vertical bar.
  - \_\_\_ d. In the **Rulesets** pane, for the **(P) Open > Classify** function, verify that the **MatchAnyBarcode ("")** action is added and select the action.
  - \_\_\_ e. In the **Properties** tab > **Parameters** section, for the **string ExpectedBarcodeValue** field, enter: CLAIM-29A

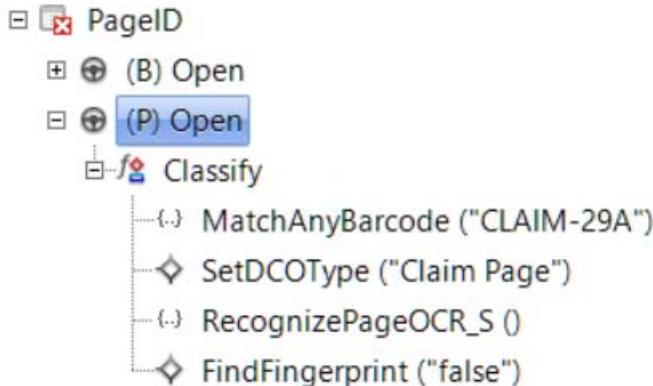
CLAIM-29A is the value of the barcode on the scanned image.

The MatchAnyBarcode action searches all the barcodes on the current page and checks if any one matches the value you enter as a parameter.

- 5. Add SetDCOType, RecognizePageOCR\_S, and FindFingerprint actions.
  - a. Select the **Classify** function and from the **Datacap.Libraries.ApplicationObjects.Actions** action library, add the **SetDCOType** action.
  - b. Select the **Classify > SetDCOType ("")** action and in the **Properties** tab, for the **string Type** field, enter: Claim Page
 

This action assigns a value to the Type property of the current object (Page) in the Document Hierarchy.
  - c. Select the **Classify** function and from the **ocr\_sr > Datacap.Libraries.ScansoftR.Actions** action library, add the **RecognizePageOCR\_S** action.
  - d. Select the **Classify** function and from the **autodoc** action library, add the **FindFingerprint** action.
  - e. Select the **FindFingerprint** action and in the **Properties** tab, for the **string StrParam** field, enter: false
 

If scanned image does not match the existing fingerprint, you have options to create more fingerprints. You set false when you don't want to create more fingerprints.
- 6. Verify that the completed PageID ruleset contains the (P) Open rule, a function, and actions as shown in the following screen capture.



- 7. Save, publish, and order the ruleset.
  - a. In the **Rulesets** tab toolbar, click the **Save changes** icon.
  - b. Select the **PageID** ruleset, on the toolbar, click the lock icon, and click **Publish ruleset** from the list.
  - c. Select the **PageID** ruleset, click the **Move up** icon on the toolbar to move up the ruleset, and place it between **VScan** and **Export** in the list.

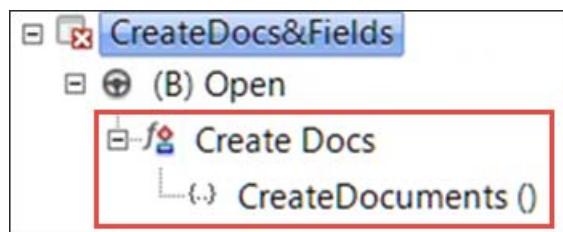
## 2.3. Configure the CreateDocs&Fields ruleset

In this section, you create the CreateDocs&Fields ruleset and configure rules.

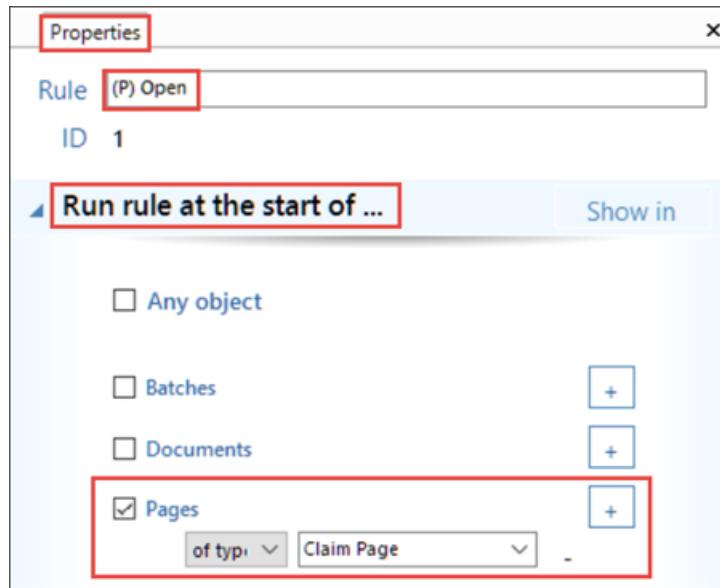
- \_\_\_ 1. In the **Rulesets** tab, select **ClaimForm** and click **Add child object** (plus) icon.
- \_\_\_ 2. Select and click each object and rename it with the following values:
  - **Ruleset1**: CreateDocs&Fields
  - **Rule1**: (B) Open
  - **Function1**: Create Docs
- \_\_\_ 3. Assign the (B) Open rule to batches.
  - \_\_\_ a. In the **Rulesets** pane, expand the **CreateDocs&Fields** and select the **(B) Open** rule.
  - \_\_\_ b. In the **Properties** tab, expand the **Run rule at start of** node.
  - \_\_\_ c. Select **Batches**.



- \_\_\_ 4. Add an action to the Create Docs function.
  - \_\_\_ a. In the **Rulesets** pane, select the **Create Docs** function.
  - \_\_\_ b. On the **Actions library** tab, expand the **ApplicationObjects > Datacap.Libraries.ApplicationObjects.Actions** action library and select the **CreateDocuments** action.
  - \_\_\_ c. Click **Add to function**.
- \_\_\_ 5. Click the **Save changes** icon in the toolbar and verify the ruleset.



- \_\_\_ 6. Add a page level rule to the **CreateDocs&Fields** ruleset.
  - \_\_\_ a. Select **CreateDocs&Fields** and click the **Add child object** (plus) icon.
  - \_\_\_ b. Select and click each object and rename it with the following values:
    - **Rule1:** (P) Open
    - **Function1:** Create Fields
- \_\_\_ 7. Assign the (P) Open rule to Pages.
  - \_\_\_ a. In the **Rulesets** pane, select the **CreateDocs&Fields > (P) Open** rule.
  - \_\_\_ b. In the **Properties** tab, expand the **Run rule at start of** node.
  - \_\_\_ c. Select **Pages** and click the plus sign next to it to select the specific page.
  - \_\_\_ d. For the **of type** field, select **Claim Page** from the list.



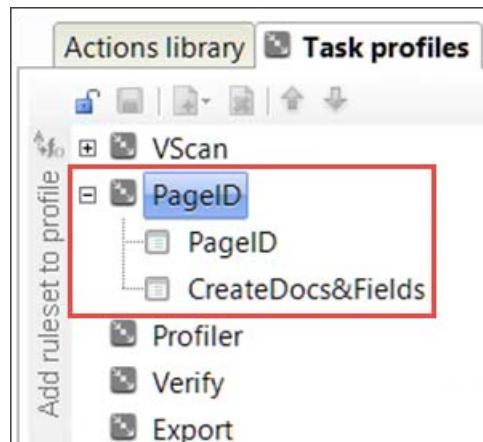
- \_\_\_ 8. Select the **Create Fields** function and from the **ApplicationObjects > Datacap.Libraries.ApplicationObjects.Actions** action library, add the **CreateFields** action.
- When you complete the ruleset, verify that it contains the two rules, functions, and actions as shown in the screen capture.



- \_\_\_ 9. Save changes and publish the ruleset.
  - \_\_\_ a. In the **Rulesets** tab toolbar, click the **Save changes** icon.
  - \_\_\_ b. Select the **CreateDocs&Fields** ruleset, on the toolbar, click the lock icon and click **Publish ruleset** from the list.
  - \_\_\_ c. Select the **CreateDocs&Fields** ruleset, click the **Move up** (green upward arrow) icon on the toolbar to move up the ruleset, and place it between **PageID** and **Export** in the list.

## 2.4. Add the two rulesets to the PageID task profile

- \_\_\_ 1. Click the **Task profiles** tab, lock it by clicking the lock  icon.
- \_\_\_ 2. Select the **PageID** task profile.
- \_\_\_ 3. On the **Rulesets** tab, select the **PageID** ruleset and on the right of the **Rulesets** pane, click **Add ruleset to profile**.
- \_\_\_ 4. Repeat the steps to add **CreateDocs&Fields** ruleset to the **PageID** task profile.
- \_\_\_ 5. On the **Task profiles** tab, expand the **PageID** task profile and verify that the completed **PageID** task profile contains two rulesets: **PageID** and **CreateDocs&Fields**
- \_\_\_ 6. On the **Task profiles** tab, click the **Save changes**  icon and unlock the **Task profiles** by clicking the lock icon.
- \_\_\_ 7. Verify that the **PageID** task profile contains the two rulesets.



## 2.5. Test the Page ID task for the application

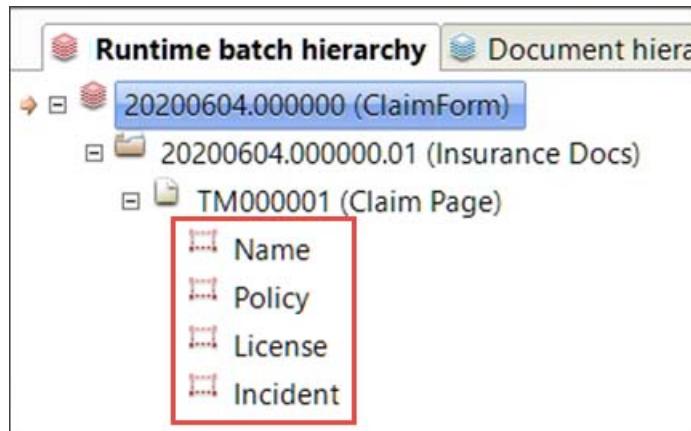
In this section, you test the rulesets that you added for the PageID task. In the Test tab, you already have a batch on hold from the previous section (when you tested VScan). In this section, you can continue with that batch.

- \_\_\_ 1. In Datacap Studio, click the **Test** tab from the top toolbar.
- \_\_\_ 2. On the upper left pane, in the **Workflow** tab, expand **ClaimForm > ClaimForm Job > VScan**, right-click the batch on hold, and click **Pick**.  
The batch now shows in running stage.

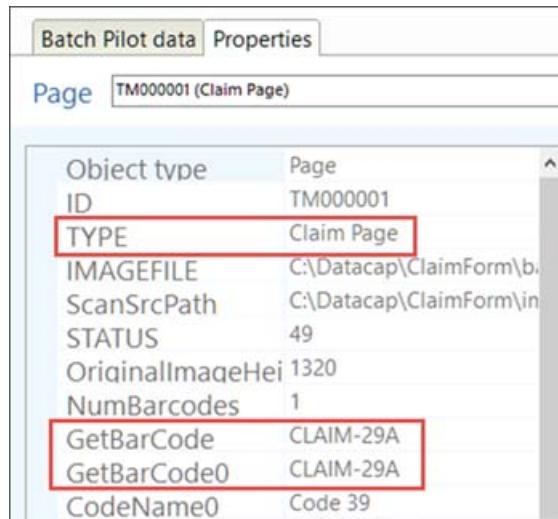
- \_\_\_ 3. In the top toolbar, select **Page ID** from the list and click the green arrow  icon to run the task.

Wait for the process to complete. This step cleans up the page, recognizes the barcode, classifies the document, and identifies the page.

- \_\_\_ 4. Verify the results of the PageID task.
- \_\_\_ a. In the **Release batch** dialog box, click **Keep Running**.
  - \_\_\_ b. In the **Runtime batch hierarchy tab**, expand the batch node and verify that the document is classified with the name: **Insurance Docs**
  - \_\_\_ c. Expand **Insurance Docs** and verify that the page is identified as **Claim Page**.
  - \_\_\_ d. Expand **Claim Page** and verify that the page contains the four fields that you defined: **Name, Policy, License, and Incident**.



- \_\_\_ e. Click **TM000001(Claim Page)** and check that the scanned image is displayed in the middle pane.
- \_\_\_ f. On the upper right, click the **Properties** tab and view the field values such as **Type: Claim Page, GetBarcode: CLAIM-29A, and Template ID: 556**



**Claim Page** is the type of your fingerprint that you configured on the **Zones** tab and **CLAIM-29A** is the barcode value on the image that was scanned.

<556> is the number of your fingerprint. This number might be different on your student image. Verify the number with the one on the **Zones** tab.

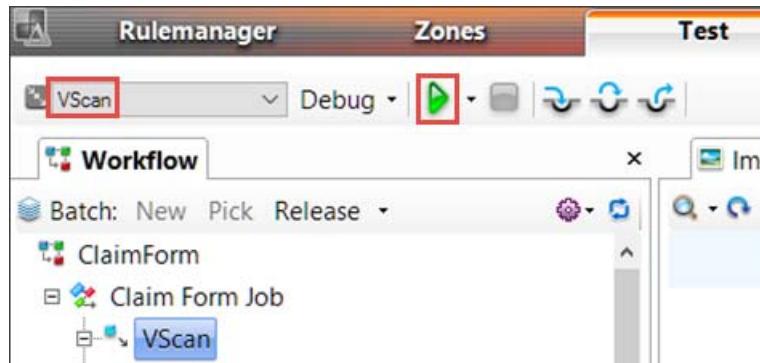
- \_\_\_ 5. In the **Workflow** tab, expand **ClaimForm > ClaimForm Job > VScan**, and right-click the running batch and select **Cancel**.



## Troubleshooting

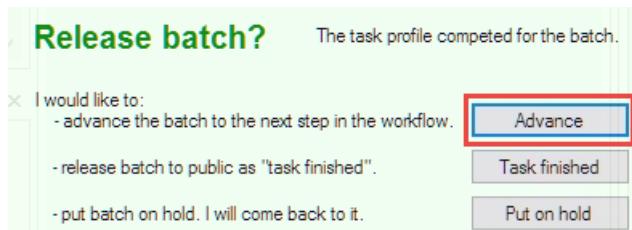
For testing, if you don't find your previous batch in the **ClaimForm > ClaimForm Job > VScan** node, start a new batch to test the application as given in the following steps.

- \_\_\_ a. On the **Workflow** tab, if it is already not expanded, expand **ClaimForm > ClaimForm Job**, right-click **VScan**, and select **New**.
- \_\_\_ b. In the top toolbar, ensure that **VScan** is selected.
- \_\_\_ c. If it is not already selected, select it from the list and click the **Process rules for target object** (green arrow icon) next to it.



Wait for the process to complete.

- \_\_\_ d. In the **Release batch** dialog box, click **Advance** to move the batch to the **PageID** task.



- \_\_\_ e. Click **TM000001 (Other)** and the scanned image is shown in the **Image** tab.
- \_\_\_ f. Continue with the **Section 8.4, Step 3**.

## Section 3. Add rulesets for the Profiler, Verify, and Export tasks

In the previous sections, you configured the application for the VScan and PageID tasks and tested them. In this section, you continue to configure the application for the Profiler (to extract values), Verify (to validate), and Export (to export data to a file) tasks.

### 3.1. Configure the Recognize ruleset

In this section, you build the Recognize ruleset for the Profiler task that extracts values from the scanned image.

- \_\_\_ 1. Select **ClaimForm** and click **Add child object** (plus) icon.
- \_\_\_ 2. Verify that a new ruleset node is added to the **ClaimForm** node with the name **Ruleset1** and it has a red x to indicate that it is locked for editing.

**Ruleset1** has **Rule1** and **Rule1** contains **Function1**.

The following table provides summary of data to add the **(B) Open** rule, function, and action to the **Recognize** ruleset.

- \_\_\_ 3. Rename the following objects:
  - **Ruleset1**: Recognize
  - **Rule1**: (B) Open

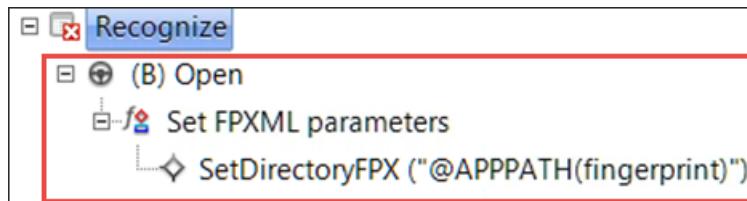
The value that you set for the action in this rule does not change per page and they need to be set only one time. Therefore, it is efficient to run the rule at the batch level.

  - **Function1** to Set FPXML parameters
- \_\_\_ 4. Assign the (B) Open rule to batches.
  - a. In the **Rulesets** pane, select the **Recognize > (B) Open** rule.
  - b. In the **Properties** tab, expand the **Run rule at start of** node.
  - c. Select **Batches**.



- \_\_\_ 5. Add an action to the Set FPXML parameters function.
  - a. Select the **Set FPXML parameters** function.

- \_\_\_ b. On the **Actions library** tab, scroll to and expand the **FPXML** action library, and select the **SetDirectoryFPX** action.
- \_\_\_ c. On the left side of the **Action library** tab, click the **Add to function** vertical bar.
- \_\_\_ d. In the **Rulesets** pane, for the **Set FPXML parameters** function, verify that the **SetDirectoryFPX ("")** action is added and select the action.
- \_\_\_ e. In the **Properties** tab > **Parameters** section, for the **string StrParam** field, enter: `@APPATH(fingerprint)`  
Recall that the Fingerprint folder is defined in the Datacap Application Manager.
- \_\_\_ f. In the **Rulesets** pane toolbar, click the **Save changes** icon and verify the ruleset.



### 3.2. Configure a rule for reading zones

In this section, you add another rule to the Recognize ruleset. The rule contains a function with actions to reading zones.

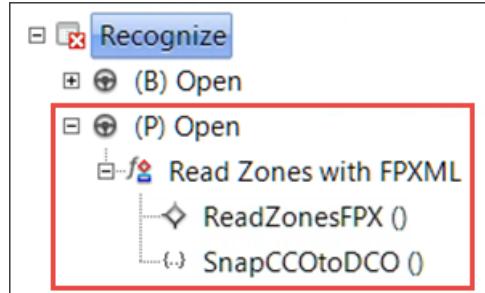
- \_\_\_ 1. Add another rule to the Recognize ruleset.
  - \_\_\_ a. Select **Recognize** and click **Add child object** (plus) icon.
  - \_\_\_ b. Verify that a new rule node is added to **Recognize** with the name **Rule1** and it has **Function1**.
- \_\_\_ 2. Rename the following objects:
  - **Rule1**: (P)Open
  - **Function1**: Read Zones with FPXML
- \_\_\_ 3. Assign the rule to Claim Page.
  - \_\_\_ a. In the **Rulesets** pane, select the **Recognize > (P) Open** rule.
  - \_\_\_ b. In the **Properties** tab, expand the **Run rule at start of** node.
  - \_\_\_ c. Select **Pages** and click the plus sign next to it to select the specific page.
  - \_\_\_ d. For the **of type** field, select **Claim Page** from the list.



- \_\_\_ 4. Select the **Read Zones with FPXML** function and from the **FPXML** action library, add the **ReadZonesFPX** action.

This action loads position information for the fingerprint.

- \_\_\_ 5. Select the **Read Zones with FPXML** function and from the **SharedRecognitionTools > SharedRecognitionTools.Actions** action library, add the **SnapCCOtoDCO** action.
- This action maps the page recognition results (field values) from CCO file into DCO fields based on the zone position of the fields.
- \_\_\_ 6. Click **Save changes** and verify that the **Recognize** ruleset contains two rules, functions, and actions as shown in the screen capture.



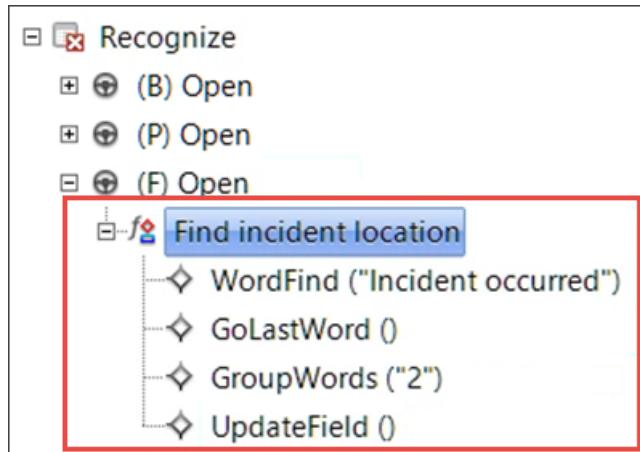
### 3.3. Configure a rule for field extraction by using keyword search

In the previous section, you configured the data extraction by using zones. In this section, you configure the Recognize ruleset to extract a field value by using keyword search.

- \_\_\_ 1. Select **Recognize** and click **Add child object** (plus) icon.  
A new rule node is added to **Recognize** with the name **Rule1**.
- \_\_\_ 2. Rename the following objects:
  - **Rule1:(F) Open**
  - **Function1: Find incident location**
- \_\_\_ 3. Assign the rule to the Incident field.
  - \_\_\_ a. In the **Rulesets** pane, select the **Recognize > (F) Open** rule.
  - \_\_\_ b. In the **Properties** tab, expand the **Run rule at start of** node.
  - \_\_\_ c. Select **Fields** and click the plus sign next to it to select the specific field.
  - \_\_\_ d. For the **of type** field, select **Incident** from the list.
  - \_\_\_ e. Click the plus sign next to **Fields** again.
  - \_\_\_ f. From the list, select the **like** option and select **Incident**.



- \_\_\_ 4. Add an action to the Find incident location function.
  - \_\_\_ a. Select the **Find incident location** function and from the **Locate** action library, add the **WordFind** action.
  - \_\_\_ b. In the **Rulesets** pane, for the **Find incident location** function, verify that the **WordFind** ("") action is added and select the action.
  - \_\_\_ c. In the **Properties** tab, set the **string StrParam** field to: `Incident occurred`  
This action searches the page for the specified word or phrase and remembers the location to use it for the subsequent actions.
- \_\_\_ 5. From the **Locate** action library, add the following actions to the **Find incident location** function.
  - **GoLastWord ()**  
Based on the location that the WordFind action identifies, this action moves the location position to the last word on the current line.
  - **GroupWords ("2")**  
This action groups the words to the left and right of a located word that are separated by the specified character width or less.
  - **UpdateField ()**  
This action updates the DCO field with the recognized data (located word or phrase).
- \_\_\_ 6. Click **Save changes** and verify that the **(F) Open** rule contains the function, and actions as shown in the screen capture.



- \_\_\_ 7. Select the **Recognize** ruleset, click the lock icon on the toolbar of the **Rulesets** tab, and click **Publish ruleset** from the list.
- \_\_\_ 8. Select the **Recognize** ruleset, click the **Move up** (green upward arrow) icon on the toolbar to move up the ruleset, and place it between **CreateDocs&Fields** and **Export** in the list.

### 3.4. Configure the Validate ruleset

In this section, you build the Validate ruleset for the Profiler task.

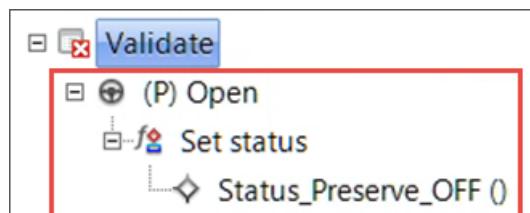
- \_\_\_ 1. Select **ClaimForm** and click **Add child object** (plus) icon.
- \_\_\_ 2. Verify that a new ruleset node is added to the **ClaimForm** node with the name **Ruleset1** and it has a red x to indicate that it is locked for editing.
- \_\_\_ 3. Rename the following objects:
  - **Ruleset1**: Validate
  - **Rule1**: (P) Open
  - **Function1**: Set status
- \_\_\_ 4. Assign the rule to Claim Page.
  - \_\_\_ a. In the **Rulesets** pane, select the **Validate > (P) Open** rule.
  - \_\_\_ b. In the **Properties** tab, expand the **Run rule at start of** node.
  - \_\_\_ c. Select **Pages** and click the plus sign next to it to select the specific page.
  - \_\_\_ d. For the **of type** field, select **Claim Page** from the list.



- \_\_\_ 5. To the **Set status** function, from the **runner** action library, add the **Status\_Preserve\_OFF** action.

This action sends the verification step to human verification.

- \_\_\_ 6. Click the **Save changes** icon and verify the ruleset.



### 3.5. Add field level rules to the Validate ruleset for checking the fields

In this section, you add field level rules to the Validate ruleset. The rules contain a function with actions to validate the fields for data integrity.

- \_\_\_ 1. Add a rule to the Validate ruleset to check the first field: Name
  - \_\_\_ a. Select **Validate** and click **Add child object** (plus) icon.
  - \_\_\_ b. Verify that a new rule node is added to **Validate** with the name **Rule1**.
  - \_\_\_ c. Rename **Rule1** to **(F)Open - Name**

This rule is specific for a field, so runs at field level.

- \_\_ 2. Assign the rule to the Name field.
- \_\_ a. In the **Rulesets** pane, select the **Validate > (F) Open - Name** rule.
  - \_\_ b. In the **Properties** tab, expand the **Run rule at start of** node.
  - \_\_ c. Select **Fields** and click the plus sign next to it to select the specific field.
  - \_\_ d. For the **of type** field, select **Name** from the list.



- \_\_ 3. Add an action to the **Function1**.
- \_\_ a. To **Function1**, from the **ValidationsAndTextAdjustments > Datacap.Libraries.ValidationsAndTextAdjustments.Actions** action library, add the **IsTextLength** action.  
This action checks whether the field has a value.
  - \_\_ b. In the **Rulesets** pane, for **Function1**, verify that the **IsTextLength("", "", "")** action is added and select the action.
  - \_\_ c. In the **Properties** tab, enter values for the following fields:
    - **string Target:** @F
    - **string Length:** 1in
    - **string ComparisonOperator:** >=
 The parameter values check whether the text length of the field value is more than one character.
  - \_\_ d. Click the **Save changes** icon and verify the rule.



- \_\_ 4. Add a rule to the Validate ruleset to check the second field: Policy
- \_\_ a. Select **Validate** and click **Add child object** (plus) icon.
  - \_\_ b. Rename **Rule1** to **(F) Open - Policy**

- \_\_\_ c. Select the **(F) Open - Policy** rule and in the **Properties > Run rule at start of node**, assign it to the **Fields > of type: Policy**



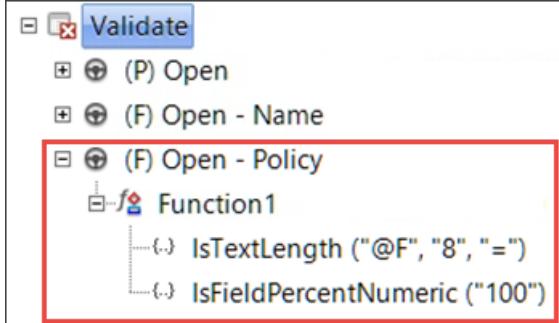
- \_\_\_ d. To **Function1**, from the **ValidationsAndTextAdjustments > Datacap.Libraries.ValidationsAndTextAdjustments.Actions** action library, add the **IsTextLength** action and assign the following parameters:
- **string Target:** @F
  - **string Length:** 8
  - **string Comparison:** =

This action checks if the field value is 8 digits long. In all the scanned images in this application, the Policy value contains 8 characters.

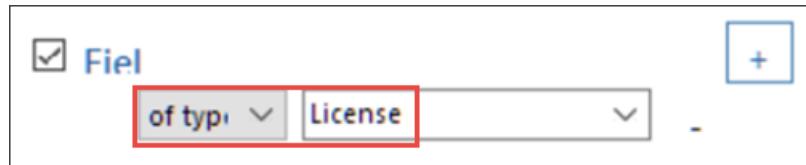
- \_\_\_ e. To **Function1**, from the **ValidationsAndTextAdjustments > Datacap.Libraries.ValidationsAndTextAdjustments.Actions** action library, add another action **IsFieldPercentNumeric** and assign the following parameter:
- **string Percentage:** 100

This action checks if the field value is numeric.

- \_\_\_ f. Click the **Save changes** icon and verify the rule.



5. Add a rule to the Validate ruleset to check the third field: License
- \_\_\_ a. Select **Validate** and click **Add child object** (plus) icon.
  - \_\_\_ b. Rename **Rule1** to **(F) Open - License**
  - \_\_\_ c. Select the **(F) Open - License** rule and in the **Properties > Run rule at start of node**, assign it to the **Fields > of type: License**



- \_\_\_ d. To **Function1**, from the **ValidationsAndTextAdjustments > Datacap.Libraries.ValidationsAndTextAdjustments.Actions** action library, add the **IsTextLength** action and assign the following parameters:

- **string Target:** @F
- **string Length:** 5
- **string Comparison:** >=

This action checks if the value for the field is 5 or more digits long.

- \_\_\_ e. To **Function1**, from the **ValidationsAndTextAdjustments > Datacap.Libraries.ValidationsAndTextAdjustments.Actions** action library, add the **IsTextLength** action second time and assign the following parameters:

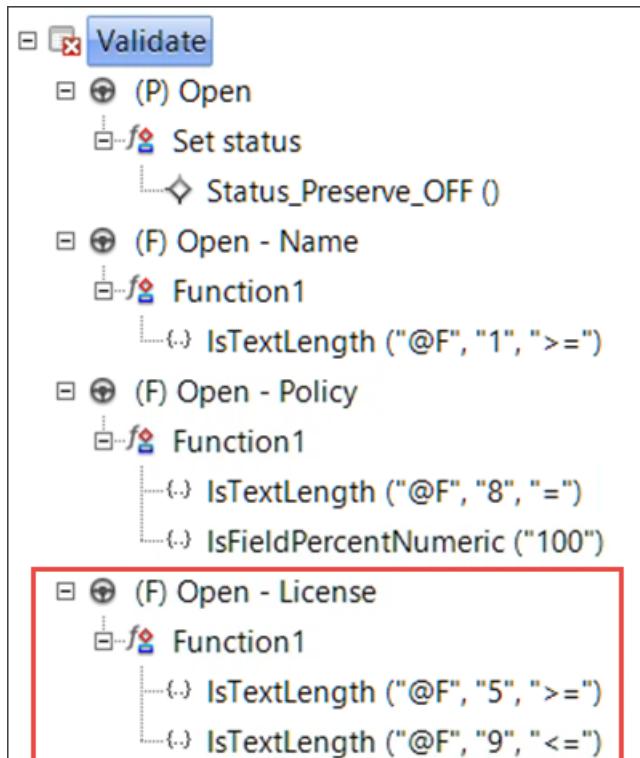
- **string Target:** @F
- **string Length:** 9
- **string Comparison:** <=

This action checks if the value for the field is 9 or less digits long.

The License field can have a value with 6 - 9 digits.

- \_\_\_ 6. Save and publish the ruleset.

- \_\_\_ a. Click the **Save changes** icon.
- \_\_\_ b. Verify that the ruleset contains the rules, functions, and actions as shown in the screen capture.



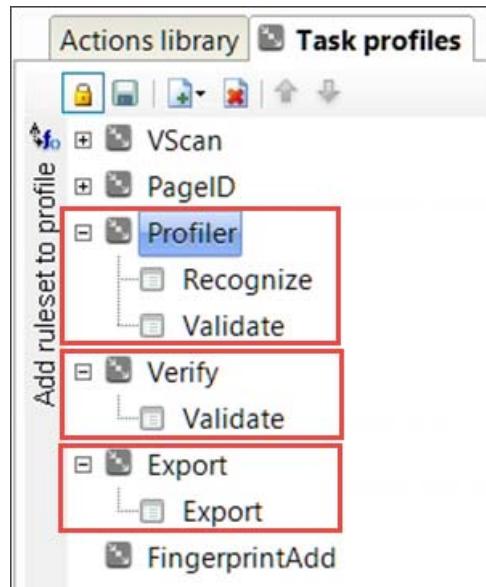
- \_\_\_ c. Select the **Validate** ruleset, on the **Rulesets** toolbar, click the lock icon, and click **Publish ruleset** from the list.
- \_\_\_ d. Select the **Validate** ruleset, click the **Move up** (green upward arrow) icon on the toolbar to move up the ruleset, and place it between **Recognize** and **Export** in the list.

### 3.6. Add rulesets to the Profiler, Validate, and Export task profiles

- \_\_\_ 1. On the upper right pane, click the **Task profiles** tab, lock it by clicking the lock  icon.
- \_\_\_ 2. Select the **Profiler** task profile.
- \_\_\_ 3. On the **Rulesets** tab, select the **Recognize** ruleset and on the right of the **Rulesets** pane, click the **Add ruleset to profile** vertical bar.
- \_\_\_ 4. Select the **Validate** ruleset and add it to the **Profiler** task profile.
- \_\_\_ 5. On the **Task profiles** tab, expand the **Profiler** task profile and verify that it contains the two rulesets: **Recognize** and **Validate**
- \_\_\_ 6. Add the **Validate** ruleset again to the **Verify** task profile.
- \_\_\_ 7. On the **Task profiles** tab, expand the **Verify** task profile and check that it contains the **Validate** ruleset.
- \_\_\_ 8. Add the **Export** ruleset to the **Export** task profile.

The application wizard created the Export ruleset. You can expand it to see the actions that are added to it.

- \_\_\_ 9. On the **Task profiles** tab, expand the **Export** task profile and check that it contains the **Validate** ruleset.
- \_\_\_ 10. On the **Task profiles** tab, click the **Save changes**  icon.
- \_\_\_ 11. Verify that the rulesets are added to the task profiles in the order as shown in the following screen capture.

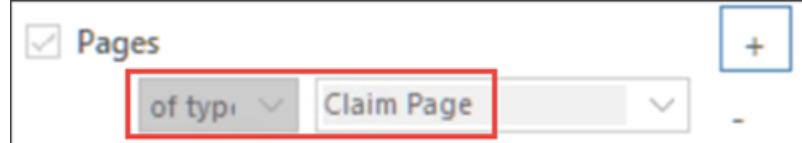


- \_\_\_ 12. Click the lock icon to unlock **Task profiles**.

### 3.7. Configure the Export task

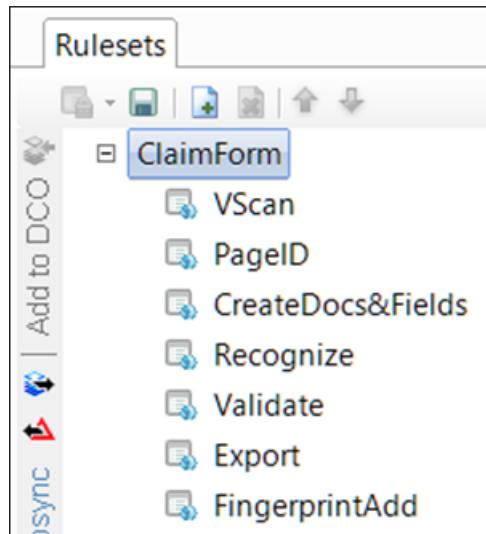
You added the Export ruleset to the task profile in the previous section. In this section, you check the ruleset and make any necessary changes.

- \_\_\_ 1. On the **Rulesets** tab, select the **Export** ruleset, click the lock icon, and from the list, click **Lock ruleset for editing**.
- \_\_\_ 2. Check the DCO association for the Set Export Params rule.
  - \_\_\_ a. In the **Rulesets** pane, expand **Export** and select the **Set Export Params** rule.
  - \_\_\_ b. In the **Properties** tab, expand the **Run rule at start of** node.
  - \_\_\_ c. Verify that the **batches** option is selected and if it is not already selected, select it.
- \_\_\_ 3. Check the DCO association for the Export Page Fields rule.
  - \_\_\_ a. In the **Rulesets** pane, select the **Export > Export Page Fields** rule.
  - \_\_\_ b. In the **Properties** tab > **Run rule at start of** node, verify that the **Pages** option is selected and it has an entry below that.
  - \_\_\_ c. If no entry is found, click the plus sign (+) next to **Pages** to create an entry.
  - \_\_\_ d. From the list, select **of type** and select **Claim Page**.



- \_\_\_ 4. Click the **Save changes** icon in the toolbar.
- \_\_\_ 5. Select the **Export** ruleset, click the lock icon, and click **Publish ruleset**.

For your reference, final list of rulesets and their order is shown in the following screen capture.

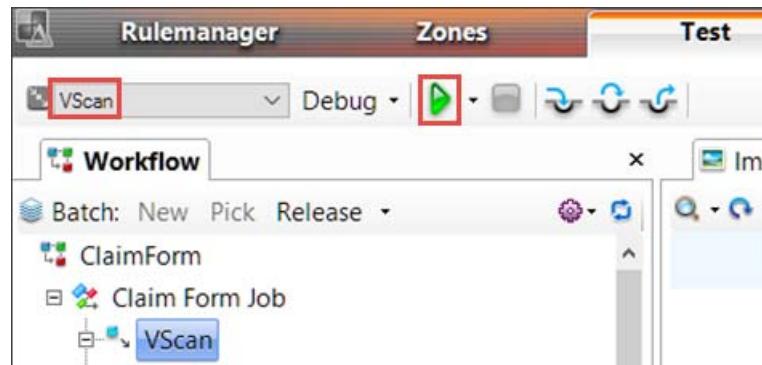


## Section 4. Test the application

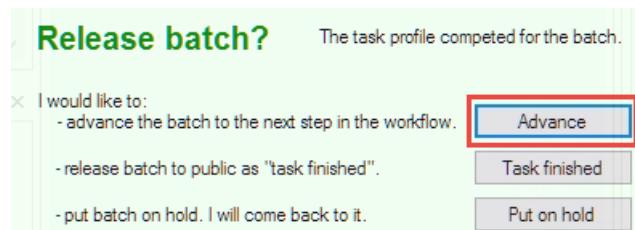
In this section, you start a batch and run all the tasks in the workflow to test the application.

### 4.1. Run the VScan and PageID tasks

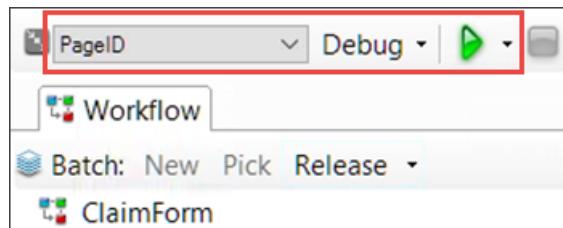
- \_\_\_ 1. In Datacap Studio, click the **Test** tab.
- \_\_\_ 2. Run the VScan task to start a batch.
  - \_\_\_ a. On the **Workflow** tab, expand **ClaimForm > ClaimForm Job**, right-click **VScan**, and select **New**.
  - \_\_\_ b. In the top toolbar, ensure that **VScan** is selected.
  - \_\_\_ c. If it is not already selected, select it from the list and click the **Process rules for target object** (green arrow icon) next to it.



- \_\_\_ 3. In the **Release batch** dialog box, click **Advance** to move the batch to the **PageID** task.



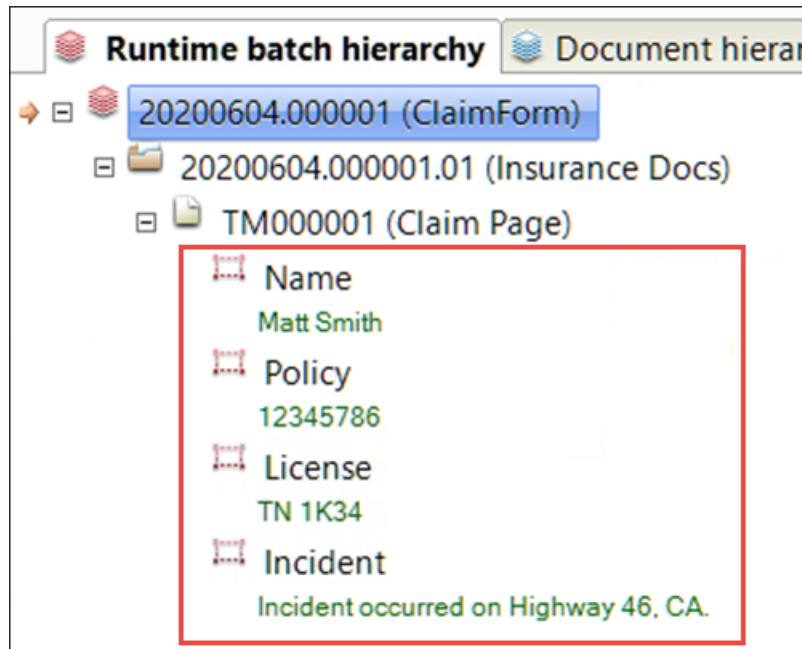
- \_\_\_ 4. In the **Runtime batch hierarchy tab**, verify that it has a batch item now and the page is listed as **Other**.
- \_\_\_ 5. Click **TM000001 (Other)** and the scanned image is shown in the **Image** tab.
- \_\_\_ 6. In the top toolbar, ensure that **PageID** is selected and click the green arrow icon to run the **PageID** task.



- \_\_\_ 7. Verify the results of running the PagelD task.
  - \_\_\_ a. In the **Release batch** dialog box, click **Advance** to move the batch to the **Profiler** task.
  - \_\_\_ b. In the **Runtime batch hierarchy tab**, expand the batch node and verify that the document is classified with the name: **Insurance Docs**
  - \_\_\_ c. Expand **Insurance Docs** and verify that the page is identified as **Claim Page**.
  - \_\_\_ d. Expand **Claim Page** and verify that the page contains the four fields that you defined: **Name, Policy, License, and Incident**.
  - \_\_\_ e. Click **Claim Page** and check that the scanned image is displayed in the middle pane.

## 4.2. Run the Profiler task

- \_\_\_ 1. In the top toolbar, ensure that **Profiler** is selected and click the green arrow icon to run the **Profiler** task.  
This step extracts the data from the fields.
- \_\_\_ 2. In the **Release batch** dialog box, click **Keep running**.
- \_\_\_ 3. In the **Runtime batch hierarchy tab**, expand the batch nodes to verify that the values for these fields: **Name, Policy, License, and Incident**

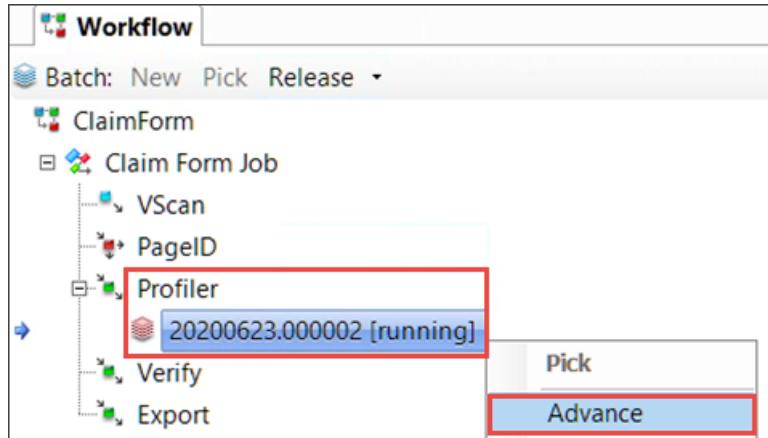


- \_\_\_ 4. Click a field (for example, **Name**), the image is shown in the **Image** tab on the middle pane and the fields Zones are highlighted.
- \_\_\_ 5. Optionally, compare the extracted values with the values on the scanned image in the C:\Datacap\ClaimForm\images folder.

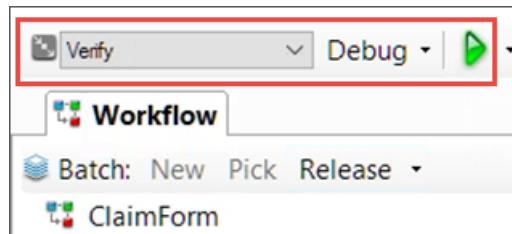
### 4.3. Advance the batch to Verify task

- On the **Workflow** tab, expand **ClaimForm > ClaimForm Job > Profiler**, right-click the batch (**running**), and select **Advance**.

The workflow moves to the Verify step.



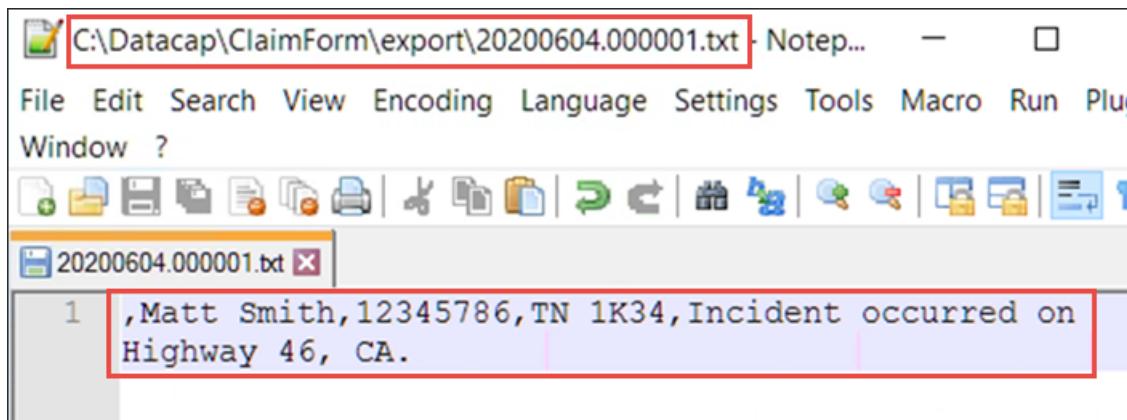
- In the top toolbar, ensure that **Verify** is selected and click the green arrow icon to run the **Verify** task.



### 4.4. Process the Export task and verify the export file

- In the **Release batch** dialog box, click **Advance** to move the batch to the **Export** task.
- In Datacap Studio **Test** tab, on the **Runtime batch hierarchy** pane, take a note of the batch name: <Date>.<batch number> (Example: 20200604.000001).  
You use this value to identify the export file in the following steps.
- In the top toolbar, ensure that **Export** is selected and click the green arrow icon to run the task.
- In the **Release batch** dialog box, click **Advance** to complete the batch.
- Check the export text file.
  - Open Windows Explorer, browse to the `C:\Datacap\ClaimForm\export` folder and verify that a text file with your batch name (that you noted in step 2) is created with today's date.

- \_\_\_ b. Open the file in Notepad++ and check that the file contains the values for the four fields (Name, policy number, license, and incident location).



The values vary depending on which image was used for scanning.

You can configure the application to format these values the way you want.

- \_\_\_ c. Close the file.

## 4.5. Test the application with scanning multiple images in a batch

In the previous section, you configured the application to scan up to five images in a single batch. In this section, you copy five images to the image folder for scanning and test the application.

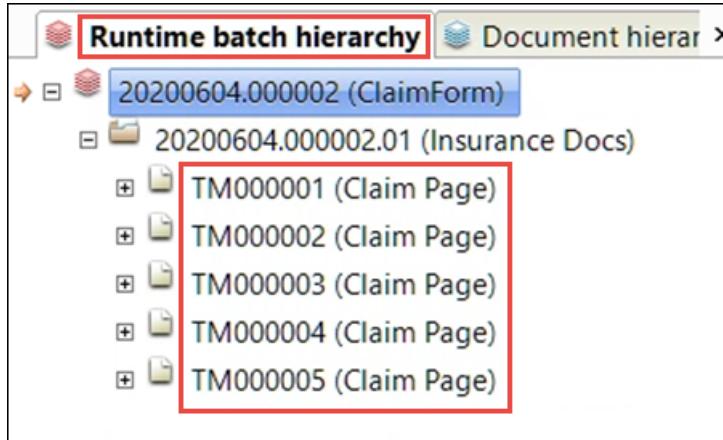
- \_\_\_ 1. In Windows Explorer, copy the following files from the C:\labfiles\Ex.3.Images\_for\_scanning folder to the C:\Datacap\ClaimForm\images folder.
  - **2-ClaimForm-BC.tif**
  - **3-ClaimForm-BC.tif**
  - **4-ClaimForm-BC.tif**
  - **5-ClaimForm-BC.tif**

The field values are different but all the forms have the same layout.

- \_\_\_ 2. Follow the instructions in the sections 4.1 to 4.4 of the [Section 4, "Test the application"](#) to run a batch and test the application.

Compare the results with the following screen captures and descriptions.

- \_\_\_ a. After you run the **PageID** task, in the **Runtime batch hierarchy tab**, verify that five pages are created (one for each image scanned).



- \_\_\_ b. After you complete the **Profiler** task, expand the batch (**ClaimForm**) > document (**Insurance Docs**) and expand each page (**Claim Page**) node, and verify that the four fields (**Name**, **Policy**, **License**, and **Incident**) are populated with values.
- \_\_\_ 3. After you complete the **Export** task, check the export file.
  - \_\_\_ a. In Windows Explorer, browse to the `C:\Datacap\ClaimForm\export` folder and verify that a text file with your latest batch name is created with today's date.
  - \_\_\_ b. Open the file in Notepad++ and check that the file contains five sets of values for the four fields (Name, policy, license, and incident).

```

1 ,Matt Smith,12345786,TN 1K34,Incident occurred on
Highway 46, CA.
2 ,Ava Crimson,56789124,4457YL,Incident occurred on
Highway 23, OK.
3 ,Ava Crimson,56789124,4457YL,Incident occurred on
Highway 23, OK.
4 ,Chris Jones,48998813,G2967R,Incident occurred on
Highway 85, NY.
5 ,James Sydney,47123025,327X54,Incident occurred on
Highway 46, OR.

```



## Troubleshooting

Check the following items:

- Document hierarchy (DCO) structure is defined correctly.
- Each ruleset is configured with rules, functions, and actions exactly as shown in the screen captures.
- Each rule is associated with the correct Datacap object.
- Task profiles have correct the rulesets that are associated with each profile.

## End of exercise

## Exercise review and wrap-up

This exercise showed how to configure rulesets for a Datacap application in Datacap Studio. You also learned how to associate the rules to process specific objects in the document hierarchy, assign the rulesets to specific task profiles, and test the configuration.

# Exercise 4. Troubleshooting a Datacap application

## Estimated time

01:00

## Overview

In this exercise, you become familiar with the log files and the troubleshooting capabilities in Datacap Studio for application development.

## Objectives

After completing this exercise, you should be able to:

- Examine the logs files for debugging
- Troubleshoot the applications in the Datacap Studio Test tab

## Introduction

Datacap Studio provides testing and debugging capabilities for application development. In the previous exercise, you used the Datacap Studio Test tab to test an application during development. In this exercise, you access logs and use the Datacap Studio Test tab to debug an application.

This exercise includes the following sections:

- [Section 1, "Examine the logs files for debugging"](#)
- [Section 2, "Troubleshoot the application in Datacap Studio"](#)
- [Section 3, "Correct the error in the application and test"](#)

## Requirements

This exercise builds on [Exercise 3, "Configuring rulesets for a Datacap application"](#). Before starting this exercise, you must complete [Exercise 3, "Configuring rulesets for a Datacap application"](#).

You built the **ClaimForm** application in the previous exercise. If this application is not available, you can use the solution **SolutionClaimForm** that is provided in C:\Datacap folder. The steps in the exercise instruct you how to use the solution application.

# Section 1. Examine the logs files for debugging

In this section, you learn how to examine log files.

## 1.1. Examine the logs for a batch

In this section, you examine various logs in the ClaimForm application batch folder.

- 1. Before you begin the exercise, ensure that the components are running. If you did not start the components already, refer to [Section 1, "Prepare your system: Start the components"](#) to start them.

If you restart your system, you must start all the components.

- 2. In Windows Explorer, browse to the C:\Datacap\ClaimForm\batches\<batch name> folder.

Datacap places all the output from the batch processing in the application batches folder.



### Troubleshooting

For this exercise, you use the **ClaimForm** application that you built in the previous exercises. If the application is not available, you can use **SolutionClaimForm**.



### Important

For this section, you use the folders and files in the C:\Datacap\ClaimForm\batches\ folder. If your batch folder is empty, you need to run a batch by using the instructions in the Exercise 3, [Section 4, "Test the application"](#)

- 3. Verify that the following runtime logs are created:

- vscan\_rrs.log
- pageid\_rrs.log
- profiler\_rrs.log
- verify\_rrs.log
- export\_rrs.log

C:\Datacap\ClaimForm\batches\20200604.000001		
Name	Date modified	Type
export_rrs.log	6/4/2020 8:16 PM	Text Document
pageid_rrs.log	6/4/2020 7:50 PM	Text Document
profiler_rrs.log	6/4/2020 7:53 PM	Text Document
verify_rrs.log	6/4/2020 8:11 PM	Text Document
vscan_rrs.log	6/4/2020 7:48 PM	Text Document

4. Examine the log file and trace through the processing sequence.

In Datacap Studio, rulesets are associated with a task profile. Each ruleset calls actions. In the log file, you can check the action names for their complete status.

- \_\_ a. Open the **vscan\_rrs.log** file in Notepad++.

The **scan()** action is defined in the **VScan** ruleset and is run for the **VScan** task.

- \_\_ b. If you prefer color coding for the text, click **Language** and click **V > Visual Basic**.
- \_\_ c. Observe the following information on the log file:

The line numbers are based on the sample file in the following screen capture, the line numbers of your log file might be slightly different.

You can search for the keyword (**mvScan**) to find the lines.

- On line 139 at the top, the **scan** action is called.
- On line 144, a previous action (set\_folder) correctly identified the **images** folder that you defined in the application. This folder is where the images are stored for scanning.
- In the following lines, scan finds one file (**1-ClaimForm-BC.tif**) and ingests the file.  
If you open the **vscan\_rrs.log** file for the batch where you processed five images, you find that the information for five files.
- Line 156 shows that the batches folder location where the file is ingested
- Line 158 shows that the file (**page 1**) is ingested

These lines indicate that the VScan task is successful.

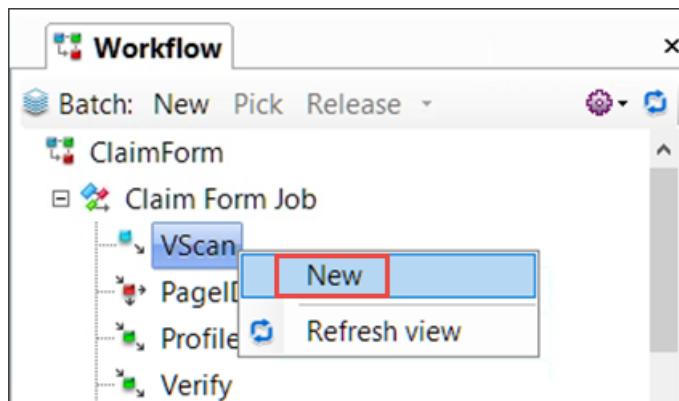
```
vscan_rss.log
138 19:48:30.703 (0) /execute statement On Action Start
139 19:48:30.703 (0) call
140 19:48:30.703 (0) Cache file: C:\Users\Administrator\AppData\Local\Temp\DCmvscan374895062.xml
141 19:48:30.703 (0) Demo mode enabled
142 19:48:30.703 (0) Refresh cache
143 19:48:30.703 (0) Scan folder:C:\Datacap\ClaimForm\images
144 19:48:30.704 (0) Scan folder: C:\Datacap\ClaimForm\images
145 19:48:30.704 (0) Found 1 files
146 19:48:30.704 (0) Check file 1-ClaimForm-BC.tif modified 5/26/2020 6:33:25 PM
147 19:48:30.704 (0) Add file to list: 1-ClaimForm-BC.tif input folder: C:\Datacap\ClaimForm\images
148 19:48:30.704 (0) Found 0 folders
149 19:48:30.705 (0) Wrote cached file list to C:\Users\Administrator\AppData\Local\Temp\DCmvscan374895062.xml
150 19:48:30.705 (0) Root folders contain 1 selected files and folders
151 19:48:30.705 (0) Try to ingest 1 files and folders
152 19:48:30.705 (0) Start ingest file 0 of 1
153 19:48:30.705 (0) Ingest file C:\Datacap\ClaimForm\images\1-ClaimForm-BC.tif
154 19:48:30.705 (0) Ingesting C:\Datacap\ClaimForm\images\1-ClaimForm-BC.tif
155 19:48:30.705 (0) Demo mode enabled
156 19:48:30.706 (0) Ingest to C:\Datacap\ClaimForm\batches\20200604.000001\TM000001.tif
157 19:48:30.706 (0) Demo mode enabled
158 19:48:30.709 (0) Ingested page 1
159 19:48:30.710 (0) Deleted lock file C:\Datacap\ClaimForm\images\1-ClaimForm-BC.tif.lock
160 19:48:30.710 (0) Finished batch. Cached 1 filenames, next 1
```

## 1.2. Empty the images folder and run a batch

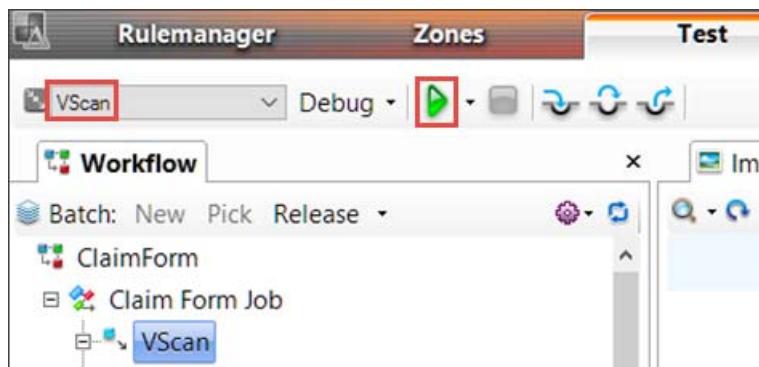
In the previous section, you examined a log file where VScan ran successfully. In this section, you remove all the images from the images folder from where the images are scanned, run the ClaimForm application, and examine the logs in the batch folder.

- 1. Back up the images and empty the images folder.
  - a. Open Windows Explorer, browse to the C:\Datacap\ClaimForm folder, and create a subfolder with the name: images\_backup
  - b. From the C:\Datacap\ClaimForm\images folder, copy all the images and paste it to the C:\Datacap\ClaimForm\images\_backup folder.
  - c. In the C:\Datacap\ClaimForm\images folder, delete all the files.
- 2. If the Datacap Studio is not already open, start it.
  - a. Double-click the **Datacap Studio** shortcut on the Windows desktop.
  - b. In the **Applications** dialog box, select **ClaimForm** and click **Next**.

- \_\_\_ c. On the **Taskmaster Login** page, enter the following values:
    - **User ID:** admin
    - **Password:** admin
    - **Station ID:** 1
  - \_\_\_ d. Click **Finish**.
- \_\_\_ 3. Start a batch.
- \_\_\_ a. In Datacap Studio, click the **Test** tab from the top toolbar.
  - \_\_\_ b. On the **Workflow** tab in the upper left pane, expand **ClaimForm > ClaimForm Job**, right-click **VScan**, and select **New**.



- \_\_\_ c. In the top toolbar, ensure that **VScan** is selected.
- \_\_\_ d. If it is not already selected, select it from the list and click the **Process rules for target object** (green arrow icon) next to it.

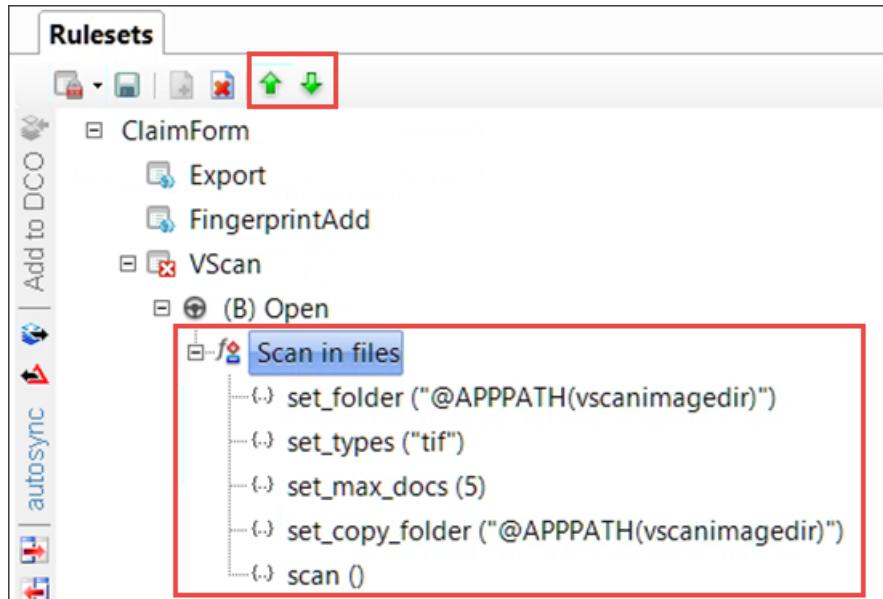


Wait for the process to complete.

- \_\_\_ 4. In the **Release batch** dialog box, click **Put on hold** and note down the batch number.
- \_\_\_ 5. Leave Datacap Studio open for the next section.

### 1.3. Examine the logs

- 1. In Windows Explorer, browse to the C:\Datacap\ClaimForm\ batches\ folder and open the latest batch folder that you processed.  
Use the batch number that you noted down in the previous section.
- 2. Open the **vscan\_rrs.log** file in Notepad++.
- 3. In Datacap Studio Rulemanager tab, observe the actions that you defined for the VScan ruleset.



- 4. In Notepad++, search for each action and check how the actions are run.
  - The **set\_folder** action

The action is run successfully as indicated by the text: action returned true and Image source folder set to: C:\Datacap\ClaimForm\images

```

61 04:04:26.476 (94) C:\Datacap\dcshared\NET\mvscan.dll.
AssemblyVersion: 9.1.7.0. AssemblyFileVersion: 9.1.7.41.
62 04:04:26.491 (16) call
"Datacap.Libraries.mvScan.Actions->Datacap.Libraries.mvScan.Actions->set_folder"
63 04:04:26.503 (15) DCSmart Metaword sArgument: @APPPATH(
vscanimagedir)
64 04:04:26.503 (0) 1 Smart Parameter element found
65 04:04:26.503 (0) Parsing Smart Parameter element {0} value:
"@APPPATH(vscanimagedir)"
66 04:04:26.503 (0) matched '@' index 0
67 04:04:26.503 (0) sArg: @APPPATH.
68 04:04:26.503 (0) @APPPATH key root value: 'vscanimagedir'
69 04:04:26.504 (0) @APPPATH looking for workflow key:
'*/dco_ClaimForm/vscanimagedir'
70 04:04:26.505 (0) workflow key found:
'C:\Datacap\ClaimForm\images'
71 04:04:26.505 (0) Smart return: 'C:\Datacap\ClaimForm\images'
72 04:04:26.505 (0) Image source folder set to: C:\Datacap\
ClaimForm\images
73 04:04:26.505 (0) result 0[x0] = true
74 04:04:26.505 (0) /call
75 04:04:26.505 (0) action returned true
76 04:04:26.505 (0) execute statement On Action
True

```

- The **set\_types** action

The action ran successfully as indicated by the text: file types to be ingested: tif

- The **set\_copy\_folder** action

The action ran successfully as indicated by the text: Copy folder set to:  
C:\Datacap\ClaimForm\images

- The **scan** action

The action did not run successfully as indicated by the text: Found 0 files, Found 0 folders and No images to process

The message indicates that no files found in the image folder and no files were ingested.

## 1.4. Copy the files back to the images folder

In the previous section, you deleted the files in the image folder to examine a log file for troubleshooting. In this section, you place all the images back in the images folder.

- \_\_ 1. Open Windows Explorer, browse to the C:\Datacap\ClaimForm\images\_backup folder, and copy the **1-ClaimForm-BC.tif** file that you backed up.
- \_\_ 2. Paste the file to the C:\Datacap\ClaimForm\images folder.

## Section 2. Troubleshoot the application in Datacap Studio

In the previous exercise, you learned how to use the Test tab in Datacap Studio. This tab is used to debug and test the applications. In this section, you step through the task profiles and observe the status of actions.

### 2.1. Edit a ruleset to introduce an error

In this section, you edit the CreateDocs&Fields ruleset to introduce an error to use it for debugging.

- \_\_\_ 1. In Datacap Studio **Test** tab, right-click the batch that was on **hold** in **VScan** and click **Delete**.  
You used this batch for the previous section..
- \_\_\_ 2. Click the **Rulemanager** view.
- \_\_\_ 3. On the **Rulesets** tab, select the **CreateDocs&Fields** ruleset, click the lock icon on the toolbar, and click **Lock ruleset for editing** from the list.
- \_\_\_ 4. Remove the association of DCO for **(P) Open** rule.
  - \_\_\_ a. In the **Rulesets** pane, expand **CreateDocs&Fields** and select the **(P) Open** rule.
  - \_\_\_ b. In the **Properties** tab, expand the **Run rule at start of** node.
  - \_\_\_ c. Verify that the **Pages** option is selected and it has an entry below that: **of type: Claim Page**
  - \_\_\_ d. Click the minus sign (-) next to the entry to remove it.
  - \_\_\_ e. Clear the selection for the **Pages** and verify that nothing is selected.  
The removal causes the actions in the rule not to run. You use this scenario to troubleshoot.
- \_\_\_ 5. In the **Rulesets** tab toolbar, click the **Save changes** icon.
- \_\_\_ 6. Select the **CreateDocs&Fields** ruleset, click the lock icon, and click **Publish ruleset** from the list.

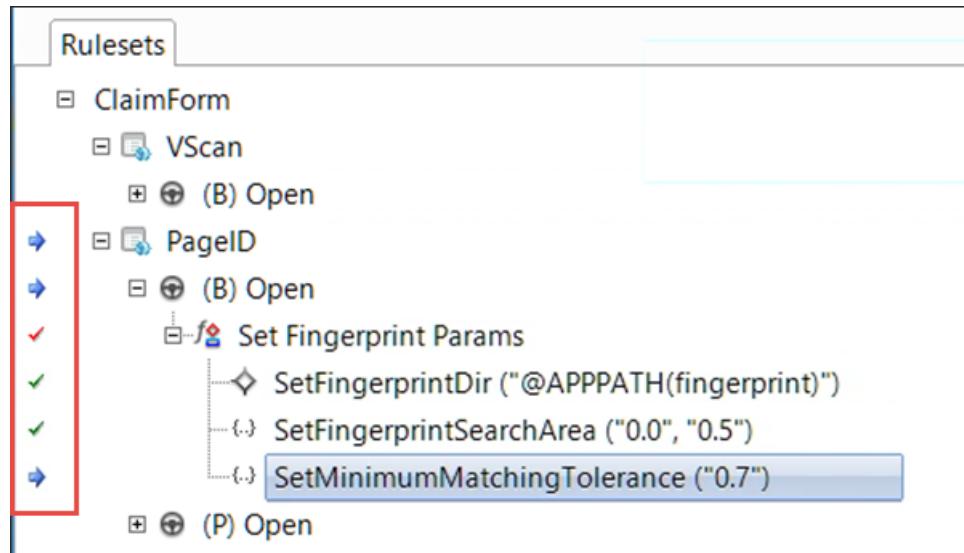
### 2.2. Start a batch to test the application

- \_\_\_ 1. In Datacap Studio, click the **Test** tab from the top toolbar.
- \_\_\_ 2. On the **Workflow** tab in the upper left pane, expand **ClaimForm > ClaimForm Job** right-click **VScan**, and select **New**.
- \_\_\_ 3. In the top toolbar, ensure that **VScan** is selected.
- \_\_\_ 4. If it is not already selected, select it from the list and click the **Process rules for target object** (green arrow icon) next to it.
- \_\_\_ 5. In the **Release batch** dialog box, click **Advance** to move the batch to the **PageID** task.

## 2.3. Step the batch through PageID

- 1. In the **Workflow** tab, expand the **PageID** task and notice that the batch is **running**.
- 2. On the toolbar, click the **Step in**  icon and on the lower left pane, in the **Rulesets** tab, observe that a blue arrow points to **PageID**.
- 3. Click the **Step in** icon some more times and after each click, notice the changes in the **Rulesets** tab.

The blue arrows points to the ruleset, rule, and action that is running. The completed actions show a green tick mark. The function is not yet completed and shows red tick mark.

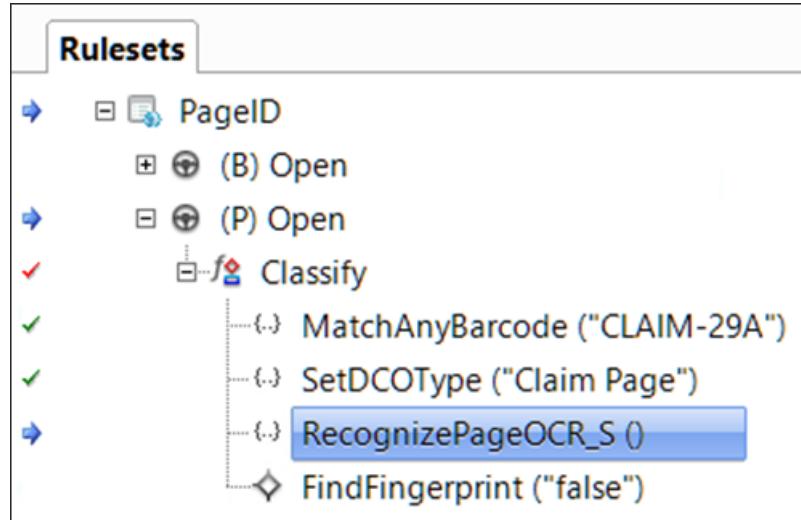


- 4. On the left middle pane, in the **Runtime batch hierarchy** tab, notice that the batch object is selected because the batch level rule is running.



- 5. Continue to click the **Step in**  icon until the **Runtime batch hierarchy** indicates the **TM000001 (Other)** page is selected.
- 6. Notice that the image is shown in the **Image** tab of the middle pane.
- 7. Continue to click the **Step in** icon a few more times.

Because the **PageID > (P) Open** rule is assigned at the page level, Datacap expands this rule > the **Classify** function and runs the **MatchAnyBarcode ("CLAIM-29A")** and other actions.



8. Complete the Classify function.
- \_\_\_ a. Continue to click the **Step in** icon a few more times and observe the **Rulesets** tab until all the actions in **Classify** are run.
  - \_\_\_ b. In the **Runtime batch hierarchy** tab, notice that the page is now identified as **Claim Page**.



- \_\_\_ c. Continue to click the **Step in** icon and observe the **Rulesets** tab until the blue arrow moves to **CreateDocs&Fields > (B) Open** rule.  
It takes several clicks.
9. Step the batch through the PagelD.
- \_\_\_ a. Continue to click the **Step in** icon.
  - \_\_\_ b. When it completes the **CreateDocuments()** action in the **(B) Open** rule, notice that it skips the **(P) Open** rule.  
Because you edited the rule to remove the association with DCO object, the actions in that rule is not run.
  - \_\_\_ c. In the **Runtime batch hierarchy** tab, notice that it creates the **Insurance Docs** document.
  - \_\_\_ d. Continue to click the **Step in** icon and observe the **Rulesets** and **Runtime batch hierarchy** tabs until the **Release batch** dialog box is shown.  
It takes several clicks.

- \_\_\_ 10. Examine the creation of fields.
- \_\_\_ a. In the **Release batch** dialog box, click **Keep running**.
- \_\_\_ b. In the **Runtime batch hierarchy** pane, notice that the step created the document and page objects, but not the fields.



For this scenario, assume that the actions for the PageID are configured correctly. This step suggests that the **(P) Open** rule was not associated with the Pages DCO and so it did not create fields.

## 2.4. Cancel the batch

- \_\_\_ 1. On the **Workflow** tab in the upper left pane, expand **ClaimForm > ClaimForm Job> PageID**.
- \_\_\_ 2. Right-click the running batch and click **Cancel**.

This step cancels the running batch. You edit the application and test it by starting a new batch.

## Section 3. Correct the error in the application and test

In the previous section, you introduced an error for debugging purposes. In this section, you correct the error and test the application.

### 3.1. Edit a ruleset to correct the error

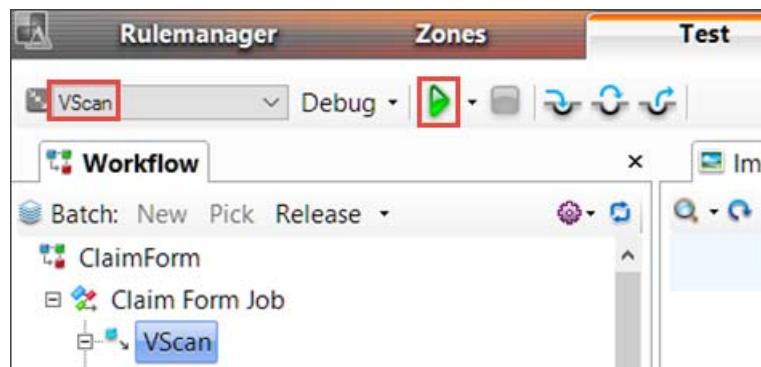
In this section, you edit the CreateDocs&Fields ruleset to correct the error.

- \_\_\_ 1. In Datacap Studio, click the **Rulemanager** view.
- \_\_\_ 2. On the **Rulesets** tab, select the **CreateDocs&Fields** ruleset, click the lock icon, and click **Lock ruleset for editing**.
- \_\_\_ 3. Associate the (P) Open rule to a page.
  - \_\_\_ a. In the **Rulesets** pane, expand the **CreateDocs&Fields > (P) Open** rule.
  - \_\_\_ b. In the **Properties** tab, expand the **Run rule at start of** node.
  - \_\_\_ c. Select **Pages** and click the plus sign next to it to select the specific page.
  - \_\_\_ d. For the **of type** field, select **Claim Page** from the list.
- \_\_\_ 4. Click the **Save changes** icon.
- \_\_\_ 5. Select the **CreateDocs&Fields** ruleset, click the lock icon, and click **Publish ruleset**.

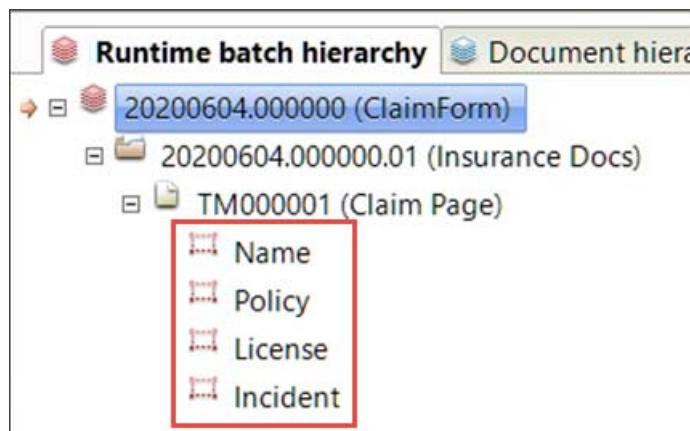
### 3.2. Test the application

In this section, you run the batch again to check whether the Page ID is creating the fields.

- \_\_\_ 1. Start a batch to test the application.
  - \_\_\_ a. In Datacap Studio, in the top toolbar, click the **Test** tab.
  - \_\_\_ b. On the **Workflow** tab, expand **ClaimForm > ClaimForm Job**, right-click **VScan**, and select **New**.
  - \_\_\_ c. In the top toolbar, ensure that **VScan** is selected and click the **Process rules for target object** (green arrow icon) next to it.



- \_\_\_ 2. Run the Page ID task.
  - \_\_\_ a. In the **Release batch** dialog box, click **Advance** to move the batch to the **PageID** task.
  - \_\_\_ b. In the top toolbar, ensure that **PageID** is selected and click the green arrow icon again to run the **PageID** task.
- \_\_\_ 3. Verify the results of running the PageID task.
  - \_\_\_ a. In the **Release batch** dialog box, click **Advance** to move the batch to the **Profiler** task.
  - \_\_\_ b. In the **Runtime batch hierarchy tab**, expand the batch node and verify that the document is classified with the name: **Insurance Docs**
  - \_\_\_ c. Expand **Insurance Docs** and verify that the page is identified as **Claim Page**.
  - \_\_\_ d. Expand **Claim Page** and verify that the page contains the fields that you defined: **Name**, **Policy**, **License**, and **Incident**.



This step confirms that the **(P) Open** rule is associated with the Pages DCO and so it creates fields.

- \_\_\_ 4. On the **Workflow** tab, expand **ClaimForm > ClaimForm Job > Profiler** right-click the running batch item and click **Cancel** from the list.
- \_\_\_ 5. At the upper right corner of Datacap Studio, click **Exit**.



### Information

You can right-click and select Set breakpoint to set breakpoints for debugging the application.

## End of exercise

## Exercise review and wrap-up

This exercise showed how to examine the logs files for debugging and troubleshoot the applications in the Datacap Studio Test tab.

# Exercise 5. Building an application to process multiple page types

## Estimated time

02:15

## Overview

In this exercise, you learn how to build a Datacap application for processing multiple documents of different page types in a single batch.

## Objectives

After completing this exercise, you should be able to:

- Create an application by using an existing application
- Configure the application in IBM Content Navigator
- Define document hierarchy and configure rulesets
- Configure fingerprinting and zones

## Introduction

In this exercise, you copy an existing application that is based on the Forms template to create a new application. You configure the new application for processing multiple page types in a batch.

This exercise includes the following sections:

- [Section 1, "Copy a Datacap application"](#)
- [Section 2, "Configure the new application in Datacap Navigator"](#)
- [Section 3, "Verify the application in Datacap Application Manager"](#)
- [Section 4, "Define Document hierarchy"](#)
- [Section 5, "Configure rulesets and task profiles"](#)
- [Section 6, "Configure Fingerprints and Zones"](#)
- [Section 7, "Test the application"](#)

## Requirements

Before starting this exercise, you must complete [Exercise 3, "Configuring rulesets for a Datacap application."](#) You built the **ClaimForm** application in the previous exercise. If this application is not available, you can use the solution **SolutionClaimForm** that is provided in C:\Datacap folder. The steps in the exercise instruct you how to use the solution application.

## Section 1. Copy a Datacap application

In the previous exercise, you built a new Datacap application by using a template to understand various components that makes up an application. In this exercise, you copy the existing application to create a new one to make the application development faster. You customize the application to process multiple page types in one batch.



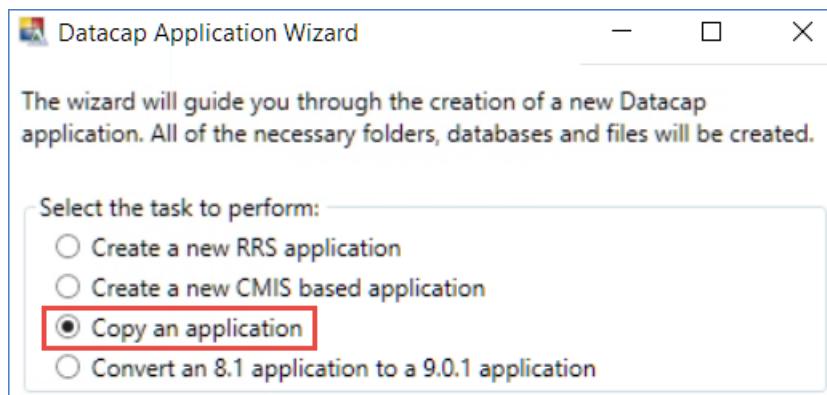
### Note

A solution of the Datacap application for this exercise is provided in C:\Datacap folder with the application name: **SolutionMultiPageApp**

You can access this application like any other Datacap application.

### 1.1. Copy a Datacap application

- 1. Before you begin the exercise, ensure that the components are running. If you did not start the components already, refer to [Section 1, "Prepare your system: Start the components"](#) to start them.  
If you restart your system, you must start all the components.
- 2. Start Datacap Studio without opening any application.
  - a. On the Windows desktop, double-click the **Datacap Studio** shortcut.
  - b. In the **Applications** dialog box, click **Close**.  
Datacap Studio opens with the **Rulemanager** tab selected.
- 3. On the upper right of the toolbar, click the **Datacap application wizard**  icon.
- 4. In **Datacap Application Wizard**, select the **Copy an application** option.

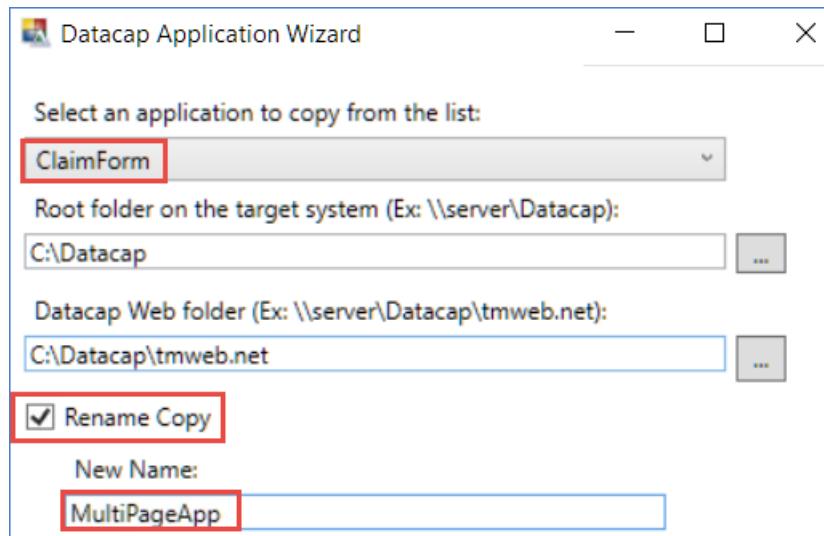


- 5. Click **Next** and for the **Select an application to copy from the list** field, select **ClaimForm**.

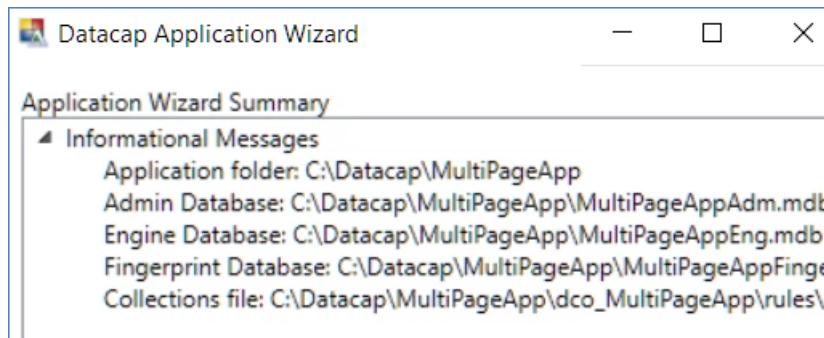
**Hint**

You built the **ClaimForm** application in the previous exercise. If this application is not available, you can use **SolutionClaimForm**.

- \_\_\_ 6. Select the **Rename Copy** option and for the **New Name** field, enter: **MultiPageApp**  
Leave the default folder values for the **Root folder on the target system** (**C:\Datacap**) and for the **Datacap Web folder** (**C:\Datacap\tmweb.net**).



- \_\_\_ 7. Click **Next** and on the next page, click **Finish**.
- \_\_\_ 8. Review the Application wizard summary for a list of folders, files, and databases.



- \_\_\_ 9. Click **Close** to close the wizard and on the upper right corner of the page, click **Exit** to close Datacap Studio.
- \_\_\_ 10. Check the entry for the application folder in the **Datacap.xml** file.
  - \_\_\_ a. Open Windows Explorer, browse to the **C:\Datacap** folder, and open the **Datacap.xml** file in NotePad++.
  - \_\_\_ b. At the end of the page, verify that an entry is added for the new application (**MultiPageApp**).

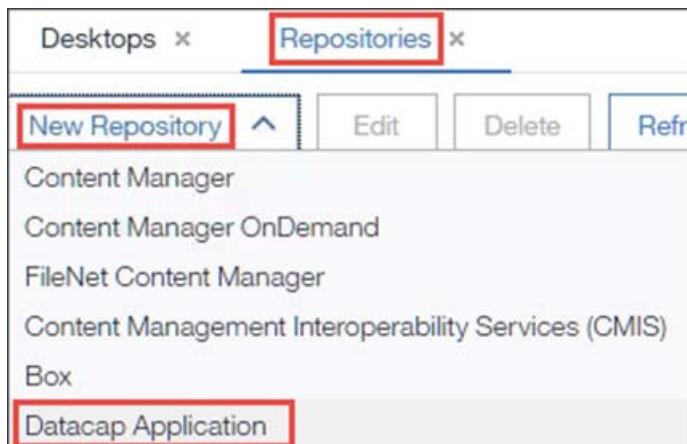
- \_\_ c. Close the file and close NotePad++.
- \_\_ d. In Windows Explorer, in the C:\Datacap folder, verify that a folder with the name of the new application (**MultiPageApp**) is created.

## Section 2. Configure the new application in Datacap Navigator

In this section, you configure the new application as a repository in Datacap Navigator administration tool. You associate the repository with the desktop so that the new application is accessible in Datacap Navigator client.

### 2.1. Add the new Datacap application as a repository

- 1. Start the IBM Content Navigator administration tool.
  - a. Open the Firefox browser and click the **ICN Admin** shortcut. Or go to the **URL**:  
`http://ecmedu01:9081/navigator/?desktop=admin`
  - b. Enter the following values:
    - **User name:** p8admin
    - **Password:** FileNet1
  - c. Click **Log In**.
- 2. Open the New Repository tab.
  - a. On the left pane, click **Repositories**.
  - b. On the **Repositories** tab, click **New Repository** and from the list, select **Datacap Application**.



- 3. On the **New Repository** tab > **General** subtab, set the values for the following properties:

Field	Value
Display name	<b>MultiPageApp</b>
Datacap wTM URI	<code>http://ecmedu01:85/ServicewTM.svc</code>
Application	<b>MultiPageApp</b>
Default Station	<b>1</b>
Use ActiveX in IE	<b>No</b>
Use Virtual Viewer	<b>Yes</b>



## Information

Details for the fields:

- For **Display name**, you can provide any name, but use of the Datacap application name is easier to track.
- The **ID** field is automatically populated based on the Display name field entry. The ID value can be edited but it is better to maintain the same names.
- For **Datacap wTM URI**, the value `ecmedu01` is the server name where wTM is installed.
- For **Application**, select the value from the list. The list contains all the available applications as listed in Datacap.xml file.

The screenshot shows the 'Configuration Parameters' dialog box with the 'General' tab selected. The form contains the following fields:

* Display name:	<input type="text" value="MultiPageApp"/> <span>(i)</span>
* ID:	<input type="text" value="MultiPageApp"/> <span>(i)</span>
* Datacap wTM URI:	<input type="text" value="http://ecmedu01:85/ServicewTM.svc"/> <span>(i)</span>
* Application:	<input type="text" value="MultiPageApp"/> <span>(i)</span> <span>▼</span>
* Default Station:	<input type="text" value="1"/> <span>(i)</span>
Use ActiveX in IE:	<input type="radio"/> Yes <input checked="" type="radio"/> No <span>(i)</span>
Use Virtual Viewer:	<input type="radio"/> Yes <input checked="" type="radio"/> No <span>(i)</span>
<span>Connect...</span>	

Red boxes highlight the 'Display name' input field, the 'Datacap wTM URI' input field, and the 'Use ActiveX in IE' radio button group.

- \_\_\_ 4. Click **Connect** to test the repository connection.
- \_\_\_ 5. On the **Log In** page, enter the following credentials for a user that has administration access to the repository.
  - **User name:** admin
  - **Password:** admin
- \_\_\_ 6. Click **Log in** and verify that the login is successful.

The Configuration Parameters subtab becomes active and you can open it now.

- \_\_\_ 7. Click **Save and Close** to save the settings and close the new repository tab.



## Troubleshooting

If the **Save and Close** action is not enabled, do the following steps:

- Leave the **New Repository** tab open (important step, otherwise you will lose your entries).
- Close the **Repositories** tab.
- In the **New Repository** tab, click **Save and Close**.
- Reopen the **Repositories** tab for the next step.

- \_\_\_ 8. On the **Repositories** tab, verify that the new repository (**MultiPageApp**) is listed.

Repositories			
Display		ID	Server Type
	ClaimForm	ClaimForm	datacapRepoType
	LoanProcess	LoanProcess	FileNet Content Manager
	MultiPageApp	MultiPageApp	datacapRepoType

- \_\_\_ 9. Close the **Repositories** tab.

## 2.2. Associate the new repository with the Datacap user desktop

- 1. In the **Desktops** tab, select the **Datacap** desktop from the list and click **Edit**.

The screenshot shows the 'Desktops' tab interface. At the top, there are several buttons: 'New Desktop' (blue), 'Edit' (red box), 'Copy', 'Delete', 'Refresh', 'Export', and 'Import'. Below the buttons is a table with columns: 'Name', 'ID', and 'Description'. There are three rows of data:

Name	ID	Description
Admin Desktop	admin	Desktop for users with administrative privi
<b>Datacap</b>	<b>datacap</b>	Default desktop for Datacap Navigator
Datacap Admin Desktop	dadmin	Datacap Navigator administration feature

- 2. On the **Datacap** tab, select the **Repositories** tab.
- 3. From the **Available Repositories** pane, select **MultiPageApp** and click forward arrow to move it to **Selected Repositories** pane.

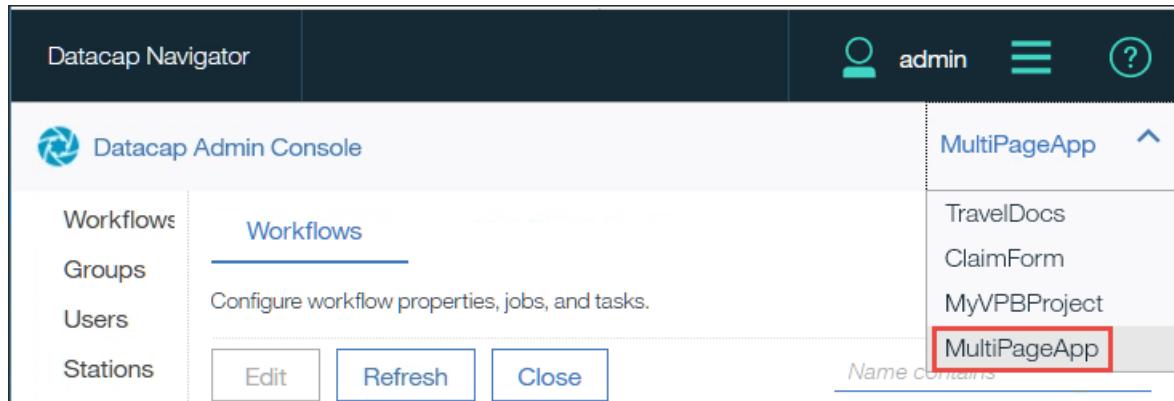
The screenshot shows the 'Repositories' tab interface. It has two main panes: 'Available Repositories' on the left and 'Selected Repositories' on the right. The 'Available Repositories' pane contains 'LoanProcess'. The 'Selected Repositories' pane contains 'TravelDocs', 'Sales', 'MyVPBProject', 'ClaimForm', and 'MultiPageApp' (which is highlighted with a red box). Between the two panes is a central area with a right-pointing arrow (highlighted with a red box) and a left-pointing arrow.

- 4. Click **Save and Close** to save the settings and close the **Datacap** tab.
- 5. Associate the new repository with the Datacap admin desktop.
  - a. In the **Desktops** tab, select the **Datacap Admin Desktop** and click **Edit**.
  - b. On the **Datacap Admin Desktop** tab, select the **Repositories** subtab.
  - c. From the **Available Repositories** pane, select **MultiPageApp** and click forward arrow to move it to **Selected Repositories** pane.
  - d. Click **Save and Close** to save the settings and close the tab.
- 6. On the banner, click the user name (**admin**) and select **Log Out** from the list and click **Log Out**.

## 2.3. Edit the workflow for the new Datacap application

Since you copied an existing application, some names in the application must be changed for the new application. In this section, you review the tasks and edit them.

- \_\_\_ 1. Start the Datacap Administration console.
  - \_\_\_ a. In the Firefox browser, click the **DC Admin** shortcut. Or go to the **URL**:  
<http://ecmedu01:9081/navigator/?desktop=dcadmin>
  - \_\_\_ b. Enter the following values:
    - **User name:** admin
    - **Password:** admin
  - \_\_\_ c. Click **Log In**.
- \_\_\_ 2. In the Datacap Administration Console, open your application.
  - \_\_\_ a. On the upper right of the page, click **TravelDocs**, and from the applications list, select **MultiPageApp**.



- \_\_\_ b. On the **Log In** page, enter the following values:
  - **User name:** admin
  - **Password:** admin
- \_\_\_ c. Click **Log In**.
- \_\_\_ d. On the upper right of the page, verify that **MultiPageApp** is selected.
- \_\_\_ 3. Rename the Claim Form Job.
  - \_\_\_ a. In the **Workflows** tab, select **MultiPageApp** and click **Edit**.
  - \_\_\_ b. In the **MultiPageApp** tab, select **Jobs** subtab.
  - \_\_\_ c. Select **Claim Form Job** and click **Edit**.
  - \_\_\_ d. On the **Claim Form Job** tab, **General** subtab, for the **Name** field, edit the value to: **MultiPageApp Job**.

Job: Claim Form Job

General	Tasks
* Name <span style="color: blue;">(i)</span>	MultipageApp Job
Description	Standard processing of single page images
* Priority <span style="color: blue;">(i)</span>	5

- \_\_\_ e. Click **Save** and verify that the tab is renamed as **MultiPageApp Job**.
- \_\_\_ 4. Edit the Verify task.
- \_\_\_ a. In the **MultiPageApp Job** tab, select the **Tasks** subtab, select the **Verify** task row, and click **Edit**.
- \_\_\_ b. In the **Verify** tab, click the **Advanced** subtab.
- \_\_\_ c. Scroll down toward the end of the page and in the **DCO tree view > Actions** section, select the **Check Integrity** option.

The Document Integrity setting is used when a document has multiple pages to arrange the pages in the correct order.

Disassemble Documents <span style="color: blue;">(i)</span>	<input checked="" type="checkbox"/>
Check Integrity <span style="color: blue;">(i)</span>	<input checked="" type="checkbox"/>
Undo <span style="color: blue;">(i)</span>	<input checked="" type="checkbox"/>

- \_\_\_ d. Click **Save and Close** to save the settings and close the **Verify** tab.
- \_\_\_ e. Click **Save and Close** on all other opened tabs.
- \_\_\_ 5. Log out of Datacap Admin Console.
- \_\_\_ a. On the banner, click the user name (**admin**) and select **Log Out** from the list.
- \_\_\_ b. Click **Log Out** and close the browser.

## Section 3. Verify the application in Datacap Application Manager

In this section, you review the settings for your application in the Datacap Application Manager.

- \_\_\_ 1. On the Windows desktop, double-click the **Datacap Application Manager** shortcut.
- \_\_\_ 2. On the **Datacap Application Manager** window, on the left pane, verify that **MultiPageApp** is listed and click it.
- \_\_\_ 3. On the **Main** tab on the right, verify that all the following fields have the values with the **MultiPageApp** name:
  - **Administration**, **Engine**, and **Fingerprint database** fields
  - **Batch folder**, **Export folder**, **Fingerprint folder**, and **VScan source folder** fields
  - **Enable FPXML** (selected)
- \_\_\_ 4. Close the Datacap Application Manager.

## Section 4. Define Document hierarchy

In this section, you configure Datacap objects (DCO) for your application in Datacap Studio.

### 4.1. Copy the images to be scanned to the new application folder

- 1. Open Windows Explorer, browse to the C:\Datacap\MultiPageApp\images folder, verify that the folder has the **1-ClaimForm-BC.tif** file and delete any other files.



#### Important

Keep the **1-ClaimForm-BC.tif** file.

If the file is not in the folder, copy it from the C:\labfiles\Ex. 3.Images\_for\_scanning folder.

- 2. Open another Windows Explorer window, browse to the C:\labfiles folder, copy the **Car1.tif** image, and paste it to the C:\Datacap\MultiPageApp\images folder.  
You add the second file to test the application for multiple documents processing.
- 3. Verify that the C:\Datacap\MultiPageApp\images folder contains the following two files:
  - **1-ClaimForm-BC.tif**
  - **Car1.tif**
- 4. Double-click the **Car1.tif** file to open and observe the fields.

You use the information on this file and create the following Datacap objects (DCO) in the next section:

- Document type: **Car Rental**

You keep the Insurance Docs and Car Rental as two separate document types, by creating a separate Document type.

- Page type: **Rental Form**



#### Information

Instead of two separate document types, you can also add the Rental Form as a separate page type to the same Document type (Insurance Docs).

For this exercise, you use a separate document type for each type of image.

- Fields:
  - **Pickup** (for Pickup Date)
  - **Return** (for Return Date)
  - **Total**

**Note**

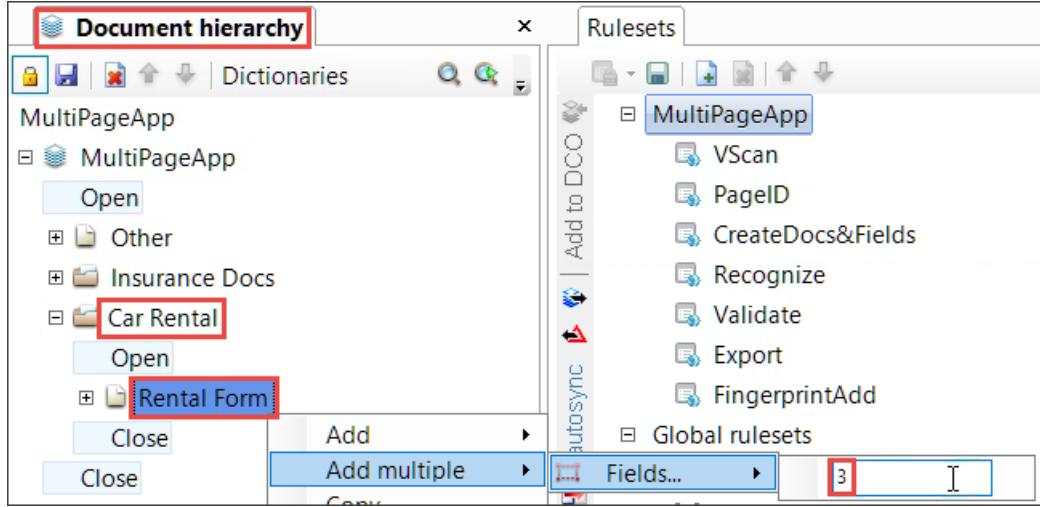
You use OMR recognition for the multiple choices field, **Options**. You configure OMR in a later section in this exercise.

- \_\_\_ 5. Close the **Car1.tif** file.

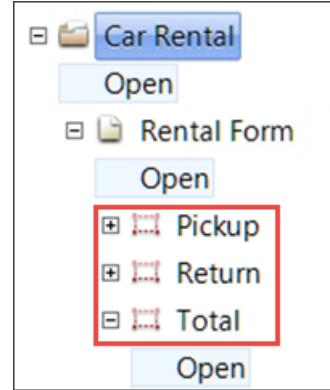
## 4.2. Create a document hierarchy for the new application

- \_\_\_ 1. Log in to Datacap Studio.
  - \_\_\_ a. On the Windows desktop, click the **Datacap Studio** shortcut.
  - \_\_\_ b. In the **Applications** dialog box, select **MultiPageApp** and click **Next**.
  - \_\_\_ c. On the **TaskMaster Login** page, enter the following values:
    - **User ID:** admin
    - **Password:** admin
    - **Station ID:** 1
  - \_\_\_ d. Click **Finish**.
 

Datacap Studio opens with the **Rulemanager** tab selected.
- \_\_\_ 2. In the **Document hierarchy** pane, click the **Lock DCO for editing** icon to lock it.
- \_\_\_ 3. Expand **MultiPageApp** and observe that the **Insurance Docs** Document type is listed.  
The **Insurance Docs** type comes from the source application that you built earlier.
- \_\_\_ 4. Add a second Document type for the new image.
  - \_\_\_ a. Right-click **MultiPageApp** and click **Add > Document**.
  - \_\_\_ b. Select **Document1**, click, and rename it as: **Car Rental**
- \_\_\_ 5. Add a page to the document type.
  - \_\_\_ a. Right-click **Car Rental** and click **Add > Page**.
  - \_\_\_ b. Select **Car Rental > Page1**, click, and rename it as: **Rental Form**
- \_\_\_ 6. Add fields to the page.
  - \_\_\_ a. Right-click **Rental Form** and select **Add multiple > Fields**.
  - \_\_\_ b. In the text box, type **3** for the number of fields and press enter.



- \_\_\_ c. Expand the **Rental Form** page and verify that three fields are listed.
- \_\_\_ d. Rename the fields with the following names in the order it is shown.
  - **Field1:** Pickup
  - **Field2:** Return
  - **Field3:** Total



- \_\_\_ 7. Click **Save Changes** and click the **Unlock DCO** lock icon.
- \_\_\_ 8. On the upper right corner of the page, click **Exit** to close Datacap Studio.

You must restart Datacap Studio to refresh the fields to be available for ruleset configuration in the next section.

## Section 5. Configure rulesets and task profiles

You copied the new application from the ClaimForm application that you built in the previous exercise. The rulesets that you configured earlier are applicable to this application. The scan images for the ClaimForm application have a barcode that you used for page identification. The new application uses a second set of images that do not have a barcode. Therefore, you must configure a ruleset with another type of page identification.

In this section, you review the rulesets and add edits to process the second page.



### Note

Each of the following subsections has similar instructions for creating rulesets. If you are already familiar with the steps, you can refer to the tables or the screen captures to create a ruleset and skip the instructions.

### 5.1. Configure the PageID ruleset for page recognition without barcode

The PageID ruleset already contains rules, functions, and actions for classifying the document and identifying the page with a barcode from the previous exercise. In this section, you configure page identification for the new page without a barcode.

- \_\_\_ 1. Log in to Datacap Studio.
  - \_\_\_ a. On the Windows desktop, click the **Datacap Studio** shortcut.
  - \_\_\_ b. In the **Applications** dialog box, select **MultiPageApp** and click **Next**.
  - \_\_\_ c. On the **TaskMaster Login** page, enter the following values:
    - **User ID:** admin
    - **Password:** admin
    - **Station ID:** 1
  - \_\_\_ d. Click **Finish**.

Datacap Studio opens with the **Rulemanager** tab selected.
- \_\_\_ 2. Check the Classify function.
  - \_\_\_ a. On the **Rulesets** tab, select the **PageID** ruleset, click the lock icon, and click **Lock ruleset for editing** from the list.
  - \_\_\_ b. Expand the **PageID > (P) Open > Classify** function and check the actions.

**Note**

The Datacap actions return a Boolean value (True or False). When an action returns True, the function continues to run. When an action returns False, the function stops and does not run the rest of the actions in that function.

For the scan images without a barcode that you use as the second page, the **MatchAnyBarcode** action returns False. The function stops even though the other two (**RecognizePageOCR\_S** and **FindFingerprint**) actions are applicable to the second page for page identification.

You add another function to identify the second page.

- 
- \_\_\_ 3. Add a function to the (P) Open rule.
    - \_\_\_ a. In the **PageID** ruleset, right-click the **(P) Open** rule and click the **Add Function**.
    - \_\_\_ b. Select and click **Function1** and rename it to: **Classify by Fingerprint**
  - \_\_\_ 4. Add an action to the **Classify by Fingerprint** function.
    - \_\_\_ a. Select the **Classify by Fingerprint** function.
    - \_\_\_ b. In the **Actions library** tab, expand the **ocr\_sr > Datacap.Libraries.ScansoftR.Actions** node, and select the **RecognizePageOCR\_S** action.
    - \_\_\_ c. Click the **Add to function** vertical bar on the left side of the **Action library** tab.
    - \_\_\_ d. In the **Rulesets** pane, verify that the **RecognizePageOCR\_S ()** action is added to the **Classify by Fingerprint** function.

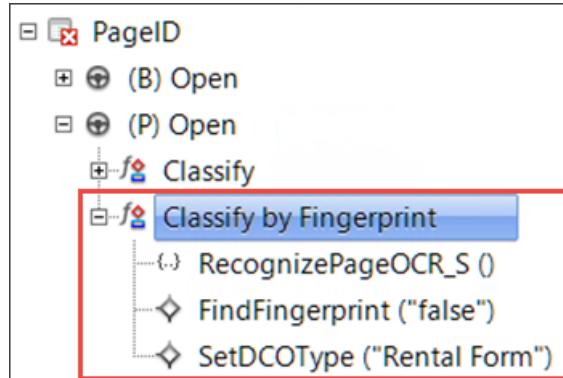
**Note**

If you already added an action to a function, you can copy the action from the function and paste it to another function.

- 
- \_\_\_ 5. Add the **FindFingerprint** action.
    - \_\_\_ a. Expand **PageID > (P) Open > Classify** and right-click the **FindFingerprint ("False")** action and click **Copy**.
    - \_\_\_ b. Right-click the **Classify by Fingerprint** function and click **Paste**.

You can also add the action from the **autodoc** action library to the **Classify by Fingerprint** function.
  - \_\_\_ 6. Add the **SetDCOType** action.
    - \_\_\_ a. In the **Rulesets** tab, select the **Classify by Fingerprint** function.
    - \_\_\_ b. In the **Action library** tab, from the **ApplicationObjects > Datacap.Libraries.ApplicationObjects.Actions** action library, add the **SetDCOType** action.
    - \_\_\_ c. In the **Rulesets** tab, select the **Classify by Fingerprint > SetDCOType ("")** action.

- \_\_\_ d. In the **Properties > Parameters** section, for the **string Type** field, enter: Rental Form
- \_\_\_ 7. Verify that the completed PageID ruleset contains the new function and actions as shown in the following screen capture.



- \_\_\_ 8. Click the **Save changes** icon in the toolbar.
- \_\_\_ 9. Select the **PageID** ruleset, click the lock icon, and click **Publish ruleset** from the list.

## 5.2. Configure the CreateDocs&Fields ruleset

In this section, you assign the (P) Open rule in the CreateDocs&Fields ruleset to the Rental Form page type.

- \_\_\_ 1. On the **Rulesets** tab, select the **CreateDocs&Fields** ruleset, click the lock icon, and click **Lock ruleset for editing** from the list.
- \_\_\_ 2. Expand the **CreateDocs&Fields** ruleset and select the **(P) Open** rule.
- \_\_\_ 3. In the **Properties** tab, expand the **Run rule at start of** node.  
The **Pages** option and for the **of type** field, **Claim Page** are already selected.
  - \_\_\_ a. Click the minus sign next to the **of type: Claim Page** line to remove that item.
  - \_\_\_ b. Verify that the **Pages** option is still selected.



- \_\_\_ 4. Click the **Save changes** icon in the toolbar.
- \_\_\_ 5. Select the **CreateDocs&Fields** ruleset, click the lock icon on the toolbar of the **Rulesets** tab, and click **Publish ruleset** from the list.

### 5.3. Configure the Recognize ruleset

In this section, you assign the (P) Open rule in the Recognize ruleset to the Rental Form page type. The steps are similar to the previous section for the CreateDocs&Fields ruleset.

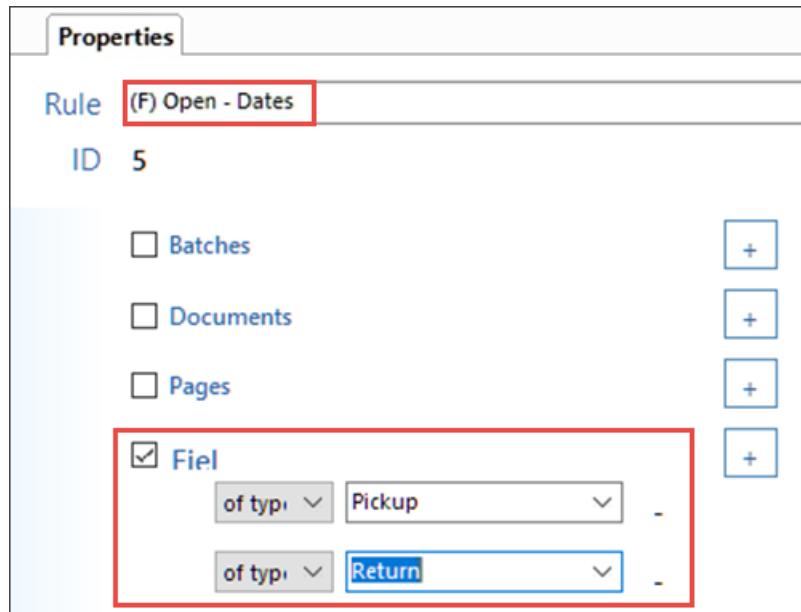
- \_\_\_ 1. On the **Rulesets** tab, lock the **Recognize** ruleset for editing.
- \_\_\_ 2. Assign the existing (P) Open rule in the Recognize ruleset to the Rental Form page type.
  - \_\_\_ a. Expand the **Recognize** ruleset and select the **(P) Open** rule.
  - \_\_\_ b. In the **Properties** tab, expand the **Run rule at start of** node.
  - \_\_\_ c. Click the minus sign next to the **of type: Claim Page** line to remove that item.
  - \_\_\_ d. Verify that the **Pages** option is still selected.
- \_\_\_ 3. Save changes and publish the **Recognize** ruleset.

### 5.4. Configure the Validate ruleset

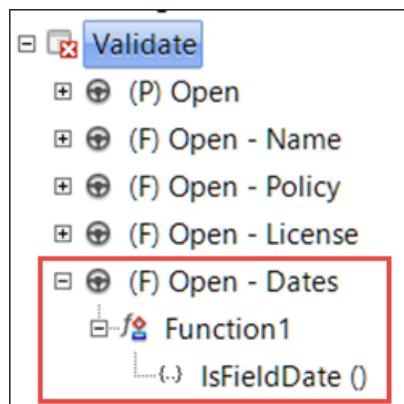
The application already has the Validate ruleset. In this section, you edit the Validate ruleset to add validation for the new fields. You also assign the rule in the Validate ruleset to the Rental Form page type.

- \_\_\_ 1. On the **Rulesets** tab, lock the **Validate** ruleset for editing.
- \_\_\_ 2. Assign the existing (P) Open rule to the Rental Form page type.
  - \_\_\_ a. Expand the **Validate** ruleset and select the **(P) Open** rule.
  - \_\_\_ b. In the **Properties** tab, expand the **Run rule at start of** node.
    - The **Pages** option and for the **of type** field, **Claim Page** are already selected.
  - \_\_\_ c. Click the minus sign next to the **of type: Claim Page** line to remove that item.
  - \_\_\_ d. Verify that the **Pages** option is still selected.
  - \_\_\_ e. In the **Rulesets** tab, click the **Save changes** icon.
- \_\_\_ 3. Add a rule to the Validate ruleset to check the date fields.
  - \_\_\_ a. Right-click the **Validate** ruleset and click **Add Rule** from the list.
  - \_\_\_ b. Rename **Rule1** to: **(F) Open - Dates**
- \_\_\_ 4. Assign the rule to the Pickup field.
  - \_\_\_ a. In the **Rulesets** pane, select the **Validate > (F) Open - Dates** rule.
  - \_\_\_ b. In the **Properties** tab > **Run rule at start of** node, select **Fields** and click the plus sign next to it.
  - \_\_\_ c. For the **of type** field, select **Pickup** from the list.

- \_\_\_ d. Click the plus sign again and for the **of type** field, select **Return** from the list.



- \_\_\_ e. In the **Rulesets** tab, click the **Save changes** icon.
5. Add an action to the **Function1**.
- Select **(F) Open - Dates > Function1**.
  - On the **Actions library** tab, from the **ValidationsAndTextAdjustments > Datacap.Libraries.ValidationsAndTextAdjustments.Actions** action library, select the **IsFieldDate** action.  
This action accepts any valid date.
  - Click the **Add to function** vertical bar on the left side of the **Action library** tab.
  - Click the **Save changes** icon in the toolbar.
  - In the **Rulesets** pane, verify that the **IsFieldDate ()** action is added to **Function1**.

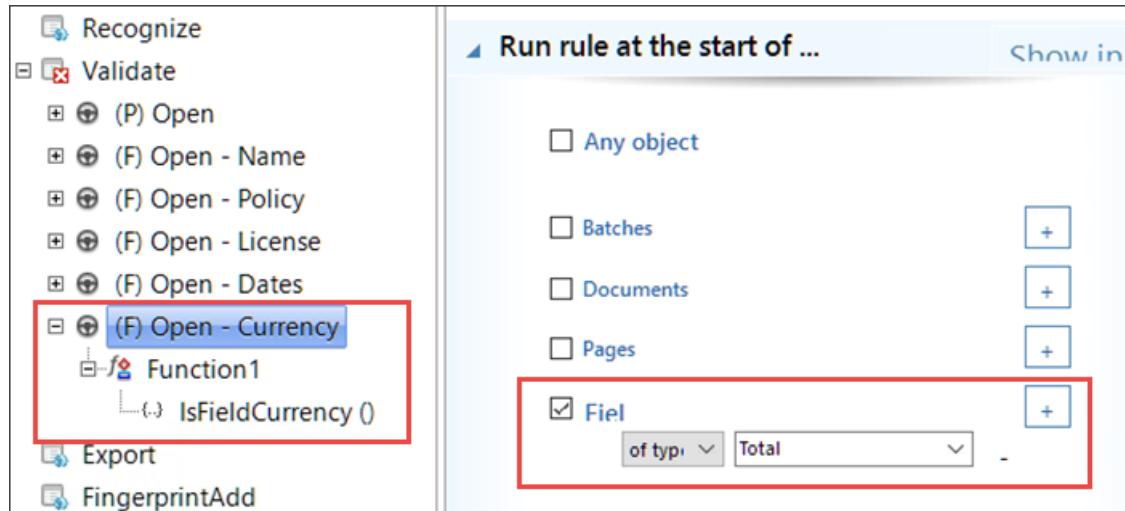


6. Add a rule to the **Validate** ruleset to check the **Total** field.

Steps are similar to the rule for the **Dates** field.

- Right-click the **Validate** ruleset and click **Add Rule** from the list.

- \_\_\_ b. Rename **Rule1** to: (F) Open - Currency
  - \_\_\_ c. To (F) Open - Currency > **Function1**, from the **ValidationsAndTextAdjustments** > **Datacap.Libraries.ValidationsAndTextAdjustments.Actions** action library, add the **IsFieldCurrency ()** action.  
This action checks whether the field value is in the currency format.
  - \_\_\_ d. Select the **(F) Open - Currency** rule and in the **Properties** > **Run rule at start of** node, select the **Fields**.
  - \_\_\_ e. Click the plus sign and for the **of type:** field, select **Total**.
- \_\_\_ 7. Verify that the completed rule contains the action and it is associated with the Total field as shown in the screen capture.



- \_\_\_ 8. Save the changes and publish the **Validate** ruleset.

## 5.5. Configure the Export ruleset

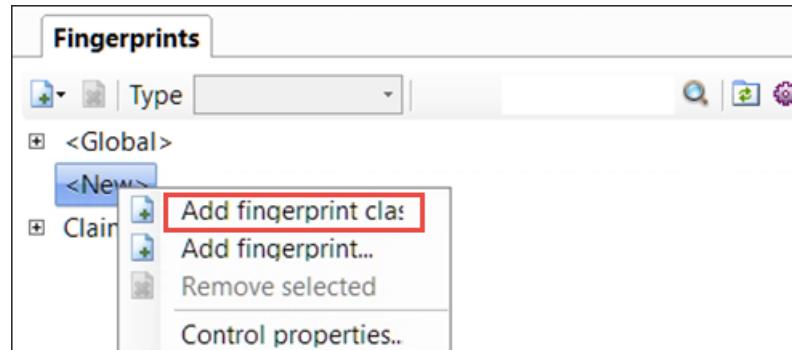
In this section, you assign the Export Page Fields rule in the Export ruleset to the Rental Form page type. The steps are similar to the previous section for the CreateDocs&Fields ruleset.

- \_\_\_ 1. On the **Rulesets** tab, lock the **Export** ruleset for editing.
- \_\_\_ 2. Expand the **Export** ruleset and select the **Export Page Fields** rule.
- \_\_\_ 3. In the **Properties** tab, expand the **Run rule at start of** node.
- \_\_\_ 4. Click the minus sign next to the **of type: Claim Page** line to remove that item.
- \_\_\_ 5. Verify that the **Pages** option is still selected.
- \_\_\_ 6. **Save the changes** and publish the **Export** ruleset.

## Section 6. Configure Fingerprints and Zones

### 6.1. Configure Fingerprinting

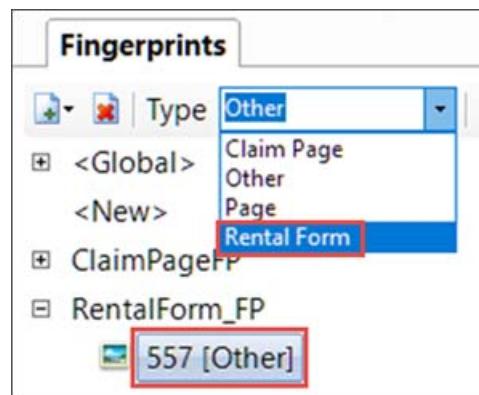
- \_\_\_ 1. In Datacap Studio, click the **Zones** tab.
- \_\_\_ 2. Create a fingerprint class for the second page.
- \_\_\_ a. On the upper left pane, in the **Fingerprints** tab, right-click **New** and select **Add fingerprint class** from the list.



- \_\_\_ b. In the **New fingerprint class** dialog box, for the **Class name** field, enter: **RentalForm\_FP** and click **OK**.
- \_\_\_ 3. Create a fingerprint.
  - \_\_\_ a. In the **Fingerprints** pane, right-click **RentalForm\_FP** and select **Add fingerprint**.
  - \_\_\_ b. In the **Open** window, browse to the **C:\Datacap\MultiPageApp\images** folder, select the **Car1.tif** file, and click **Open**.

It takes a few moments to complete the cleanup of the image and create a fingerprint.

- \_\_\_ c. In the **Fingerprints** pane, expand **RentalForm\_FP** and verify that a fingerprint is created (Example: **557[Other]**).  
You might get a different number.
- \_\_\_ d. Select **557[Other]** and on the toolbar, for the **Type** field, select **Rental Form** from the list.





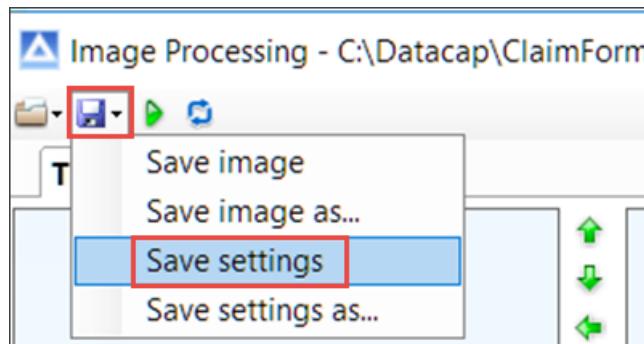
## Information

The items in the Type list are stored in a file that is called MultiPageApp.XML in the application folder (C:\Datacap\MultiPageApp\dco\_MultiPageApp).

- \_\_\_ 4. Configure the image settings.
  - \_\_\_ a. Select **557[Rental Form]** and in the **Image** tab, verify that the image is opened.
  - \_\_\_ b. On the **Image** tab toolbar, on the upper right, click the gear icon.
  - \_\_\_ c. In the **Image Processing** page, in the **Properties** tab on the right pane, verify and set the values for the following fields:
    - Deskew: **True**
    - Despeckle: **True**



- \_\_\_ d. In the toolbar, click the **Save** icon and from the list, select **Save settings**.



- \_\_\_ e. On the **Setting saved** dialog box, click **OK** and close the **Image Processing** page.
- \_\_\_ 5. Check the fingerprint files that are created.
  - \_\_\_ a. Open Windows Explorer and browse to the C:\Datacap\MultiPageApp\fingerprint folder.
  - \_\_\_ b. Verify that the following files that start with **557** are created:

You might have a different number on your system. <555> is the default that the application wizard created. <556> is for the Claim Page.

- **557.tif**
- **557.xml**
- **557.cco**

## 6.2. Configure zones

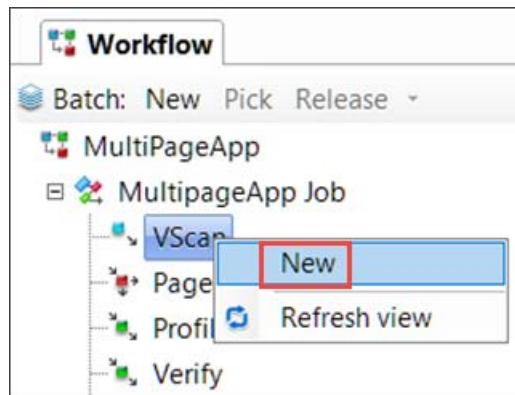
- 1. On the **Zones** tab, in the **Fingerprints** tab, select **RentalForm\_FP > 557 [Rental Form]** and on the **Image** tab, verify that the image is shown.
- 2. On the left middle pane, in the **Document hierarchy** tab, click the lock icon to enable the DCO editing.
- 3. Expand **MultiPageApp > Car Rental > Rental Form**.
- 4. Draw zones for the fields for OCR recognition.
  - a. In the **Document hierarchy** pane, select the **Pickup** field.
  - b. On the **Image** tab, draw the zone around the date value for the **Pickup** column.
  - c. Draw from upper left corner to the lower right corner and avoid the border line in the drawing.
  - d. Repeat the steps to select each field and draw zones to other two fields: **Return** (date) and **Total** (dollars).

- 5. In the **Document hierarchy** pane, click **Save** and click the lock icon to unlock the DCO.
- 6. Notice that the fields have a blue color  on the icons after you complete the zones configuration.

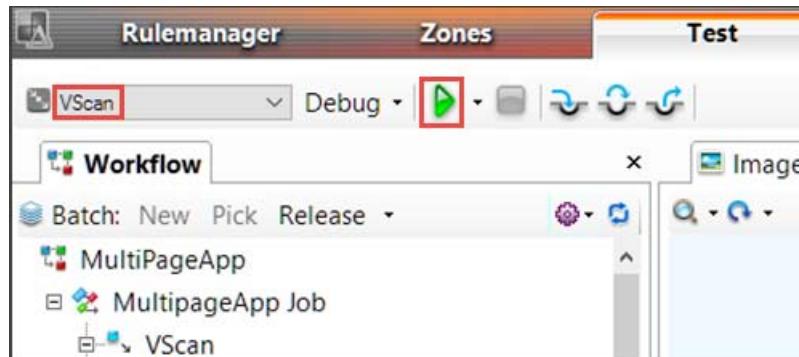
## Section 7. Test the application

### 7.1. Start a batch to test the application

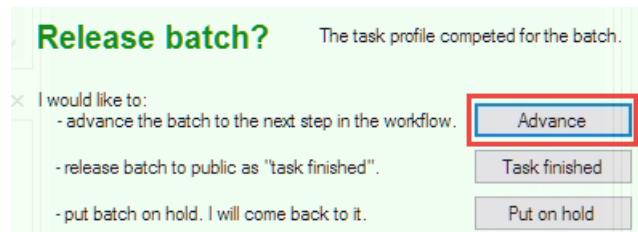
- 1. Verify that the **1.ClaimForm-BC.tif** and **Car1.tif** files are available in the **C:\Datacap\MultiPageApp\images** folder.  
In a previous section, you placed these two files in this folder.
- 2. In Datacap Studio, click the **Test** tab from the top toolbar.
- 3. In the upper left pane, on the **Workflow** tab, expand **MultiPageApp > MultiPageApp Job** right-click **VScan**, and select **New**.



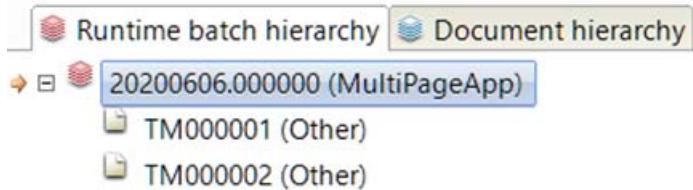
- 4. In the top toolbar, ensure that **VScan** is selected and click the **Process rules for target object** (green arrow icon) next to it.



- 5. In the **Release batch** dialog box, click **Advance** to move the batch to the **PageID** task.



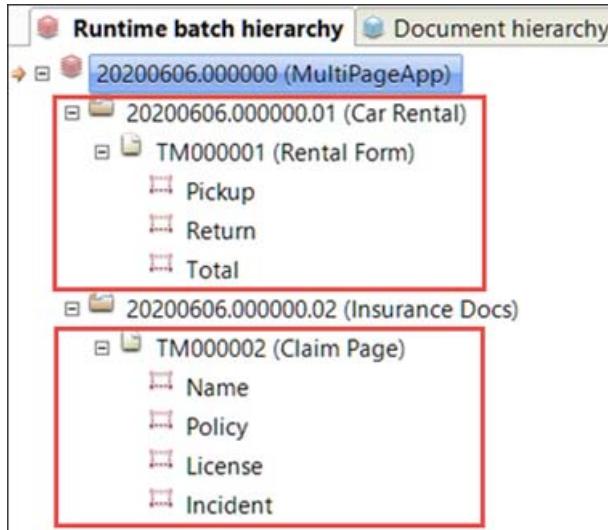
- \_\_\_ 6. In the **Runtime batch hierarchy tab**, verify that it has two batch items now and the pages are listed as **Other**.



- \_\_\_ 7. Click each item (**TM000001 (Other)** and **TM000002 (Other)**) and verify that the corresponding scanned image is shown in the **Image** tab.

## 7.2. Run the Page ID task

- \_\_\_ 1. In the top toolbar, ensure that **PageID** is selected and click the green arrow icon to run the **PageID** task.  
This step classifies the documents and identifies the pages.
- \_\_\_ 2. In the **Release batch** dialog box, click **Advance** to move the batch to the **Profiler** task.
- \_\_\_ 3. In the **Runtime batch hierarchy tab**, expand the batch node and verify that the documents are classified with the names: **Car Rental** and **Insurance Docs**
- \_\_\_ 4. Expand **Car Rental** and verify that the page is identified with the name: **Rental Form**
- \_\_\_ 5. Expand **Rental Form** and verify that the page contains the three fields that you defined: **Pickup**, **Return**, and **Total**.
- \_\_\_ 6. Expand **Insurance Docs** and verify that the page is identified with the name: **Claim Page**
- \_\_\_ 7. Expand **Claim Page** and verify that the page contains the four fields that you defined: **Name**, **Policy**, **License**, and **Incident**.

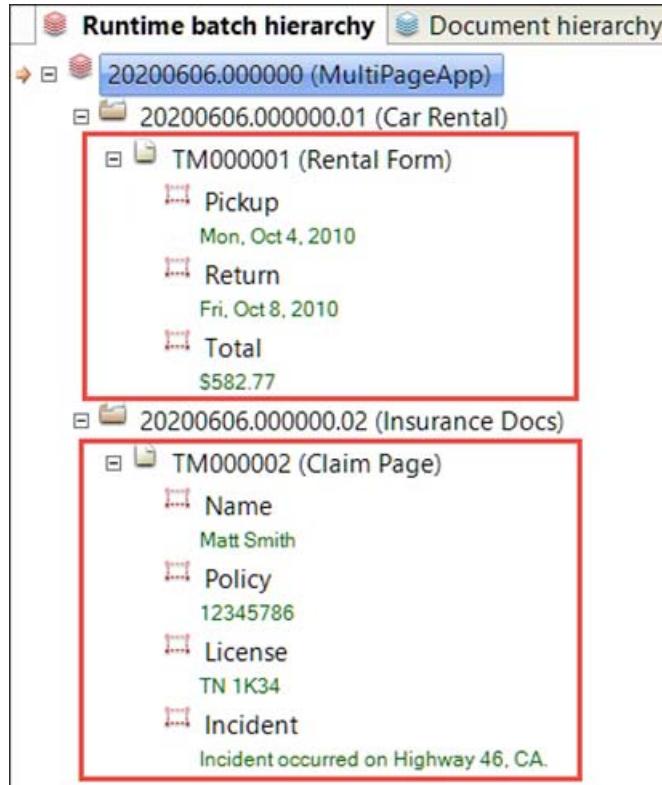


### 7.3. Run the Profiler task

- 1. In the top toolbar, ensure that **Profiler** is selected and click the green arrow icon to run the **Profiler** task.

Wait for the process to complete. This step extracts the data from the fields.

- 2. In the **Release batch** dialog box, click **Keep running**.
- 3. In the **Runtime batch hierarchy tab**, expand the batch item nodes and verify that the fields are populated with values (shown in green text).



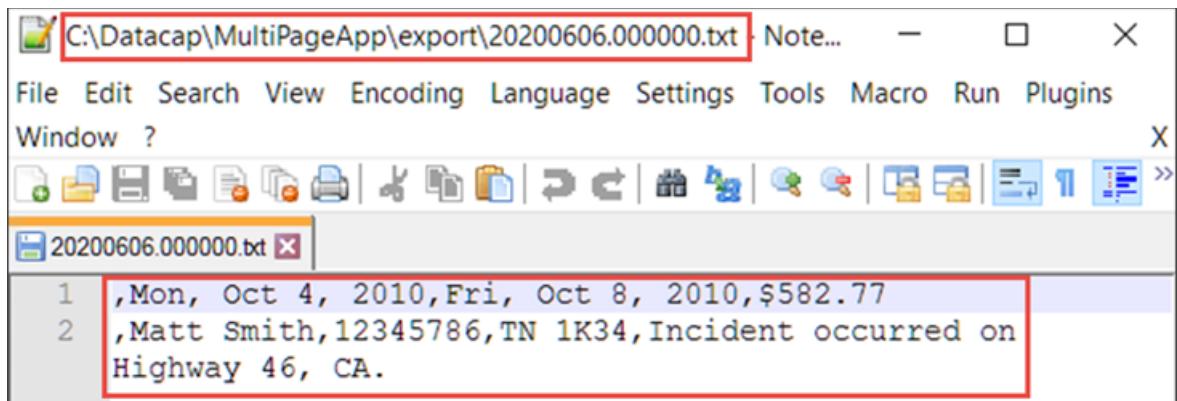
- 4. Click a field (Example: **Pickup**) and notice that the image is shown in the Image tab on the middle pane with the fields Zones selected.

### 7.4. Advance the batch to Verify task and process the batch

- 1. On the **Workflow** tab, in the **MultiPageApp Job > Profiler** node, right-click the batch (**running**), and select **Advance**.  
The workflow moves to the Verify step.
- 2. In the top toolbar, ensure that **Verify** is selected and click the green arrow icon to run the **Verify** task.

## 7.5. Process the Export task and verify the export file

- \_\_\_ 1. In the **Release batch** dialog box, click **Advance** to move the batch to the **Export** task.
- \_\_\_ 2. On the **Runtime batch hierarchy** pane, take a note of the batch name: <Date>.<batch number> (Example: 20200606.000000).  
You use this value to identify the export file in the following steps.
- \_\_\_ 3. In the top toolbar, ensure that **Export** is selected and click the green arrow icon to run the task.
- \_\_\_ 4. In the **Release batch** dialog box, click **Advance** to complete the batch.
- \_\_\_ 5. Check the export text file.
  - \_\_\_ a. In Windows Explorer, browse to the `C:\Datacap\MultiPageApp\export` folder and verify that a text file with your batch name (that you noted in step 2) is created with today's date.
  - \_\_\_ b. Open the file in Notepad++ and check that the file contains the comma-separated values for the following fields:
    - Pickup date, return date, and total dollars for the Rental Form page in the first line
    - Name, policy number, license, and incident for the Claim Page in the second line



- \_\_\_ c. Close the file and Notepad++.

**End of exercise**

## Exercise review and wrap-up

In this exercise, you copied an existing application that is based on the Forms template and created a application. You configured the new application for processing different page types in a batch.

# Exercise 6. Configuring an application for recognition with OMR

## Estimated time

01:30

## Overview

This exercise teaches how to implement the OMR capability in a Datacap application to extract field values from multiple choice check boxes.

## Objectives

After completing this exercise, you should be able to:

- Configure OMR for a Datacap application
- Test the application

## Introduction

In this exercise, you configure OMR to extract field values from multiple choice check boxes.

This exercise includes the following sections:

- [Section 1, "Configure OMR fields"](#)
- [Section 2, "Configure rulesets and zones for OMR"](#)
- [Section 3, "Test the application for OMR"](#)

## Requirements

This exercise builds on [Exercise 5, "Building an application to process multiple page types"](#). Before starting this exercise, you must complete the [Exercise 5, "Building an application to process multiple page types"](#).

A solution of the Datacap application for the exercise is provided in C:\Datacap folder with the application name, **SolutionMultiPageApp**. You can access this application like any other Datacap application.

## Section 1. Configure OMR fields

In the previous exercise, you created a new Datacap application by copying an existing application. In this section, you enhance the application for multiple option check boxes recognition with OMR.



### Note

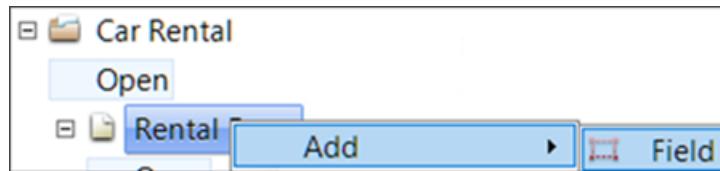
A solution of the Datacap application for this exercise is provided in C:\Datacap folder with the application name: **SolutionMultiPageApp**

You can access this application like any other Datacap application.

### 1.1. Configure document hierarchy for the new page

In this section, you configure a field in document hierarchy for the new page of your application by using Datacap Studio.

- 1. Before you begin the exercise, ensure that the components are running. If you did not start the components already, refer to [Section 1, "Prepare your system: Start the components"](#) to start them.  
If you restart your system, you must start all the components.
- 2. If it is not already open, log in to Datacap Studio.
  - a. Double-click the **Datacap Studio** shortcut on the Windows desktop.
  - b. In the **Applications** dialog box, select **MultiPageApp** and click **Next**.
  - c. On the **Taskmaster Login** page, enter the following values:
    - **User ID:** admin
    - **Password:** admin
    - **Station ID:** 1
  - d. Click **Finish**.  
Datacap Studio opens with the **Rulemanager** tab selected.
- 3. Add a field to the Rental Form page.
  - a. In the **Document hierarchy** pane, click the **Lock DCO for editing**  icon.
  - b. Expand **MultiPageApp > Car Rental > Rental Form**.
  - c. Right-click **Rental Form** and click **Add > Field**.



- d. Verify that **Field1** is listed in the **Rental Form** page and rename it to: Options

## 1.2. Create a dictionary for the Options field

After you complete the application and run the batch, the results of the Options field are shown as ones and zeros. You can define the values to be shown as the names of the options by using a dictionary.

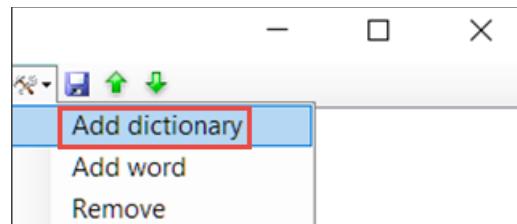
- \_\_\_ 1. Create a dictionary.

- \_\_\_ a. In the **Document hierarchy** pane toolbar, click **Dictionaries**.

If it is not visible, click the downward arrow on the toolbar.



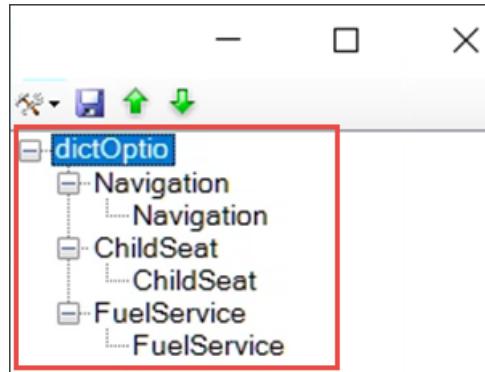
- \_\_\_ b. In the dialog box for the dictionary, click the **Edit dictionary** icon and select **Add Dictionary**.



- \_\_\_ c. Rename <new dictionary>: dictOptions
- \_\_\_ 2. Add words to the dictionary.
  - \_\_\_ a. Click the **Edit dictionary** icon and from the list, click **Add word**.
  - \_\_\_ b. For <new word>, enter Navigation and for **value**, enter: Navigation
  - \_\_\_ c. Select **dictOptions** and add the following two words (word, value):
    - ChildSeat, ChildSeat
    - FuelService, FuelService

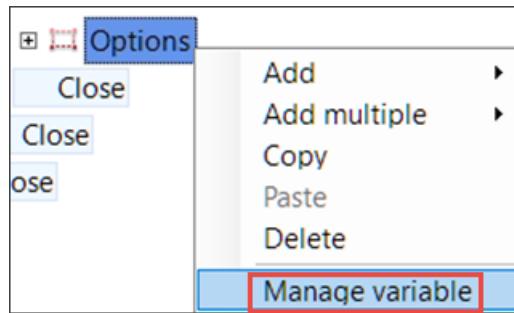
If you don't select **dictOptions**, the words are added as subnodes in **Navigation**.

- \_\_\_ 3. Verify that the values are added as shown in the following screen capture and click the **Save** icon.



### 1.3. Configure variables and child fields for the Options field

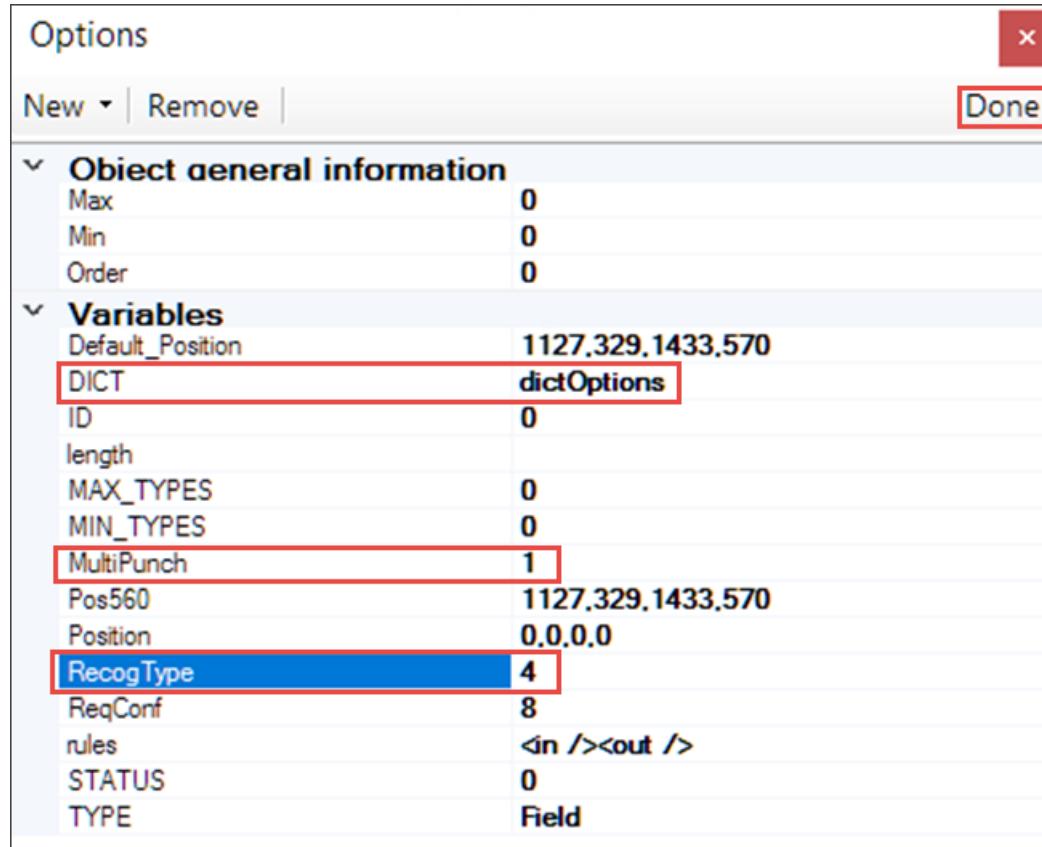
- \_\_\_ 1. Create a variable for the Options field.  
 \_\_\_ a. In the **Document hierarchy** pane, right-click **Options** and click **Manage variable**.



- \_\_\_ b. In the **Options** dialog box, click **New**, in the text box, type MultiPunch, and press enter.

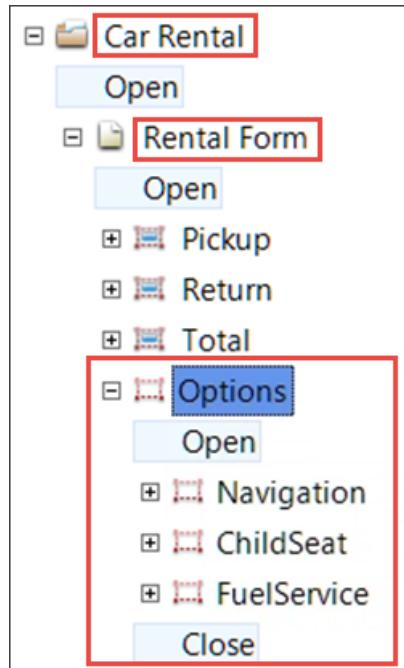


- \_\_\_ c. Create two more variables: **RecogType** and **DICT**
- \_\_\_ 2. Set values for the following fields:
- **MultiPunch:** 1  
This value indicates that multiple selections are allowed for the field.
  - **RecogType:** 4  
**RecogType = 4** indicates that the **Options** field is an OMR field.
  - **DICT:** dictOptions  
This value is the name of your dictionary that you created in the earlier step. It associates the dictionary with the Options field.



- \_\_\_ d. On the upper right of the toolbar, click **Done** to save the changes and close the dialog box.
- \_\_\_ 3. Add child fields for the Options field.
  - \_\_\_ a. In the **Document hierarchy** pane, right-click **Options** and click **Add multiple > Fields**.
  - \_\_\_ b. In the text box, type 3 and press enter.
  - \_\_\_ c. Verify that three child fields are listed for the **Options** field.  
The Options field has multiple choices and can have multiple values. You create child fields for the Options field for the multiple values.
  - \_\_\_ d. Select each of the following fields and rename with the names in the order it is shown.
    - **Field1**: Navigation
    - **Field2**: ChildSeat
    - **Field3**: FuelService

- \_\_\_ 4. Verify that the completed DCO has the fields as shown in the following screen capture.



- \_\_\_ 5. Click the **Save Changes** icon and click the **Unlock DCO** lock icon.  
\_\_\_ 6. On the upper right corner of the page, click **Exit** to close Datacap Studio.

You must restart Datacap Studio to refresh the fields to be available for ruleset configuration in the next section.

## Section 2. Configure rulesets and zones for OMR

In this section, you configure rulesets to your application to implement OMR capability.

### 2.1. Start the Datacap Studio

- 1. Before you begin the exercise, ensure that the components are running. If you did not start the components already, refer to [Section 1, "Prepare your system: Start the components"](#) to start them.

If you restart your system, you must start all the components.

- 2. Log in to Datacap Studio.
  - a. Double-click the **Datacap Studio** shortcut on the Windows desktop.
  - b. In the **Applications** dialog box, select **MultiPageApp** and click **Next**.
  - c. On the **Taskmaster Login** page, enter the following values:
    - **User ID:** admin
    - **Password:** admin
    - **Station ID:** 1
  - d. Click **Finish**.

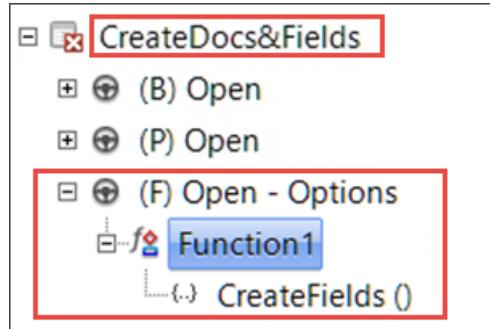
Datacap Studio opens with the **Rulemanager** tab selected.

### 2.2. Configure the CreateDocs&Fields ruleset

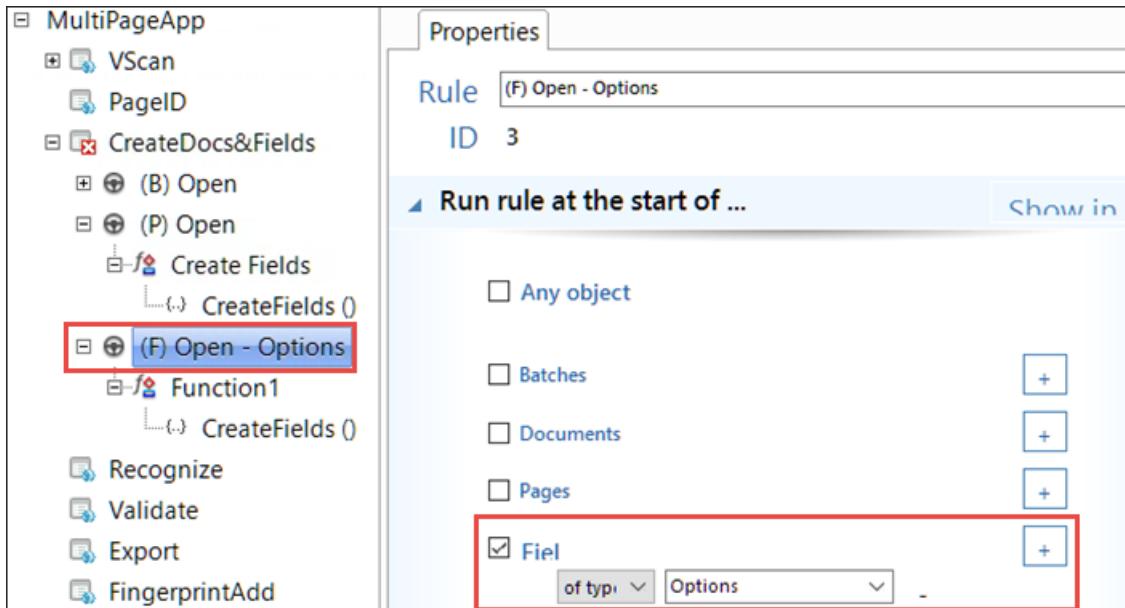
In this section, you add a rule for creating fields for the Options field and its child fields.

- 1. On the **Rulesets** tab, select the **CreateDocs&Fields** ruleset, click the lock icon on the toolbar, and click **Lock ruleset for editing** from the list.
- 2. Create a rule.
  - a. Right-click the **CreateDocs&Fields** ruleset and click **Add Rule** from the list.
  - b. Rename **Rule1** to: (F) Open - Options
- 3. Add an action to the Function1.
  - a. Expand the **CreateDocs&Fields > (P) Open > Create Fields** function and right-click the **CreateFields( )** action and click **Copy**.
  - b. Right-click the **CreateDocs&Fields > (F) Open - Options > Function1** and click **Paste**.

You can also add the action from the ApplicationObjects >  
Datacap.Libraries.ApplicationObjects.Actions action library to the Function1.



- 4. Assign the rule to the Options field.
  - a. In the **Rulesets** pane, select the **(F) Open - Options** rule.
  - b. In the **Properties** tab, expand the **Run rule at start of** node.
  - c. Select **Fields** and click the plus sign next to it to select the specific field.
  - d. For the **of type** field, select **Options** from the list.



## Troubleshooting

If **Options** is not in the list, the changes that you made in the Document hierarchy need to be refreshed. Restart Datacap Studio.

- 
- 5. In the toolbar, click the **Save changes** icon.
  - 6. Select the **CreateDocs&Fields** ruleset, click the lock icon on the toolbar of the **Rulesets** tab, and click **Publish ruleset** from the list.

## 2.3. Configure the Recognize ruleset

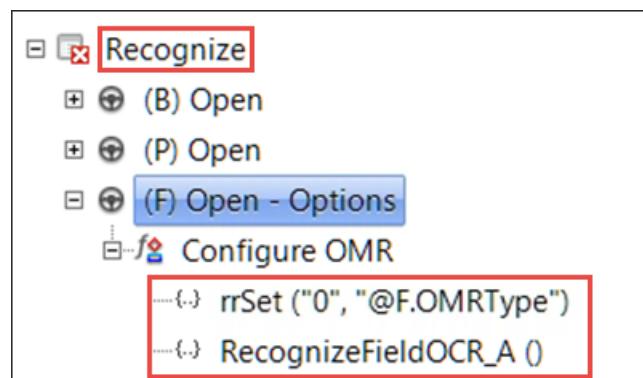
In this section, you configure the OMR recognition for the Options field.

- 1. On the **Rulesets** tab, lock the **Recognize** ruleset for editing.
- 2. Add a rule to the Recognize ruleset to recognize multiple option check boxes with OMR.
  - a. Right-click the **Recognize** ruleset and click **Add Rule** from the list.
  - b. Rename the following items.
    - **Rule1:** (F) Open – Options
    - **Function1:** Configure OMR
- 3. Assign the rule to the Options field.
  - a. In the **Rulesets** pane, select the **Recognize > (F) Open - Options** rule.
  - b. In the **Properties** tab > **Run rule at start of** node, select **Fields** and click the plus sign next to it.
  - c. For the **of type** field, select **Options** from the list.
- 4. Add actions to the Configure OMR function.
  - a. Select the **(F) Open - Options > Configure OMR** function.
  - b. From the **RuleRunnerLogic > Datacap.Libraries.RuleRunnerLogic.Actions** action library, add the **rrSet** action.
  - c. In the **Rulesets** pane, verify that the **rrSet ("", "")** action is added and select the action.
  - d. In the **Properties > Parameters** section, set the **string Source** field to **0** and set the **string Target** field to: **@F.OMRTypE**

The setting represents the Square background to read non-dropout checkbox.
- e. From the **RecognitionOCRA > RecognitionOCRA.OCRAActions** action library, add the **RecognizeFieldOCR\_A ()** action.

This field-level action recognizes the zoned field and retrieves the fields value.

- 5. For the **Configure OMR** function, verify that the actions are configured correctly in the order that is shown as shown in the screen capture.



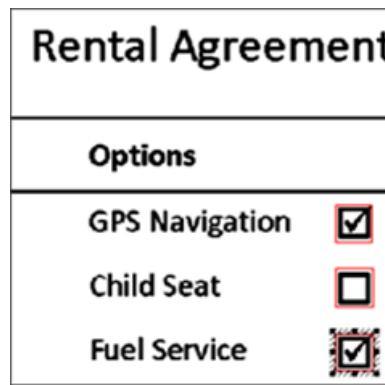
- 6. Save changes and publish the **Recognize** ruleset.

## 2.4. Configure zones

- \_\_\_ 1. On Datacap Studio, click **Zones** tab.
- \_\_\_ 2. In the **Fingerprints** tab, select **RentalForm\_FP > 557 [Rental Form]** and verify that the image is shown on the **Image** tab.  
The number on your system might be different.
- \_\_\_ 3. Draw zone for the Options field for OMR recognition.
  - \_\_\_ a. In the **Document hierarchy** pane, click the lock icon to enable DCO editing.
  - \_\_\_ b. Expand **MultiPageApp > Car Rental > Rental Form** and select the **Options** field.
  - \_\_\_ c. In the **Image** tab, draw the zone around the **Options** column.



- \_\_\_ 4. Draw zones for the Options subfields for OMR recognition.
  - \_\_\_ a. In the **Document hierarchy** pane, expand **Options** and select the **Navigation** field.
  - \_\_\_ b. In the **Image** tab, right-click and click **Zoom > Zoom In** to view the check boxes closely.
  - \_\_\_ c. Draw the zone around the checkbox for **GPS Navigation**.  
Draw zones of equal size for each checkbox.
  - \_\_\_ d. In the **Document hierarchy** pane, select the **Options > ChildSeat** field and on the **Image** tab, draw the zone around the checkbox for **Child Seat**.
  - \_\_\_ e. In the **Document hierarchy** pane, select the **Options > FuelService** field and on the **Image** tab, draw the zone around the checkbox for **Fuel Service**.

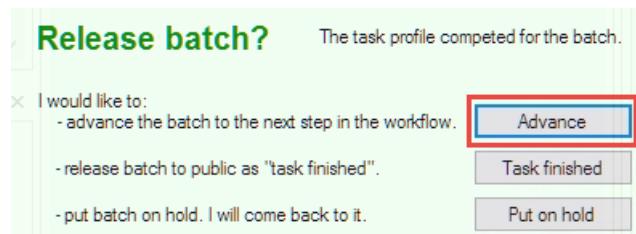


- \_\_\_ f. On the **Document hierarchy** pane toolbar, click the **Save** icon and click the lock icon to unlock DCO.

## Section 3. Test the application for OMR

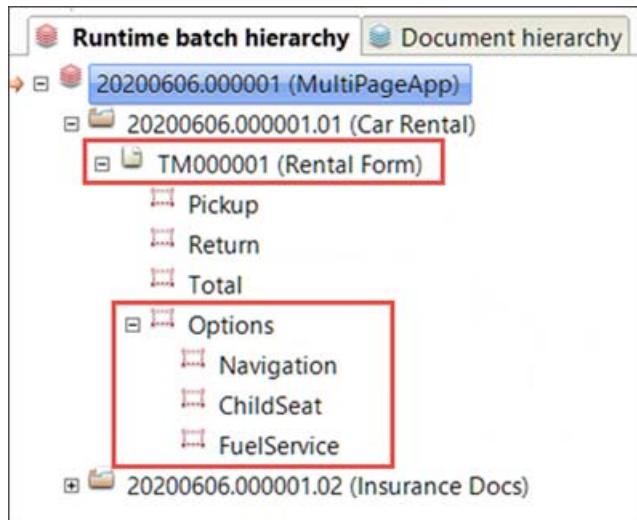
### 3.1. Run the VScan and PagID tasks

- \_\_\_ 1. Start a batch to test the application.
  - \_\_\_ a. In Datacap Studio, click the **Test** tab.
  - \_\_\_ b. On the **Workflow** tab, expand **MultiPageApp > MultiPageApp Job** right-click **VScan**, and select **New**.
  - \_\_\_ c. In the top toolbar, ensure that **VScan** is selected and click the **Process rules for target object** (green arrow icon) next to it.
- \_\_\_ 2. Run the Page ID task.
  - \_\_\_ a. In the **Release batch** dialog box, click **Advance** to move the batch to the **PagID** task.



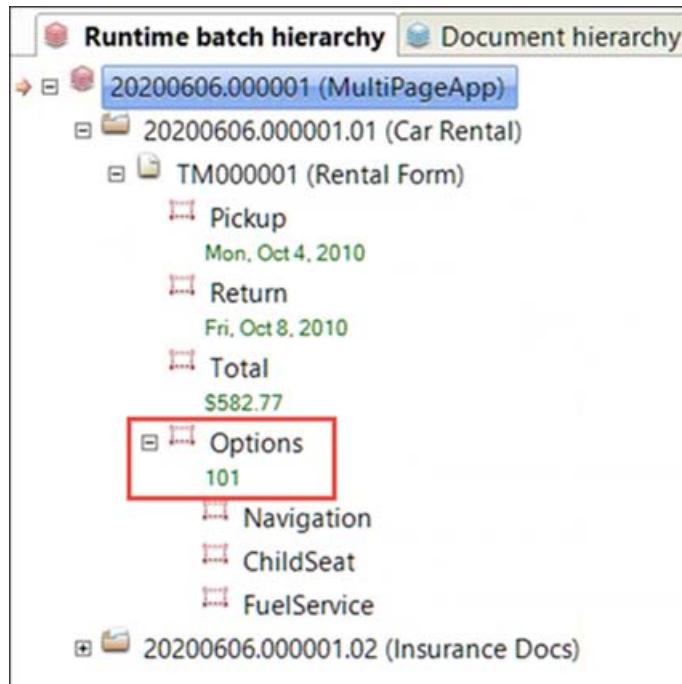
- \_\_\_ b. In the **Runtime batch hierarchy tab**, verify that it has two batch items now and that a pages are listed as **Other**.
- \_\_\_ c. In the top toolbar, ensure that **PagID** is selected, click the green arrow icon again to run the **PagID** task, and Wait for the process to complete.
- \_\_\_ 3. Verify the results of the PagID task.
  - \_\_\_ a. In the **Release batch** dialog box, click **Advance** to move the batch to the **Profiler** task.
  - \_\_\_ b. In the **Runtime batch hierarchy tab**, expand the batch node and verify that the documents are classified with the names: **Car Rental** and **Insurance Docs**
  - \_\_\_ c. Expand **Car Rental** and verify that the page is identified with the name: **Rental Form**
  - \_\_\_ d. Expand **Rental Form** and verify that the page contains the four fields that you defined: **Pickup**, **Return**, **Total**, and **Options**.

- \_\_ e. Expand **Options** and verify that the field contains the three fields that you defined: **Navigation**, **ChildSeat**, and **FuelService**.



### 3.2. Run the Profiler and Verify tasks

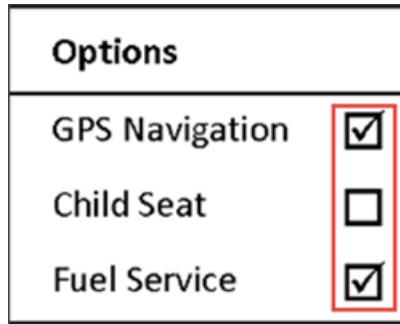
- In the top toolbar, ensure that **Profiler** is selected and click the green arrow icon to run the **Profiler** task.  
This step extracts the data from the fields.
- In the **Release batch** dialog box, click **Keep running**.
- In the **Runtime batch hierarchy tab**, expand the **MultiPageApp > Car Rental** document.
- Expand the page (**Rental Form**) node, and verify that the **Options** field is populated with a value: **101**



The value for the three check boxes is: **101** where the value **1** indicates that the checkbox is selected. Value **0** indicates that the checkbox is not selected.

When the data is exported to a file, the dictionary values are used.

- \_\_\_ 5. In the **Runtime batch hierarchy tab**, click the **Options** field and notice that the image is shown in the Image tab on the middle pane.
- \_\_\_ 6. Compare the results with the image, where two fields are selected and one is not selected.

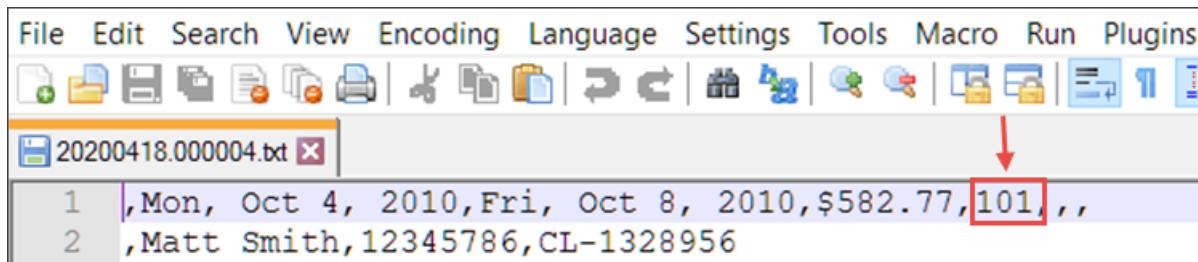


- \_\_\_ 7. Advance the batch to Verify task and process the batch
  - \_\_\_ a. On the **Workflow** tab, expand **MultiPageApp > MultiPageApp Job > Profiler**, right-click the batch (**running**), and select **Advance**.  
The workflow moves to the Verify step.
  - \_\_\_ b. In the top toolbar, ensure that **Verify** is selected and click the green arrow icon to run the **Verify** task.

### 3.3. Run the Export task and check the export file

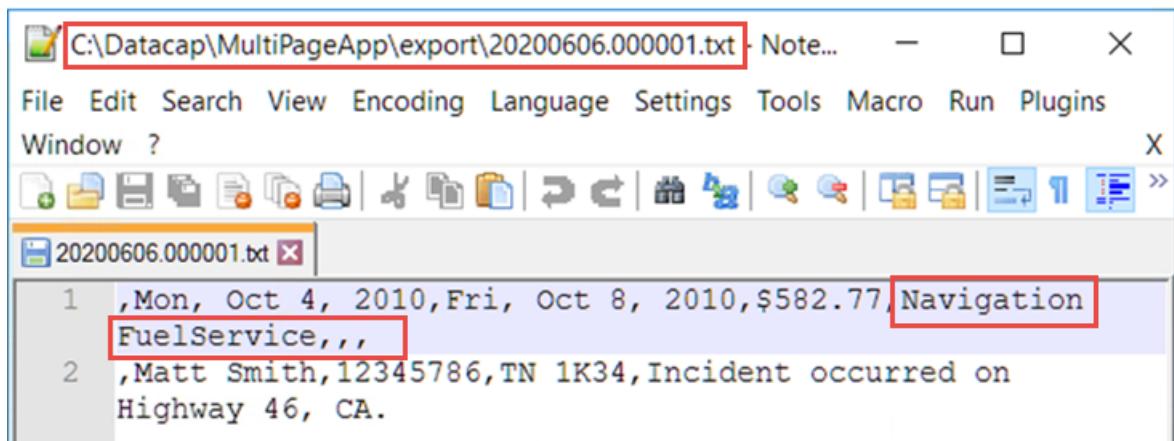
- \_\_\_ 1. In the **Release batch** dialog box, click **Advance** to move the batch to the **Export** task.
- \_\_\_ 2. In Datacap Studio **Test** tab, on the **Runtime batch hierarchy** pane, take a note of the batch name: <Date>. <batch number> (Example: 20200606.000001).  
You use this value to identify the export file in the following steps.
- \_\_\_ 3. In the top toolbar, ensure that **Export** is selected and click the green arrow icon to run the task.
- \_\_\_ 4. In the **Release batch** dialog box, click **Advance** to complete the batch.
- \_\_\_ 5. Check the export text file.
  - \_\_\_ a. In Windows Explorer, browse to the `C:\Datacap\MultiPageApp\export` folder and verify that a text file with your batch name is created with today's date.
  - \_\_\_ b. Open the file in Notepad++ and check that the file contains the comma-separated values for the following fields:
    - Pickup date, return date, total, and Options for the Rental Form page in the first line
    - Name, policy number, license, and incident for the Claim Page in the second line

A sample file that shows the options values (101) without the dictionary configuration.



The screenshot shows the Datacap Studio interface with a menu bar and toolbar. A file named "20200418.000004.txt" is open. In the text editor, the value "101" is highlighted with a red box and has a red arrow pointing down from the toolbar towards it. The text content includes two lines: "1 ,Mon, Oct 4, 2010,Fri, Oct 8, 2010,\$582.77,101,,," and "2 ,Matt Smith,12345786,CL-1328956".

A sample file that shows the options values as text with the dictionary configuration.



The screenshot shows the Datacap Studio interface with a menu bar and toolbar. A file named "20200606.000001.txt" is open. The value "101" is replaced by the word "Navigation" in the text, which is highlighted with a red box. Another value, "FuelService", is also highlighted with a red box. The text content includes two lines: "1 ,Mon, Oct 4, 2010,Fri, Oct 8, 2010,\$582.77,Navigation" and "2 ,Matt Smith,12345786,TN 1K34,Incident occurred on Highway 46, CA..".

- c. Close the file.
- 6. In Datacap Studio, at the upper right corner, click **Exit**.

Restarting Datacap Studio before the next exercise ensures that any running batches are cleared.

## End of exercise

## Exercise review and wrap-up

In this exercise, you configured OMR for your application to recognize multiple choice check boxes and extract values.

# Exercise 7. Exporting data to a FileNet Content Manager repository

## Estimated time

01:15

## Overview

This exercise teaches how to configure a Datacap application to export scanned documents and data to a FileNet Content Manager repository.

## Objectives

After completing this exercise, you should be able to:

- Export data and documents to a FileNet Content Manager repository
- Verify the exported documents

## Introduction

In a previous exercise, you created the MultiPageApp application. It has an Export ruleset that exports the field values to a text file. In this exercise, you configure the application to export data and documents to a FileNet Content Manager repository.

This exercise includes the following sections:

- [Section 1, "Export data to a FileNet Content Manager repository"](#)
- [Section 2, "Test the application"](#)

## Requirements

This exercise builds on [Exercise 6, "Configuring an application for recognition with OMR"](#). Before starting this exercise, you must complete the [Exercise 6, "Configuring an application for recognition with OMR"](#).

A solution of the Datacap application for the exercise is provided in C:\Datacap folder with the application name, **SolutionMultiPageApp**. You can access this application like any other Datacap application.

# Section 1. Export data to a FileNet Content Manager repository

In this section, you configure the application to export the data to the FileNet Content Manager repository. Your student system is configured for FileNet Content Manager.



## Note

A solution of the Datacap application for this exercise is provided in C:\Datacap folder with the application name: **SolutionMultiPageApp**

You can access this application like any other Datacap application.

## 1.1. Start the Datacap Studio

- \_\_\_ 1. Before you begin the exercise, ensure that the components are running. If you did not start the components already, refer to [Section 1, "Prepare your system: Start the components"](#) to start them.  
If you restart your system, you must start all the components.
- \_\_\_ 2. Log in to Datacap Studio.
  - \_\_\_ a. Double-click the **Datacap Studio** shortcut on the Windows desktop.
  - \_\_\_ b. In the **Applications** dialog box, select **MultiPageApp** and click **Next**.
  - \_\_\_ c. On the **Taskmaster Login** page, enter the following values:
    - User ID:** admin
    - Password:** admin
    - Station ID:** 1
  - \_\_\_ d. Click **Finish**.

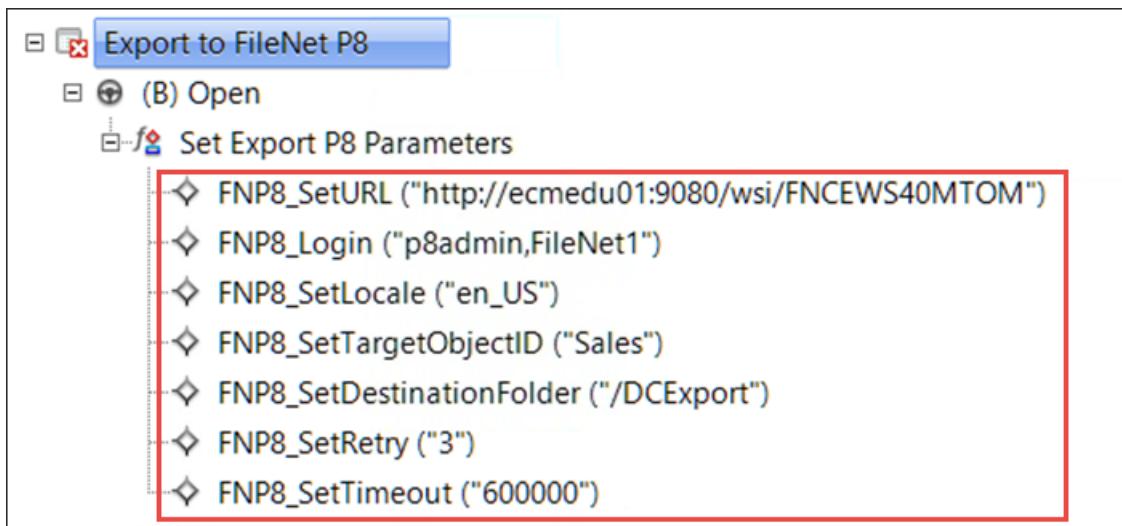
Datacap Studio opens with the **Rulemanager** tab selected.

## 1.2. Add the Export to FileNet P8 ruleset

In this section, you create the Export ruleset that exports the data to the FileNet Content Manager repository. You define all the parameters that are required for this task in a batch level rule.

- \_\_\_ 1. In the **Rulesets** pane, select **MultiPageApp** and click **Add child object** icon.  
A new ruleset node is added to **MultiPageApp** with the name **Ruleset1** and it has a red x flag to indicate that it is locked for editing.
- \_\_\_ 2. Rename the following objects:
  - **Ruleset1:** Export to FileNet P8
  - **Rule1:** (B) Open
  - **Function1:** Set Export P8 Parameters

- \_\_\_ 3. Assign the (B) Open rule to batches.
  - \_\_\_ a. In the **Rulesets** pane, select the **Export to FileNet P8 > (B) Open** rule.
  - \_\_\_ b. In the **Properties** tab > **Run rule at start of** node, select **Batches**.
- \_\_\_ 4. Add actions to the Set Export P8 Parameters function.
  - \_\_\_ a. Select the **Set Export P8 Parameters** function and from the **FileNetP8** action library, add the following actions and set the parameters.
    - FNP8\_SetURL ("http://ecmedu01:9080/wsi/FNCEWS40MTOM")
    - FNP8\_Login ("p8admin,FileNet1")
    - FNP8\_SetLocale ("en\_US")
    - FNP8\_SetTargetObjectID ("Sales")
    - FNP8\_SetDestinationFolder ("/DCExport")
    - FNP8\_SetRetry ("3")
    - FNP8\_SetTimeout ("600000")
  - \_\_\_ b. Verify that the completed ruleset contains all the actions with correct parameters as shown on the following screen capture.



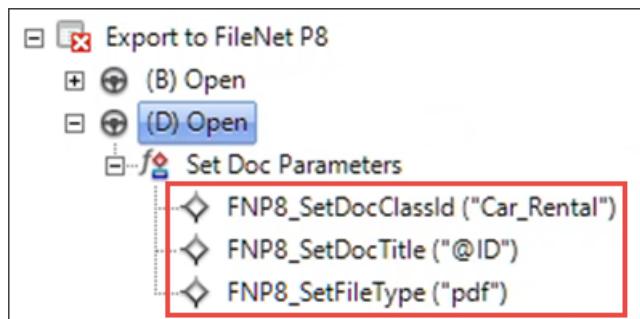
- \_\_\_ 5. In the **Rulesets** tab, click the **Save changes** icon in the toolbar.

### 1.3. Add a document level rule to set document parameters

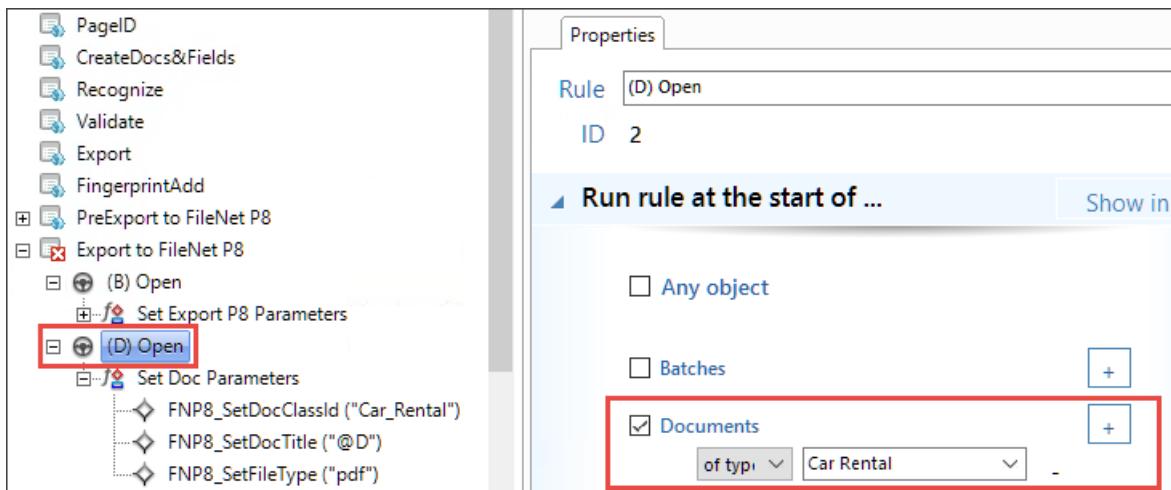
In this section, you add a document level rule to the Export to FileNet P8 ruleset to define the document class, title, and file type for upload.

- \_\_\_ 1. In the **Rulemanager** tab, right-click the **Export to FileNet P8** ruleset and click **Add Rule** from the list.
- \_\_\_ 2. Rename the new rule and its function:
  - **Rule1: (D) Open**
  - **Function1: Set Doc Parameters**

- 3. Add actions to the Set Doc Parameters function.
- a. Select the **Set Doc Parameters** function and from the **FileNetP8** action library, add the following three actions and set the parameters.
- FNP8\_SetDocClassId ("Car\_Rental")
  - FNP8\_SetDocTitle ("@ID")
  - FNP8\_SetFileType ("pdf")
- b. Verify that the completed ruleset contains the actions as shown on the following screen capture.



- 4. Assign the (D) Open rule to the Car Rental document.
- a. In the **Rulesets** pane, select the **Export to FileNet P8 > (D) Open** rule.
- b. In the **Properties** tab, expand the **Run rule at start of** node.
- c. Select **Documents** and click the plus sign next to it to select the specific field.
- d. For the **of type** field, select **Car Rental** from the list.

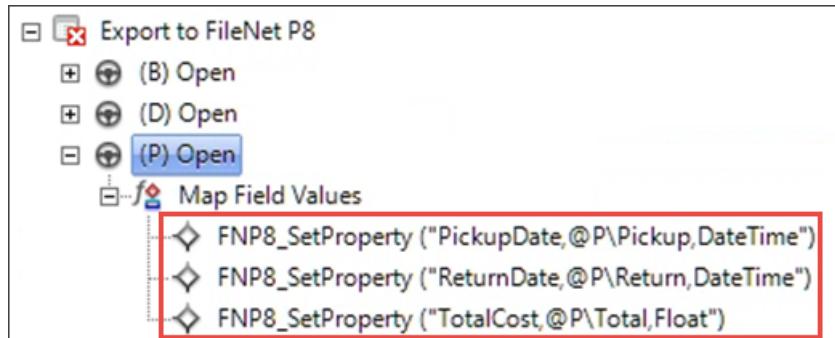


- 5. In the **Rulesets** tab, click the **Save changes** icon in the toolbar.

## 1.4. Add a page level rule to map the field values

In this section, you add a page level rule to the Export to FileNet P8 ruleset to map the field values to the document properties for upload.

- 1. Right-click the **Export to FileNet P8** ruleset and click **Add Rule** from the list.
- 2. Rename the new rule and its function:
  - **Rule1:** (P) Open
  - **Function1:** Map Field Values
- 3. Add actions to the Map Field Values function.
  - a. Select the **Map Field Values** function and from the **FileNetP8** action library, add the following three actions and set the parameters.
    - FNP8\_SetProperty("PickupDate,@P\Pickup,DateTime")
    - FNP8\_SetProperty("ReturnDate,@P\Return,DateTime")
    - FNP8\_SetProperty("TotalCost,@P\Total,Float")
  - b. Verify that the completed ruleset contains the actions in the order they are shown on the following screen capture.



### Note

The parameters values for the **FNP8\_SetProperty** action contain Property ID from the repository, the corresponding DCO Field value, and data type for the property. For example, (“**PickupDate,@P\Pickup,DateTime**”) contains PickupDate from the repository, Pickup from the DCO field value, and DateTime for the data type.

- 
- 4. Assign the (P) Open rule to the Rental Form page.
    - a. In the **Rulesets** pane, select the **Export to FileNet P8 > (P) Open** rule.
    - b. In the **Properties > Run rule at start of** node, select **Pages** and click the plus sign next to it.

- \_\_\_ c. For the **of type** field, select **Rental Form** from the list.

The screenshot shows the 'Run rule at the start of...' configuration. On the left, there's a tree view with nodes like 'scan ()', 'PageID', 'CreateDocs&Fields', etc. Under 'Export', there are three options: '(B) Open', '(D) Open', and '(P) Open'. The '(P) Open' node has a child node 'Map Field Values' which contains three entries: 'FNP8\_SetProperty ("PickupDate,@P\Pickup,DateTime")', 'FNP8\_SetProperty ("ReturnDate,@P\Return,DateTime")', and 'FNP8\_SetProperty ("TotalCost,@P\Total,Float")'. On the right, the 'Properties' panel shows 'Rule (P) Open' and 'ID 3'. The 'Run rule at the start of...' section is expanded, showing checkboxes for 'Any object', 'Batches', and 'Documents'. Below these is a section for 'Pages' with a checked checkbox and a dropdown menu set to 'Rental Form'. This 'Pages' section is highlighted with a red box.



### Note

You can also define this rule at the field level because you are using the field values. But field level requires one rule for each field. By configuring at a page level, you can define one rule for all the fields.

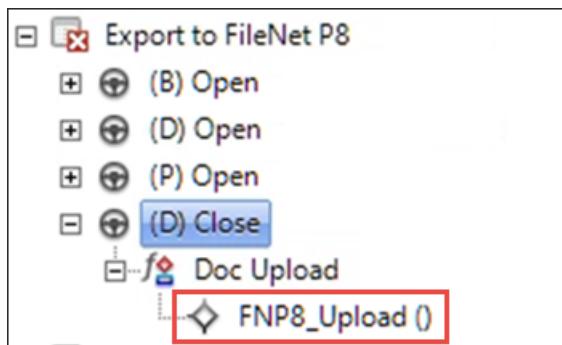
- \_\_\_ 5. In the **Rulesets** tab, click the **Save changes** icon in the toolbar.

## 1.5. Add a document level rule to upload the scan document

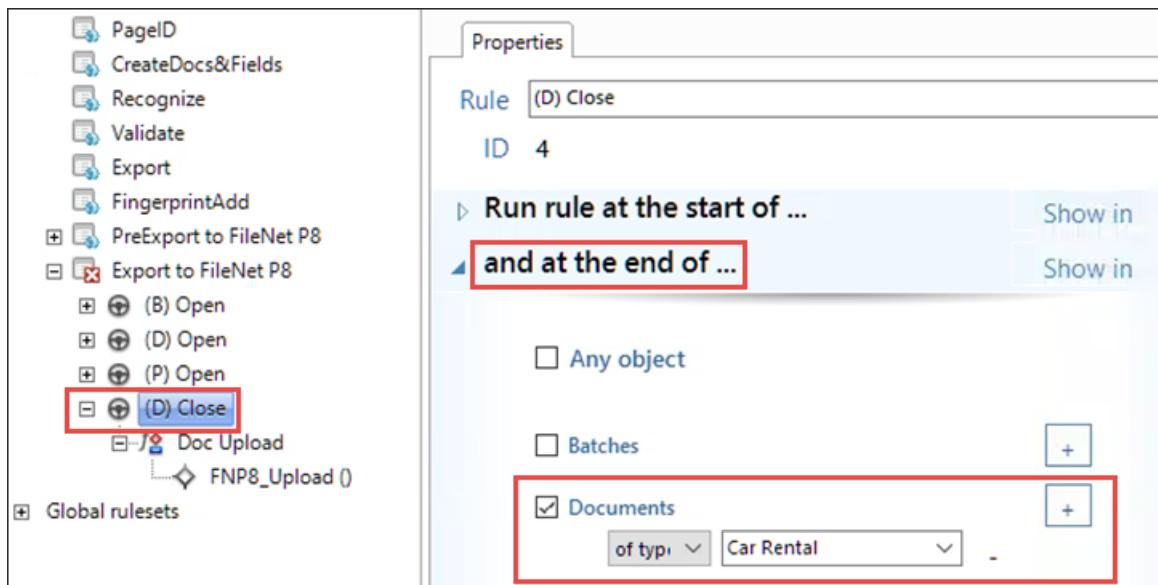
In this section, you add a document level rule to the Export to FileNet P8 ruleset to upload the document.

- \_\_\_ 1. Right-click the **Export to FileNet P8** ruleset and click **Add Rule** from the list.
- \_\_\_ 2. Rename the new rule and its function:
  - **Rule1:** (D) Close
  - **Function1:** Doc Upload
- \_\_\_ 3. Add an action to the Doc Upload function.
  - \_\_\_ a. Select the **Doc Upload** function.
  - \_\_\_ b. On the **Actions library** tab, from the **FileNetP8** library, add the **FNP8\_Upload** action to the function.

- \_\_\_ c. Verify that the completed ruleset contains the action as shown on the following screen capture.



- \_\_\_ 4. Assign the (D) Close rule to the Car Rental document.
- \_\_\_ a. In the **Rulesets** pane, select the **Export to FileNet P8 > (D) Close** rule.
- \_\_\_ b. In the **Properties** tab, expand the **and at the end of** node.  
Ensure that you select **at the end** (and not at the start).
- \_\_\_ c. Select **Documents** and click the plus sign next to it to select the specific field.
- \_\_\_ d. For the **of type** field, select **Car Rental** from the list.

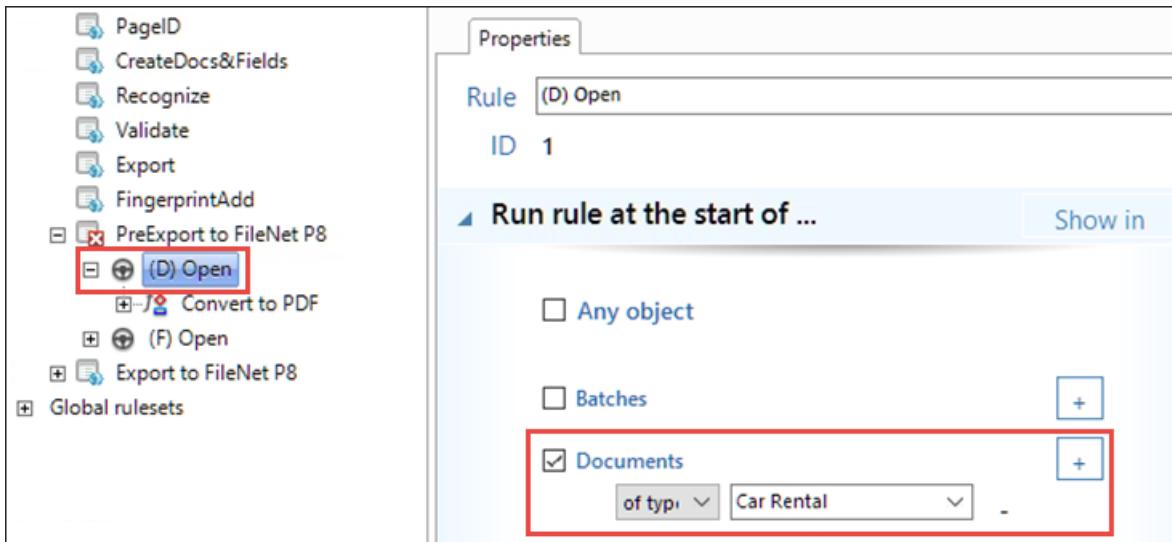


- \_\_\_ e. In the **Rulesets** tab, click the **Save changes** icon in the toolbar.
- \_\_\_ 5. Select the **Export to FileNet P8** ruleset, click the lock icon on the toolbar of the **Rulesets** tab, and click **Publish ruleset** from the list.

## 1.6. Add a ruleset to prepare the document for the upload

The scanned images are in TIF file format. You want to export the document in PDF format. In this section, you add a ruleset to convert the TIF file format to PDF.

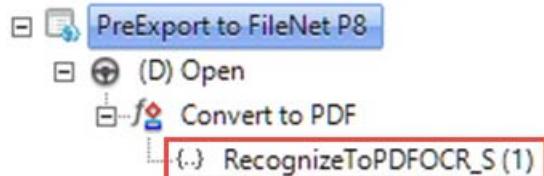
- \_\_\_ 1. In the **Rulesets** pane, select **MultiPageApp** and click **Add child object**  icon.  
A new ruleset node is added to **MultiPageApp** with the name **Ruleset1** and it is locked for editing.
- \_\_\_ 2. Rename the following objects:
  - **Ruleset1**: PreExport to FileNet P8
  - **Rule1**: (D) Open
  - **Function1**: Convert to PDF
- \_\_\_ 3. Assign the (D) Open rule to the Car Rental document.
  - \_\_\_ a. In the **Rulesets** pane, select the **PreExport to FileNet P8 > (D) Open** rule.
  - \_\_\_ b. In the **Properties** tab, expand the **Run rule at start of** node.
  - \_\_\_ c. Select **Documents** and click the plus sign next to it to select the specific field.
  - \_\_\_ d. For the **of type** field, select **Car Rental** from the list.



- \_\_\_ 4. Add an action to the Convert to PDF function.
  - \_\_\_ a. Select the **Convert to PDF** function.
  - \_\_\_ b. On the **Actions library** tab, from the **ocr\_sr > Datacap.Libraries.ScansoftR.Actions** action library, add the **RecognizeToPDFOCR\_S** action to the function.
  - \_\_\_ c. In the **Rulesets** pane, select the **RecognizeToPDFOCR\_S ("")** action and in the **Properties** tab, set the **int OutputPDFType** field to: **1**

This action converts a scanned image file (TIF) to a PDF file. The numeric value 1 for the parameter indicates the PDF output type (generic PDF document).

- \_\_\_ d. Verify that the completed ruleset contains the action as shown on the following screen capture.

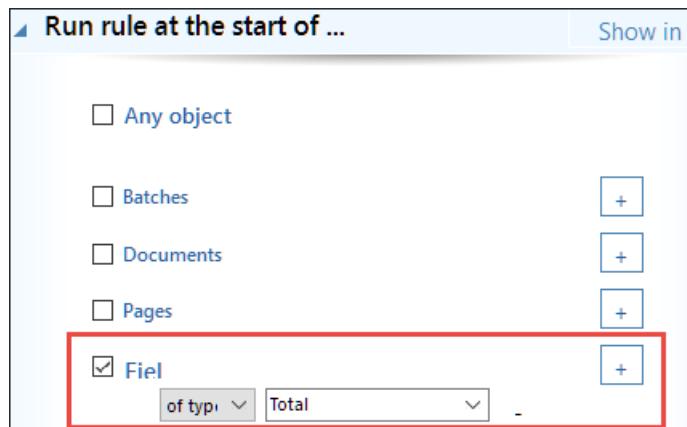


- \_\_\_ 5. In the **Rulesets** tab, click the **Save changes** icon in the toolbar.

## 1.7. Add a rule to reformat the currency value

When you export the data, the data format must be compatible on the repository. For example, the currency field that is extracted in Datacap has a dollar symbol. You need to reformat it to a float data type. In this section, you add a rule to reformat the currency value.

- \_\_\_ 1. Add a field level rule to the ruleset to reformat the currency field value (remove dollar symbol) for upload.
  - \_\_\_ a. In the **Rulesets** tab, right-click the **PreExport to FileNet P8** ruleset and click **Add Rule** from the list.
  - \_\_\_ b. Rename the new rule and its function:
    - Rule1:** (F)Open
    - Function1:** Reformat currency
- \_\_\_ 2. Assign the (F) Open rule to the Total field.
  - \_\_\_ a. In the **Rulesets** pane, select the **PreExport to FileNet P8 > (F) Open** rule.
  - \_\_\_ b. In the **Properties > Run rule at start of** node, select **Fields** and click the plus sign.
  - \_\_\_ c. For the **of type** field, select **Total** from the list.



- \_\_\_ 3. Add an action to the Reformat currency function.
  - \_\_\_ a. Select the **Reformat currency** function.
  - \_\_\_ b. On the **Actions library** tab, from the **ValidationsAndTextAdjustments > Datacap.Libraries.ValidationsAndTextAdjustments.Actions** action library, add the **DeleteAllCharactersSelected** action to the function.
  - \_\_\_ c. In the **Rulesets** pane, select the **DeleteAllCharactersSelected ("\$","","")** action.
  - \_\_\_ d. In the **Properties** tab > **Parameters** section, for the **string CharactersToDelete** field, enter: \$



### Note

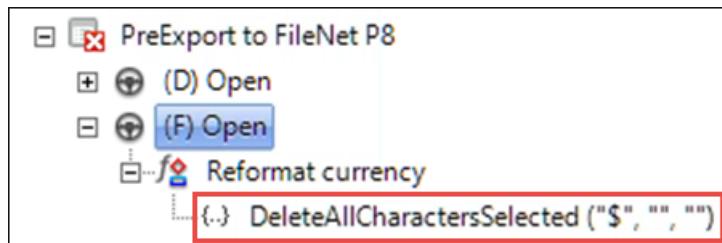
The **DeleteAllCharactersSelected ("\$","","")** action deletes the specified (\$) character from the field value.

The first parameter (**string CharactersToDelete**) indicates the character to be deleted.

The second parameter (**string StartIndex**) specifies the position. If the value is blank, it deletes the first position.

Third parameter (**string Count**) specifies the number of times to delete. A blank value assumes the default and deletes one time.

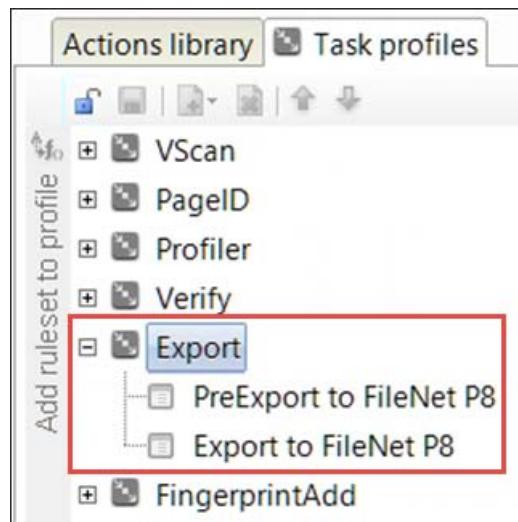
- \_\_\_ e. Verify that the completed ruleset contains the action and the parameters as shown on the following screen capture.



- \_\_\_ 4. Save the changes and publish the ruleset.
  - \_\_\_ a. Click the **Move up** icon on the toolbar to move up the ruleset, and place it between **FingerprintAdd** and **Export to FileNet P8** in the list.
  - \_\_\_ b. Click **Save changes** and publish the **PreExport to FileNet P8** ruleset.

## 1.8. Add the rulesets to the Export task profile

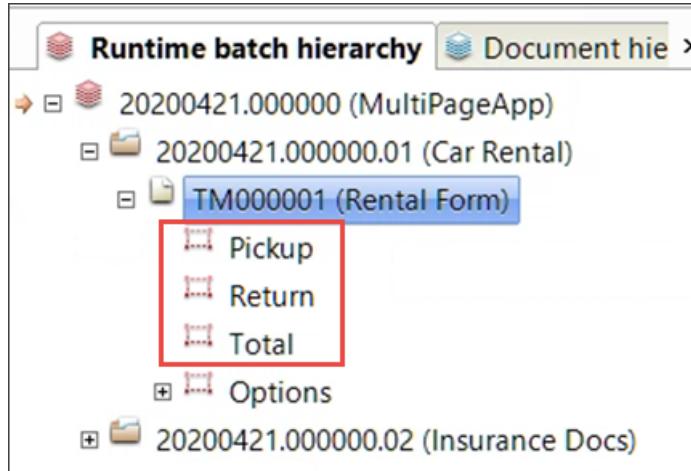
- \_\_\_ 1. Click the **Task profiles** tab, lock it by clicking the lock  icon.
- \_\_\_ 2. Expand the **Export** task profile, select the **Export** ruleset, and click **Remove** (red X) icon.  
Note that this ruleset exports the fields as comma-separated values into a text file as you learned in the previous section.
- \_\_\_ 3. Add the Export to FileNet P8 rulesets to the Export task profile.
  - \_\_\_ a. Select the **Export** task profile and on the **Rulesets** tab, select the **PreExport to FileNet P8** ruleset.
  - \_\_\_ b. Click the **Add ruleset to profile** vertical bar on the right of the **Rulesets** pane.
- \_\_\_ 4. Repeat the steps to add the **Export to FileNet P8** ruleset to the **Export** task profile.
- \_\_\_ 5. On the **Task profiles** tab toolbar, click the **Save changes**  icon and unlock **Task profiles**.
- \_\_\_ 6. Expand the **Export** task profile and verify that it contains the two rulesets.



## Section 2. Test the application

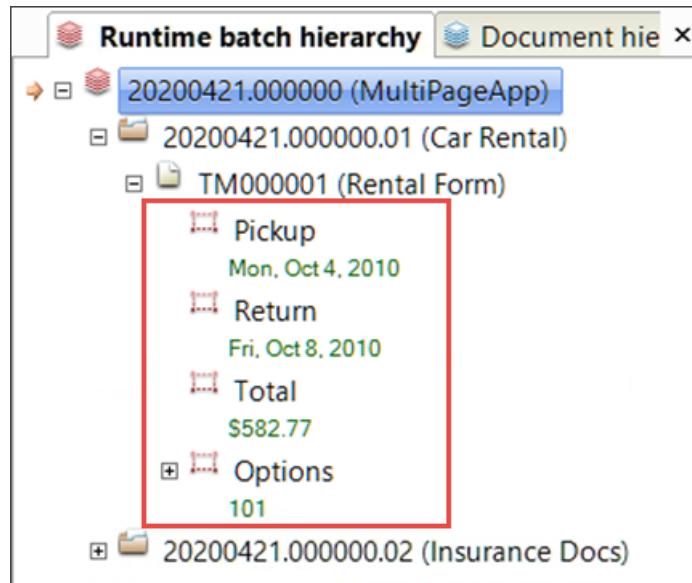
### 2.1. Run the VScan and PageID tasks

- \_\_\_ 1. In Datacap Studio top toolbar, click the **Test** tab.
- \_\_\_ 2. Start a batch to test the application.
  - \_\_\_ a. On the **Workflow** tab, expand **MultiPageApp Job** right-click **VScan**, and select **New**.
  - \_\_\_ b. In the top toolbar, ensure that **VScan** is selected and click the **Process rules for target object**  icon next to it.
- \_\_\_ 3. Run the Page ID task.
  - \_\_\_ a. In the **Release batch** dialog box, click **Advance** to move the batch to the **PageID** task.
  - \_\_\_ b. In the **Runtime batch hierarchy tab**, verify that it has batch items and pages are listed as **Other**.
  - \_\_\_ c. In the top toolbar, ensure that **PageID** is selected, click the green arrow icon again to run the **PageID** task.
- \_\_\_ 4. Verify the results of the PageID task.
  - \_\_\_ a. In the **Release batch** dialog box, click **Advance** to move the batch to the **Profiler** task.
  - \_\_\_ b. In the **Runtime batch hierarchy tab**, expand the batch node and verify that the documents are classified with the names: **Car Rental and Insurance Docs**
  - \_\_\_ c. Expand **Car Rental** and verify that the page is identified with the name: **Rental Form**
  - \_\_\_ d. Expand **Rental Form** and verify that the page contains the **Pickup, Return, Total, and Options** fields.



## 2.2. Run the Profiler and Verify tasks

- \_\_\_ 1. Run the Profiler task.
  - \_\_\_ a. In the top toolbar, ensure that **Profiler** is selected and click the green arrow icon to run the **Profiler** task.  
This step extracts the data from the fields.
  - \_\_\_ b. In the **Release batch** dialog box, click **Keep running**.
- \_\_\_ 2. Check the Profiler task results.
  - \_\_\_ a. In the **Runtime batch hierarchy tab**, expand the **MultiPageApp > Car Rental** document.
  - \_\_\_ b. Expand the page (**Rental Form**) node, and verify that the fields are populated with values.



- \_\_\_ 3. Advance the batch to Verify task and process the batch.
  - \_\_\_ a. On the **Workflow** tab, expand **MultiPageApp Job > Profiler**, right-click the batch (**running**), and select **Advance**.  
The workflow moves to the Verify step.
  - \_\_\_ b. In the top toolbar, ensure that **Verify** is selected and click the green arrow icon to run the **Verify** task.

## 2.3. Run the Export task

- 1. Process the Export task. In the **Release batch** dialog box, click **Advance** to move the batch to the **Export** task.
- 2. In Datacap Studio **Test** tab, on the **Runtime batch hierarchy** pane, take a note of the batch name: <Date>.<batch number> (Example: 20200421.000001).  
The batch name is assigned as the title to the export document.
- 3. In the top toolbar, ensure that **Export** is selected and click the green arrow icon.
- 4. In the **Release batch** dialog box, click **Advance** to complete the batch.
- 5. On the upper right corner of the page, click **Exit** to close Datacap Studio.

## 2.4. Check the export documents

In this section, you log in to the FileNet Content Manager repository and verify the export documents. The IBM Content Navigator is the client for the repository.

- 1. In the Firefox browser, open the IBM Content Navigator client.
    - a. Click the **Sample Desktop** shortcut or enter the following URL:  
`http://ecmedu01:9081/navigator/`
    - b. Enter the following values:
      - User ID: P8admin
      - Password: FileNet1
    - c. Click **Log In**.
- In IBM Content Navigator, the FileNet Content Manager **Sales** repository opens in the **Browse** view.
- Browse** is shown on the upper left and **Sales** is shown on the upper right.
- 2. From the left pane, select the **Sales > DCExport** folder.
  - 3. Verify that a document is listed with today's date and with batch ID as the title.
    - a. Single-click the document to see the properties shown in the rightmost pane.
    - b. Verify that the properties contain values that were extracted from the scan image.

The screenshot shows the IBM FileNet Content Manager interface. On the left, there is a navigation tree under the 'Sales' category, with 'DCEExport' selected. In the center, a list view shows a single item named '20200607.000000.01'. On the right, the 'Properties' pane is open, displaying the following information:

Class:	Car_Rental
Document Title:	20200607.000000.01
Return Date:	10/7/2010, 8:00 PM
Pickup Date:	10/3/2010, 8:00 PM
Total Cost:	582.77

Below the properties, it says 'Added By: p8admin'. The 'MIME Type' field in the properties pane is highlighted with a red box and contains the value 'application/pdf'.

- c. In the **Properties** pane, expand the **System Properties** and check that the **MIME type** is **application/pdf**.

MIME Type: application/pdf

The document is exported to the IBM FileNet Content Manager repository as a PDF document.

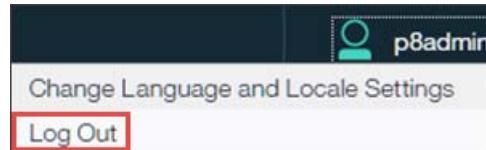
- 4. Double-click the document to open it in the viewer, view the stored document, and close the viewer.

- \_\_\_ 5. Optionally, in the browse view, right-click the document and click **Download > As Original** to verify that you are able to download the PDF file.

Name	Size	Modified By	Class:	
20200421.000001.01			Car_Rental	

Document Title: 20200421.00  
Pickup Date: 10/3/2010, 8:00 AM  
Return Date: 10/7/2010, 8:00 AM  
Total Cost: 582.77

- \_\_\_ 6. Log out of IBM Content Navigator.  
\_\_\_ a. On the banner, click the user and select **Logout** from the list.



- \_\_\_ b. Click **Logout** and close the browser.

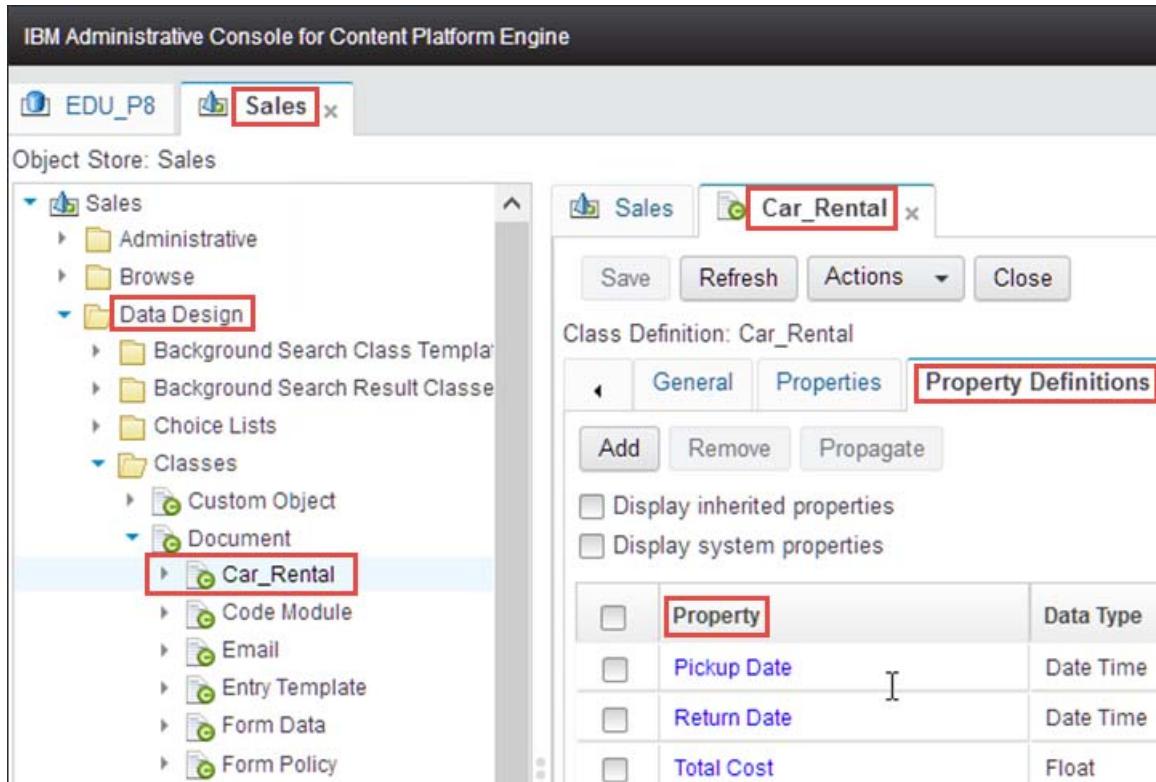


## Troubleshooting

If you receive any error, then the common problems are:

- Mismatch between the Datacap field names and the **Car\_Rental** Document class in the repository.
  - One of the property values is not initialized correctly.
  - One of the property data types is incorrect.
  - Password is not correct.
  - Incorrect repository information.
- \_\_\_ a. Open the **export\_rrs.log** file in the batches folder of your latest batch. Search for the word **Error or Abort**. This search should point you to the field that is causing the problem.
- \_\_\_ b. Check that you typed the repository field names and Datacap data field names correctly in the actions parameters for the **Export to FileNet P8 > (P) Open > Map Field Values** function.
- \_\_\_ c. If you need to check the **Document** class property names:
  - In Firefox, click the **ACCE** bookmark or enter the **URL**: <http://ecmedu01:9080/acce>

- Log in as: p8admin/FileNet1
- On the **EDU\_P8** domain tab, expand **Object Stores** on the left pane and click **Sales**.
- In the **Sales** tab, expand **Data Design > Classes > Document** and click **Car\_Rental**.
- In the **Car\_Rental** tab, click the **Property Definitions** subtab.



- In the **Property** column, click a property.

- In the **Property Definition** page, look up the values for **Symbolic name** and **Data type**.  
The Symbolic name that is entered in FastDoc must match the Symbolic Name on the server.

Property Definition

General	Alias IDs	More	Modification Access
Display name : ?	Pickup Date		
Symbolic name : ?	PickupDate		
Description : ?	Pickup Date		
Data type : ?	Date Time		
Cardinality :	Single		
Primary Id :	{8F62656B-6693-4297-8D89-DF04BB781DD3}		

**End of exercise**

## Exercise review and wrap-up

In this exercise, you copied an existing application that is based on the Forms template and created a new application for processing multiple documents. In your application, you configured OMR to extract field values from multiple choice check boxes. You also configured the application to export data and documents to a FileNet Content Manager repository.

# Exercise 8. Creating page layouts

## Estimated time

01:00

## Overview

In this exercise, you learn how to build page layouts, test the application, and verify the layout XML file.

## Objectives

After completing this exercise, you should be able to:

- Configure a Datacap application to create page layouts
- Test the application and verify the layout XML file

## Introduction

Datacap processes the PDF and image files to produce a layout XML file for each page. The Full page OCR recognition or PDF conversion actions are used to create the layout XML structure. More actions can be used to enhance the internal format of the layout.

The layout XML file groups text into blocks as a person would look at the document. Each block might have the default type of block or a specific type such as title or table.

A key feature in the Datacap Insight application is that data can be identified and extracted before the pages are identified or classified. This feature allows the information in the document to be used for identification and classification.

In this exercise, you configure rulesets to convert a PDF file into TIFF and build page layouts.

This exercise includes the following sections:

- [Section 1, "Configure an application to create page layouts"](#)
- [Section 2, "Test the application and verify the layout file"](#)

## Requirements

None

## Section 1. Configure an application to create page layouts

In the previous exercises, you learned how to build a Datacap application. For this exercise, the student system already contains a partially built application. In this section, you add the Datacap actions that build the layout XML for scanned documents.



### Note

Some of the following subsections provides instructions for creating rulesets. If you are already familiar with the steps, you can refer to the tables or the screen captures and skip the instructions.

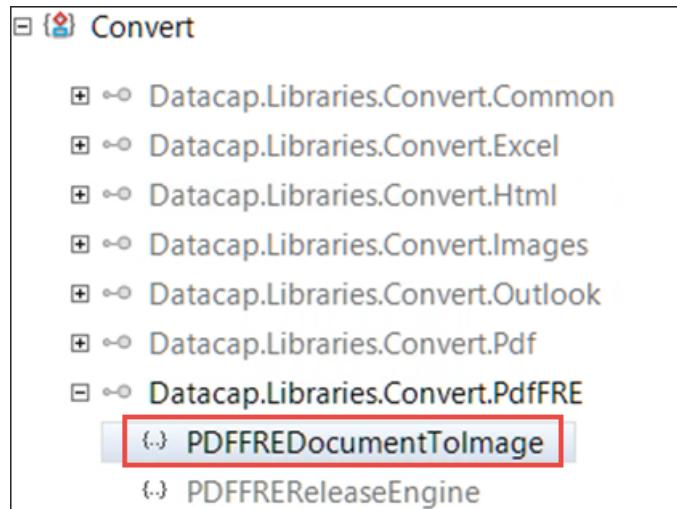
A solution of the Datacap application for this exercise is provided in C:\Datacap folder with the application name, **SolutionDataExtraction**. You can access this application like any other Datacap application.

### 1.1. Add the Convert PDF ruleset

You configure a ruleset to convert a PDF file to TIFF format. With TIFF, you can use standard Datacap methods to process and verify documents and data.

- 1. Before you begin the exercise, ensure that the components are running. If you did not start the components already, refer to [Section 1, "Prepare your system: Start the components"](#) to start them.  
If you restart your system, you must start all the components.
- 2. Log in to Datacap Studio.
  - a. On the Windows desktop, double-click the **Datacap Studio** shortcut.
  - b. In the **Applications** dialog box, select **DataExtraction** and click **Next**.
  - c. On the **TaskMaster Login** page, enter the following values:
    - **User ID:** admin
    - **Password:** admin
    - **Station ID:** 1
  - d. Click **Finish**.
- 3. In the **Rulesets** tab, select the **Convert PDF** ruleset, click the lock icon, and select **Lock ruleset for editing** from the list.
- 4. Add the PDFFREDocumentToImage action to the Convert PDF to TIF function.
  - a. Expand the **Convert PDF > (P) Open** rule and select the **Convert PDF to TIF** function.  
This function already contains an action: rrSet("1", "@P.y\_createLayout")  
This action enables the capability of the PDFFREDocumentToImage action to run recognition and create layout XML file.

- b. In the **Actions library** pane, expand the **Convert > Datacap.Libraries.Convert.PdfFRE** action library and select the **PDFFREDocumentToImage** action.



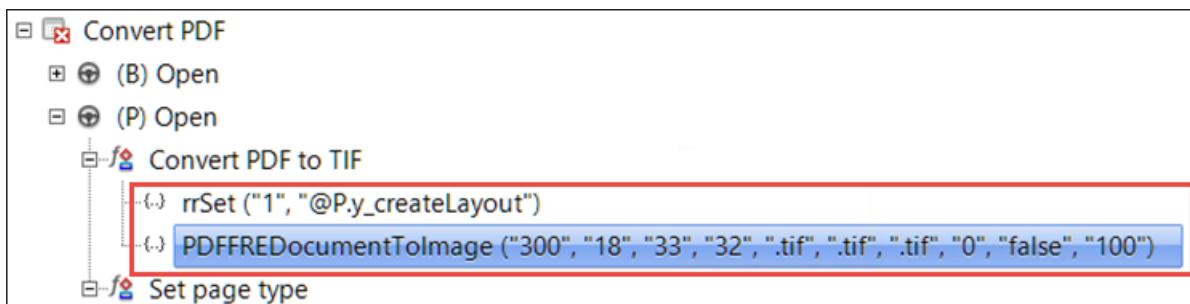
- c. Click the **Add to function** vertical bar.
- 5. Select the **PDFFREDocumentToImage** action and in the **Properties** tab, set the parameters values by using the data in the following table.

Parameter	Value
string resolution	300
string compressionBW	18
string compressionColor	33
string compressionGray	32
string extensionBW	.tif
string extensionColor	.tif
string extensionGray	.tif
string convertMode	0
string useFastBinarization	false
string jpegQuality	100

- \_\_\_ 6. Verify that the completed action contains all the parameter values.

Parameters	
string resolution	300
string compressionBW	18
string compressionColor	33
string compressionGrav	32
string extensionBW	.tif
string extensionColor	.tif
string extensionGray	.tif
string convertMode	0
string useFastBinarization	false
string ipeQuality	100

The PDFFREDocumentToImage action converts a PDF file to TIFF format and creates the layout file. The convertMode parameter value 0 preserves the color. The file extension is: TIF



- \_\_\_ 7. Save and Publish the ruleset.
- \_\_\_ a. Click the **Save changes** icon.
  - \_\_\_ b. Select the **Convert PDF** ruleset, click the lock icon, and click **Publish ruleset**.



### Hint

On the **Rulesets** tab, in the **Solution** section, you can refer to the completed Convert PDF ruleset: **Solution - Convert PDF**

## 1.2. Configure the Analyze Layout ruleset

Use the following table to configure the Analyze Layout ruleset.

If you need step by step details, refer to the instructions after the table.

Item	Value
Ruleset name	AnalyzeLayout
Rule name	(P) Open
Run rule at start of	Pages of type: CoveragePage
Function name	Analyze layout
Actions	Document Analytics > Document Analytics.Actions <ul style="list-style-type: none"> <li>• AnalyzeLayout()</li> </ul> SharedRecognitionTools > SharedRecognitionTools.Actions <ul style="list-style-type: none"> <li>• CreateCcoFromLayout ()</li> <li>• CreateTextFile()</li> </ul> Document Analytics > Document Analytics.Actions <ul style="list-style-type: none"> <li>• CreateHTML()</li> </ul>



### Information

The AnalyzeLayout action enhances and refines the layout XML and improves the labeling and grouping of text blocks that are identified. It requires a previously created layout file.

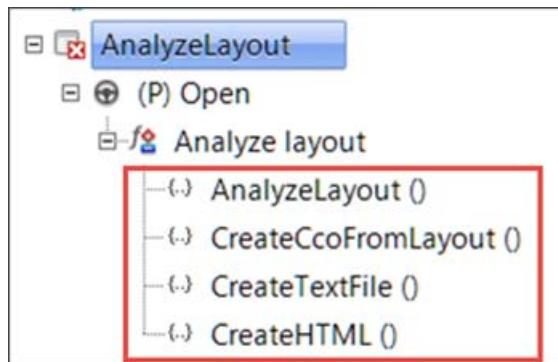
The CreateCcoFromLayout action transfers the characters, confidence, lines, and words from the layout XML to Datacap CCO format and creates a CCO for each page in the batch. The CCO is needed for any locate actions that you use and for click and key functions.

The CreateTextFile action creates a text file for the current page and adds the recognized values to page. The action is helpful for debugging purposes and you can check what recognition values are placed in the CCO file.

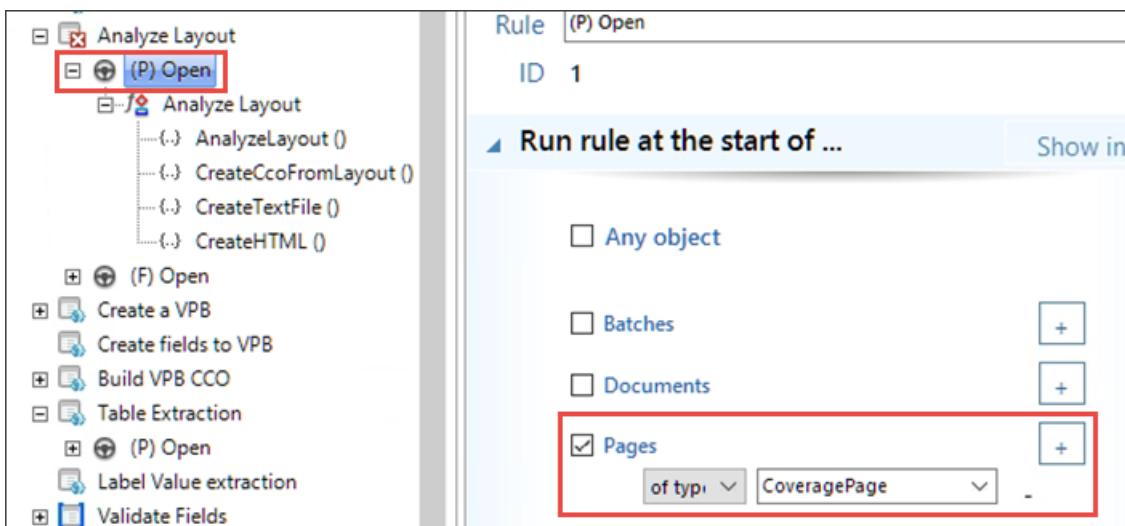
The CreateHTML action creates an HTML document based on the text and format that is captured in the layout XML file. This step is helpful where an HTML document is needed for processing or viewing the document in another product after capture is complete. It is an optional action.

- 
- \_\_ 1. Select **DataExtraction** and click **Add child object** (plus) icon.
  - \_\_ 2. Rename the objects with the following values:
    - **Ruleset1**: Analyze Layout

- **Rule1:** (P) Open
  - **Function1:** Analyze Layout
3. Assign the (P) Open rule to the CoveragePage page.
- a. In the **Rulesets** pane, select the **AnalyzeLayout > (P) Open** rule.
  - b. In the **Properties** tab, expand the **Run rule at start of** node.
  - c. Select **Pages** and click the plus sign next to it.
  - d. For the **of type** field, from the list, select **CoveragePage**.
4. Select the **Analyze Layout** function and add the following actions:
- Document Analytics > Document Analytics.Actions > **AnalyzeLayout**
  - SharedRecognitionTools > SharedRecognitionTools.Actions > **CreateCcoFromLayout**
  - SharedRecognitionTools > SharedRecognitionTools.Actions > **CreateTextFile**
  - Document Analytics > Document Analytics.Actions > **CreateHTML**
5. Verify that the completed **AnalyzeLayout** ruleset contains the actions that you defined.



6. Verify that the **AnalyzeLayout > (P) Open** rule is associated with the **Run rule at the start of > Pages** and **of type > CoveragePage** is selected.



### 1.3. Configure the Populate Layout field in DCO function

The Analyze Layout ruleset is locked for editing from the previous section. In this section, you configure the (F) Open rule for the Analyze Layout ruleset as described in the following table.

If you need step by step details, refer to the instructions after the table.

Item	Value
Ruleset name	AnalyzeLayout
Rule name	(F) Open
Run rule at start of	Fields of type: Layout
Function name	Populate Layout field in DCO
Actions	Document Analytics > Document Analytics.Actions <ul style="list-style-type: none"> <li>• CopyAllBlocks()</li> </ul> ApplicationObjects > Datacap.Libraries ApplicationObjects.Actions <ul style="list-style-type: none"> <li>• SetDCOStatus ("‐1")</li> </ul>



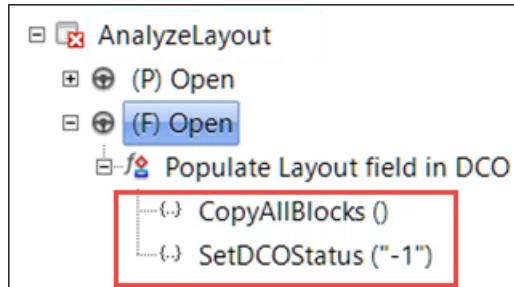
#### Note

The CopyAllBlocks action copies the fields (as subfields below Layout) to the document hierarchy for each block of text. It helps visualize the blocks within Datacap Studio.

The SetDCOStatus action helps to hide the layout fields in user interfaces. The layout fields are not intended to be used from the standard user interfaces.

- 
- \_\_\_ 1. Select the **AnalyzeLayout** ruleset and click **Add child object** (plus) icon.
  - \_\_\_ 2. Rename the following objects:
    - **Rule1:** (F) Open
    - **Function1:** Populate Layout field in DCO
  - \_\_\_ 3. Assign the (F) Open rule to the Layout field.
    - \_\_\_ a. In the **Rulesets** pane, select the **AnalyzeLayout > (F) Open** rule.
    - \_\_\_ b. In the **Properties > Run rule at start of** node, select **Fields** and click the plus sign next to it.
    - \_\_\_ c. For the **of type** field, select **Layout** from the list.
  - \_\_\_ 4. Add the following two actions to the Populate Layout field in DCO function:
    - **CopyAllBlocks()** from the Document Analytics > Document Analytics.Actions action library

- SetDCOStatus ("‐1") from the ApplicationObjects > Datacap.Libraries ApplicationObjects.Actions action library
- 5. Verify that the completed **(F) Open** rule contains the two actions as shown in the screen capture.



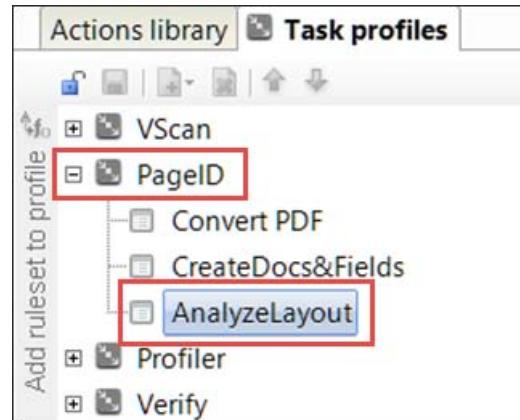
- 6. Verify that the AnalyzeLayout > **(F) Open** rule is associated with the Run rule at the start of > **Fields** and of type > **Layout** is selected.

- 7. Save, order, and publish the ruleset.
  - a. Click **Save changes**.
  - b. Select the **Analyze Layout** ruleset, click the **Move up** icon on the toolbar to place it between **CreateDocs&Fields** and **Create fields to VPB**.
  - c. Select the **AnalyzeLayout** ruleset, click the lock icon, and from the list, click **Publish ruleset**.

## 1.4. Add the AnalyzeLayout ruleset to the PageID task profile

- 1. Lock the **Task profiles** tab by clicking the lock icon.
- 2. Select the **PageID** task profile.
- 3. On the **Rulesets** tab, select the **AnalyzeLayout** ruleset and on the right of the **Rulesets** pane, click the **Add ruleset to profile** vertical bar.

- \_\_\_ 4. On the **Task profiles** tab, expand the **PageID** task profile and verify that the ruleset is listed.

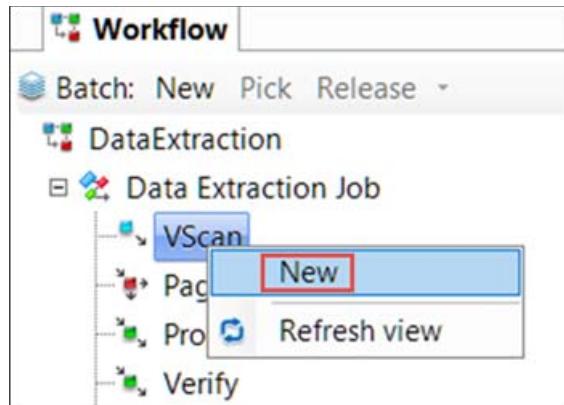


- \_\_\_ 5. On the **Task profiles** tab, click the **Save changes** icon and click the lock icon to unlock it.

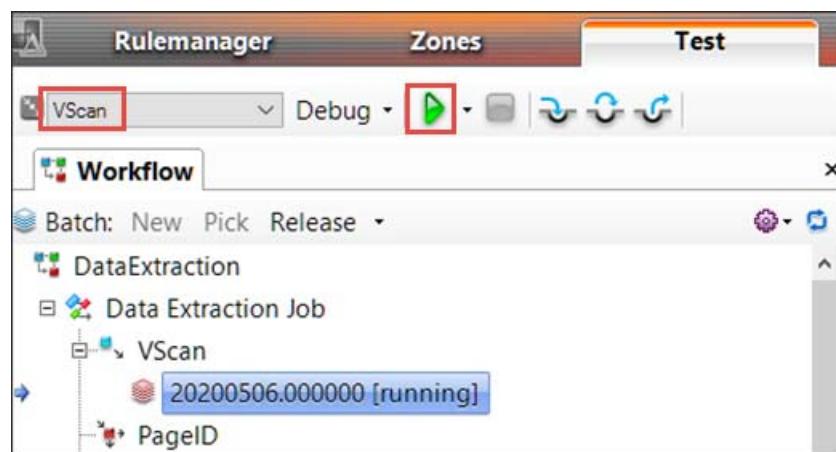
## Section 2. Test the application and verify the layout file

### 2.1. Start a batch to test the application

- \_\_\_ 1. In Datacap Studio, click the **Test** tab from the top toolbar.
- \_\_\_ 2. Start a batch.
  - \_\_\_ a. On the **Workflow** tab, right-click **DataExtraction Job > VScan** and select **New**.

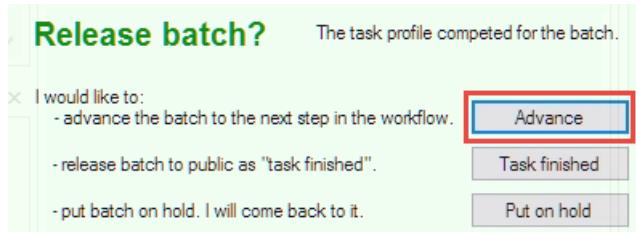


- \_\_\_ b. In the top toolbar, ensure that **VScan** is selected and click the **Process rules for target object** (green arrow icon) next to it.



## 2.2. Run the Page ID task

- \_\_\_ 1. In the **Release batch** dialog box, click **Advance** to move the batch to the **PageID** task.



- \_\_\_ 2. In the **Runtime batch hierarchy tab**, verify that the page is listed as **Other**.
- \_\_\_ 3. In the top toolbar, ensure that **PageID** is selected and click the green arrow icon to run the **PageID** task.  
Wait for the process to complete. It takes a while to complete.
- \_\_\_ 4. Verify the results of running the PageID task.
  - \_\_\_ a. In the **Release batch** dialog box, click **Keep running**.



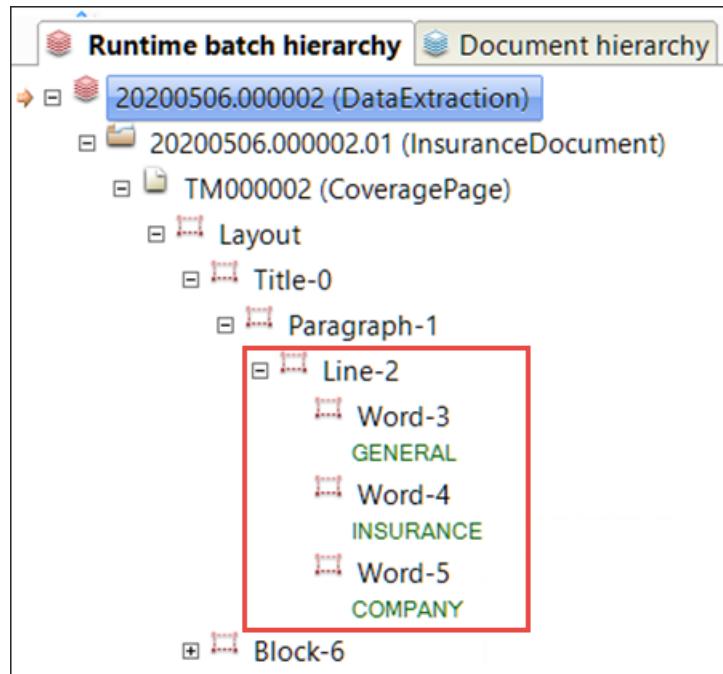
### Information

During the test, if you need to edit the rulesets or modify the task profile configuration, you can complete them and retest the application. Since the batch is still running, you can retest it by running the PageID task again. No need to start the batch again from the VScan step.

- 
- \_\_\_ b. In the **Runtime batch hierarchy tab**, verify that a node has the document name: **InsuranceDocument**
  - \_\_\_ c. Expand **InsuranceDocument** and verify that the page is identified with the name: **CoveragePage**
  - \_\_\_ d. Expand **CoveragePage** and the **Layout** field.

Each block of text from the page layout file is copied to the fields in Layout.

- \_\_\_ e. Expand **Title-0 > Paragraph-1 > Line-2** and verify that the **Word** fields have values that are shown in green text.



- \_\_\_ 5. Compare the text from the Title Block with the original scanned PDF document (`SomeInsurance Company.pdf`) in the `C:\Datacap\DataExtraction\images` folder.



- \_\_\_ 6. Optionally, expand other **Block** nodes and check the values and compare with the original scanned PDF document.

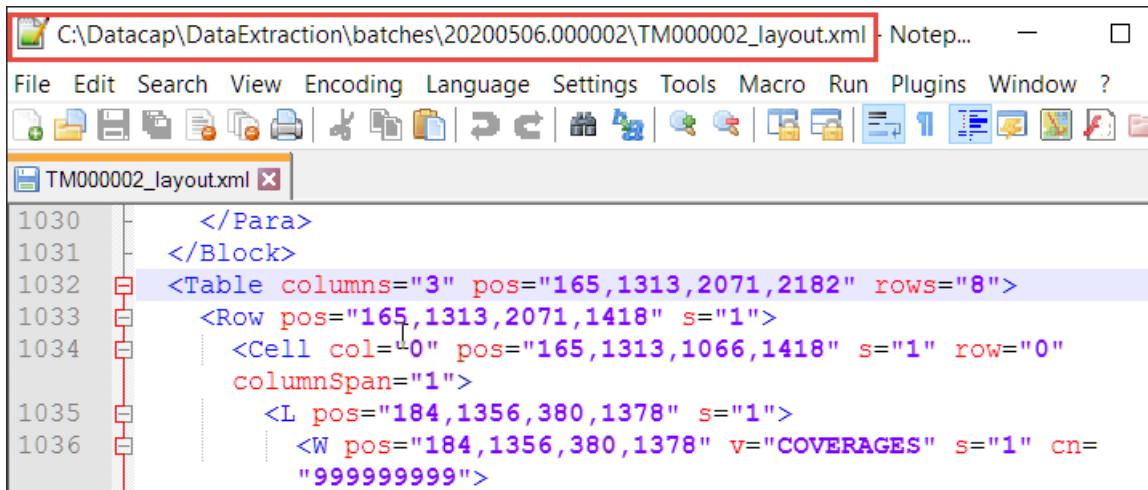
## 2.3. Verify the layout file

After you run the PageID task, the layout XML file is created for the batch in the batches folder.

- \_\_\_ 1. In the **Runtime batch hierarchy** pane, take a note of the batch name: `<Date>.batch number` (Example: 20200506.000002).

You use this value to identify the batch folder in the following steps.

- \_\_\_ 2. Check the layout XML file.
  - \_\_\_ a. In Windows Explorer, browse to the C:\Datacap\DataExtraction\batch folder and open the subfolder with the batch number that you noted down.
  - \_\_\_ b. Verify that an XML file with the term **layout** in the name is created with today's date (For example, TM000002\_layout.xml).
  - \_\_\_ c. Open the file in Notepad++ and check that the file contains the text that is extracted from the scanned document and formatted as different blocks.
  
- \_\_\_ 3. Identify the text in the table.
  - \_\_\_ a. In the layout XML file, search for the term **Table**.
  - \_\_\_ b. Notice that the table data is extracted and saved in this file with **Row** and **Cell** tags.



```

C:\Datacap\DataExtraction\batch\20200506.000002\TM000002_layout.xml - Notep...
File Edit Search View Encoding Language Settings Tools Macro Run Plugins Window ?
TM000002_layout.xml x
1030 </Para>
1031 </Block>
1032 <Table columns="3" pos="165,1313,2071,2182" rows="8">
1033 <Row pos="165,1313,2071,1418" s="1">
1034 <Cell col="0" pos="165,1313,1066,1418" s="1" row="0"
1035 <L pos="184,1356,380,1378" s="1">
1036 <W pos="184,1356,380,1378" v="COVERAGES" s="1" cn=
      "999999999">

```

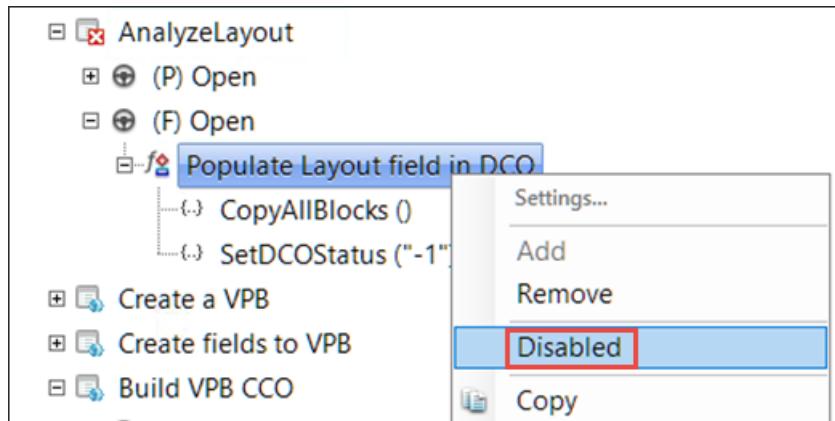
- \_\_\_ 4. Close the file.

## 2.4. Disable the Populate Layout field in DCO function

You used the Populate Layout field in DCO function to populate the layout file data in the batch document hierarchy for testing. When you complete the testing, you disable the function. Disabling prevents large amount of unwanted data that is exported into the export file. You generate the export file in the later exercises.

- \_\_\_ 1. In the Datacap Studio **Test > Workflow** tab, expand **DataExtraction Job > PageID**, right-click the running batch item, and click **Cancel**.
- \_\_\_ 2. Click the **Rulemanager** tab.
- \_\_\_ 3. In the **Rulesets** tab, lock the **AnalyzeLayout** ruleset for editing.
- \_\_\_ 4. Expand the ruleset and expand the **(F) Open** rule.

- \_\_\_ 5. Right-click the **Populate Layout field in DCO** function and select **Disabled** from the list.



- \_\_\_ 6. Verify that the function and actions are disabled (grayed out).  
 \_\_\_ 7. Click the **Save changes**  icon in the toolbar.  
 \_\_\_ 8. Select the **AnalyzeLayout** ruleset, click the lock icon > **Publish ruleset**.  
 \_\_\_ 9. Leave Datacap Studio open for the next exercise.
- 



### Troubleshooting

For this exercise, in the **Rulesets** tab > **Solution** section of the **DataExtraction** application, the completed rulesets are provided.

If the testing didn't work, you can copy the rules, functions, or actions from the solution and paste it to your ruleset for testing.

A solution of the Datacap application for this exercise is also provided in C:\Datacap folder with the application name, **SolutionDataExtraction**. You can access this application like any other Datacap application and test it.

---

## End of exercise

## Exercise review and wrap-up

This exercise showed how to configure the Datacap application to create page layouts.

# Exercise 9. Creating a virtual page block and extracting values from tables

## Estimated time

02:00

## Overview

In this exercise, you configure rulesets to create a virtual page block and extract data from a table, and verify the results.

## Objectives

After completing this exercise, you should be able to:

- Create a virtual page block
- Extract data from a table
- Test the application and verify the extracted data from the table

## Introduction

A virtual page block (VPB) is a new page that is a subset of an existing page or a subset of merged pages. You can run Datacap actions against the new page as you would any other page. A VPB can be used to reduce the amount of data that needs to be processed.

This exercise includes the following sections:

- [Section 1, "Create a Virtual Page Block"](#)
- [Section 2, "Test the VPB Configuration"](#)
- [Section 3, "Configure data extraction from tables"](#)
- [Section 4, "Test the application"](#)

## Requirements

This exercise builds on [Exercise 8, "Creating page layouts"](#). Before starting this exercise, you must complete [Exercise 8, "Creating page layouts"](#).

A solution of the Datacap application for the exercise is provided in C:\Datacap folder with the application name, **SolutionDataExtraction**. You can access this application like any other Datacap application.

## Section 1. Create a Virtual Page Block

In this task, you configure a ruleset to create a virtual page block (VPB).

### 1.1. Check the file for virtual page block

In this section, you check the file that is used for scanning and select the words to define the VPB.

- 1. In Windows Explorer, browse to the C:\Datacap\DataExtraction\images folder and open the General Insurance Company.pdf file.

A VPB is a new page that is created as a subset of an existing page. To create the VPB, you need to identify keywords that indicate the start and end of a section of a page.

- 2. Refer to the PDF file and the following screen capture of the page with the marked area.

GENERAL INSURANCE COMPANY  
VERIFICATION OF COVERAGE  
(SEE BELOW UNDER CAUTIONARY NOTE)

INSURED																										
ROBERT HANSON 1234 Oak St. McKinney, CA, 75070	Policy Number: 902-319-444712 Policy Date: From 04/08/2017 To 10/08/2017 Registered State: TEXAS																									
<p>To whom it may concern: This letter is to verify that we have issued the policyholder coverage under the above policy number for the dates indicated in the effective and expiration date fields for the vehicle listed. This should serve as proof that the below mentioned vehicle meets or exceeds the financial responsibility requirement for your state.</p> <p>This verification of coverage does not amend, extend or alter the coverage afforded by this policy.</p> <p>Vehicle Year: 2015 Make: BMW Model: 528i</p> <p><b>VIN: WBSFH4419F83241810</b></p> <table border="1" style="width: 100%; border-collapse: collapse; text-align: center;"> <thead> <tr> <th>COVERAGES</th> <th>LIMITS</th> <th>DEDUCTIBLES</th> </tr> </thead> <tbody> <tr> <td>BODILY INJURY LIABILITY</td> <td>\$100,000 / \$300,000</td> <td></td> </tr> <tr> <td>PROPERTY DAMAGE LIABILITY</td> <td>\$50,000</td> <td></td> </tr> <tr> <td>UNINSURED &amp; UNDERINSURED MOTORISTS</td> <td>\$30,000 / 360,000</td> <td></td> </tr> <tr> <td>COMPREHENSIVE</td> <td></td> <td>ACV LESS \$1,000</td> </tr> <tr> <td>COLLISION</td> <td></td> <td>\$1000 DED/WAIVER</td> </tr> <tr> <td>EMERGENCY ROAD SERVICE</td> <td>FULL</td> <td>NON-DED</td> </tr> <tr> <td>RENTAL REIMBURSEMENT</td> <td>\$35/DAY - \$1,050 MAX</td> <td></td> </tr> </tbody> </table> <p><b>Lienholder X      Additional Insured X      Interested Party</b></p> <p>BMW FINANCIAL SERVICES NA P.O. BOX 3331 ATTENTION INSURANCE SALT HILLS, MA, 12345</p>			COVERAGES	LIMITS	DEDUCTIBLES	BODILY INJURY LIABILITY	\$100,000 / \$300,000		PROPERTY DAMAGE LIABILITY	\$50,000		UNINSURED & UNDERINSURED MOTORISTS	\$30,000 / 360,000		COMPREHENSIVE		ACV LESS \$1,000	COLLISION		\$1000 DED/WAIVER	EMERGENCY ROAD SERVICE	FULL	NON-DED	RENTAL REIMBURSEMENT	\$35/DAY - \$1,050 MAX	
COVERAGES	LIMITS	DEDUCTIBLES																								
BODILY INJURY LIABILITY	\$100,000 / \$300,000																									
PROPERTY DAMAGE LIABILITY	\$50,000																									
UNINSURED & UNDERINSURED MOTORISTS	\$30,000 / 360,000																									
COMPREHENSIVE		ACV LESS \$1,000																								
COLLISION		\$1000 DED/WAIVER																								
EMERGENCY ROAD SERVICE	FULL	NON-DED																								
RENTAL REIMBURSEMENT	\$35/DAY - \$1,050 MAX																									

To create a VPB that contains the table on the page, you select the following keywords:

- Word to start: **VIN**
- Word to end: **Lienholder**

In the following sections, you use these words and the Locate action library to identify the start and the end of the block.



### Note

Some of the following subsections provide instructions for creating rulesets. If you are already familiar with the steps, you can refer to the data in the table or the screen captures and skip the instructions.

A solution of the Datacap application for this exercise is provided in C:\Datacap folder with the application name, **SolutionDataExtraction**. You can access this application like any other Datacap application.

## 1.2. Start Datacap Studio

- 1. Before you begin the exercise, ensure that the components are running. If you did not start the components already, refer to [Section 1, "Prepare your system: Start the components"](#) to start them.  
If you restart your system, you must start all the components.
- 2. If the Datacap Studio is not already open, log in to Datacap Studio.
  - a. On the Windows desktop, double-click the **Datacap Studio** shortcut.
  - b. In the **Applications** dialog box, select **DataExtraction** and click **Next**.
  - c. On the **TaskMaster Login** page, enter the following values:
    - **User ID:** admin
    - **Password:** admin
    - **Station ID:** 1
  - d. Click **Finish**.

## 1.3. Configure the Create a VPB ruleset

Use the following table to configure the Create a VPB ruleset.

If you need step by step details, refer to the instructions after the table.

Item	Value
Ruleset name	Create a VPB
Rule name	(P) Open
Run rule at start of	Pages of type: CoveragePage

Function name	Create Virtual Page Block
Actions	<p>Locate</p> <ul style="list-style-type: none"> <li>• RegExFind ("VIN")</li> <li>• SetVirtualPageStartPosition ()</li> <li>• RegExFind ("Lienholder")</li> <li>• SetVirtualPageEndPosition ()</li> <li>• CreateVirtualPage ("VPB")</li> </ul>



## Information

The RegExFind action locates the first occurrence of a word or phrase (that you passed as the parameter) on the current page. The search is case-sensitive.

After identifying the start position for the VPB, the SetVirtualPageStartPosition action sets it as the start position.

You add the RegExFind action the second time to locate the first occurrence of a word or phrase (Lienholder) on the current page. This word is used to mark the end position.

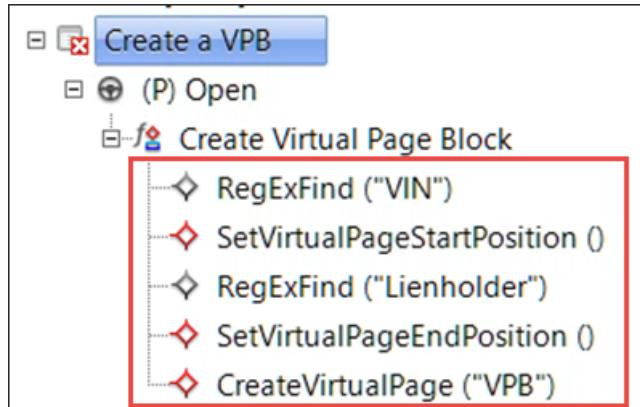
The SetVirtualPageEndPosition action establishes the end of the VPB.

The CreateVirtualPage action creates the virtual page with the start and end that you defined. The name, which you passed in the parameter for the virtual page (VPB), is already defined in the DCO.

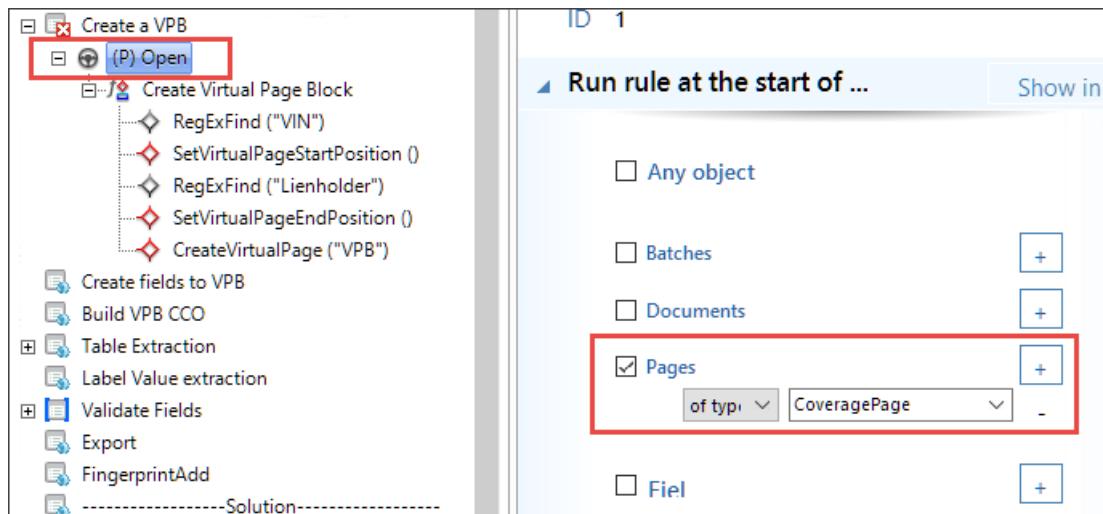
In the previous exercise, you already checked the Document hierarchy tab for the **VPB** page type.

- 
- \_\_\_ 1. In the **Rulemanager** tab, select **DataExtraction** and click **Add child object** (plus) icon.
  - \_\_\_ 2. Select and click each object and rename it with the following values:
    - **Ruleset1**: Create a VPB
    - **Rule1**: (P) Open
    - **Function1**: Create Virtual Page Block
  - \_\_\_ 3. Assign the (P) Open rule to Pages.
    - \_\_\_ a. In the **Rulesets** pane, select the **Create a VPB > (P) Open** rule.
    - \_\_\_ b. In the **Properties > Run rule at start of** node, select **Pages** and click the plus sign next to it.
    - \_\_\_ c. For the **of type** field, select **CoveragePage** from the list.
  - \_\_\_ 4. Select the **Create Virtual Page Block** function, and from the **Locate** action library, add the following actions in the order they are listed, and set the parameters:
    - RegExFind ("VIN")
    - SetVirtualPageStartPosition ()
    - RegExFind ("Lienholder")

- SetVirtualPageEndPosition ()
  - CreateVirtualPage ("VPB")
- \_\_\_ 5. Verify that the completed **Create a VPB** ruleset contains the actions that you defined.



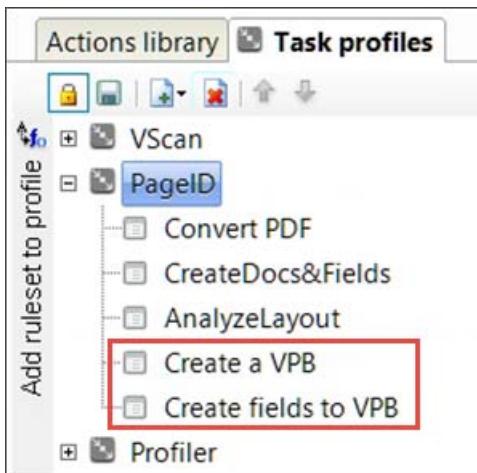
- \_\_\_ 6. Verify that the Create a VPB > **(P) Open** rule is associated with the Run rule at the start of > **Pages** and of type > **CoveragePage** is selected.



- \_\_\_ 7. Save the changes, order, and publish the ruleset.
- \_\_\_ a. In the toolbar, click the **Save changes**  icon.
  - \_\_\_ b. Select the **Create a VPB** ruleset, click the lock icon, and click **Publish ruleset**.
  - \_\_\_ c. Select the **Create a VPB** ruleset, click the **Move up**  icon to place it between the **AnalyzeLayout** and **Create fields to VPB** rulesets.

## 1.4. Add both VPB rulesets to the PageID task profile

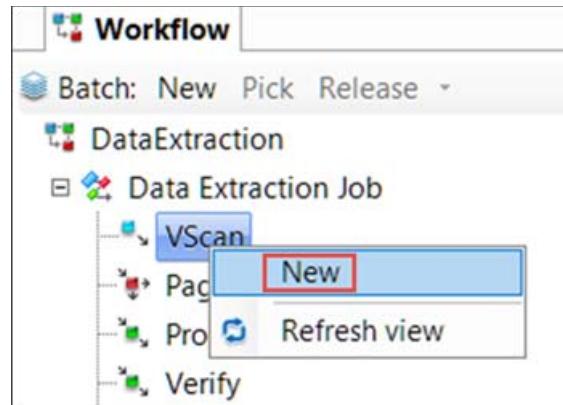
- \_\_\_ 1. On the upper right pane, click the **Task profiles** tab, lock it by clicking the lock  icon.
- \_\_\_ 2. Select the **PageID** task profile.
- \_\_\_ 3. On the **Rulesets** tab, select the **Create a VPB** ruleset and on the right of the **Rulesets** pane, click the **Add ruleset to profile** vertical bar.
- \_\_\_ 4. On the **Task profiles** tab, expand the **PageID** task profile and check that it contains the **Create a VPB** ruleset.
- \_\_\_ 5. Add the **Create fields to VPB** ruleset to the **PageID** task profile.  
The ruleset contains actions to add the VPB page to the document and also create fields for the VPB page.
- \_\_\_ 6. Click the **Save changes**  icon on the **Task profiles** tab and unlock the Task profiles by clicking the lock icon.
- \_\_\_ 7. Verify that the rulesets are added to the task profiles in the order as shown in the following screen capture.



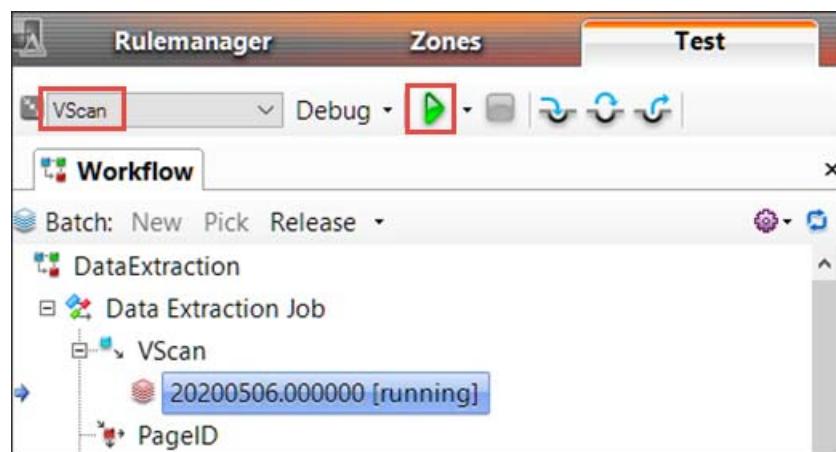
## Section 2. Test the VPB Configuration

### 2.1. Start a batch to test the application

- \_\_\_ 1. In Datacap Studio, click the **Test** tab from the top toolbar.
- \_\_\_ 2. Start a batch.
  - \_\_\_ a. On the **Workflow** tab, expand **DataExtraction Job** right-click **VScan**, and select **New**.

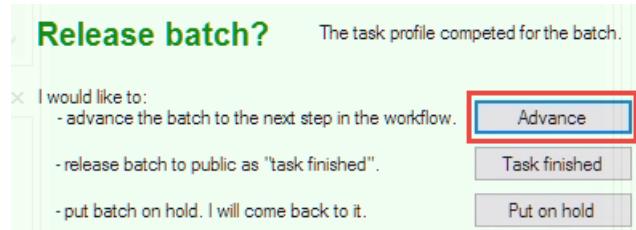


- \_\_\_ b. In the top toolbar, ensure that **VScan** is selected and click the **Process rules for target object** (green arrow icon) next to it.



## 2.2. Run the Page ID task

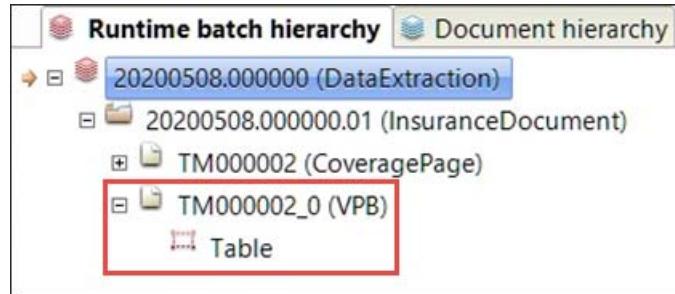
- \_\_\_ 1. In the **Release batch** dialog box, click **Advance** to move the batch to the **PageID** task.



- \_\_\_ 2. In the **Runtime batch hierarchy tab**, verify that the page is listed as **Other**.
- \_\_\_ 3. In the top toolbar, ensure that **PageID** is selected and click the green arrow icon to run the **PageID** task.
- \_\_\_ 4. Verify the results of running the PageID task.
  - \_\_\_ a. In the **Release batch** dialog box, click **Keep running**.
  - \_\_\_ b. In the **Runtime batch hierarchy** tab, verify that the node has the document name: **InsuranceDocument**
  - \_\_\_ c. Expand **InsuranceDocument** and verify that two pages are listed: **CoveragePage** and **VPB**

You created **CoveragePage** in the previous exercise. In this exercise, you defined a virtual page block called **VPB**.
  - \_\_\_ d. Expand **VPB** and notice that it has the **Table** field.

In the next section, you configure data extraction from a table.



- \_\_\_ 5. In the **Runtime batch hierarchy** pane, take a note of the batch name: <Date>. <batch number> (Example: 20200508.000000).  
You use this value to identify the batch folder in the following steps.
- \_\_\_ 6. In the Datacap Studio **Test** tab, on the **Workflow** tab, expand **DataExtraction Job > PageID**, right-click the running batch item and click **Cancel** from the list.

## 2.3. Check the newly created page (VPB)

After you run the PagelD task, a layout XML file is created for the VPB in the batches folder. In this section, you check this file and compare it with the original full page layout file that you created in the previous exercise.

- 1. Open Windows Explorer, navigate to the batch folder:

C:\Datacap\DataExtraction\batches\<batch ID>

The <batch ID> is the value that you noted down in the previous section.

- 2. Open the TM000002\_layout.xml file in Notepad++.

The layout XML file is for the complete page that you scanned.

In the same folder, one more layout XML file similar to this file is created for VPB.

- 3. Search for the word, **VIN**, and notice that it is listed on **Line 1002** for the first time.

The line number on your file could be different.

```

1000      </L>
1001
1002      <L pos="167,1228,626,1254" s="2">
          <W pos="167,1228,238,1254" v="VIN:" s="2" cn="9999">
          <C pos="167,1228,192,1254" v="V" s="2" cn="10" />
          <C pos="192,1228,204,1254" v="I" s="2" cn="10" />
1003
1004

```

- 4. Search for the word, **Table**, and notice that it is listed on **Line 1032** for the first time.

The line number on your file could be different.

```

1031      </Block>
1032      <Table columns="3" pos="165,1313,2071,2182" rows="8">
          <Row pos="165,1313,2071,1418" s="1">
          <Cell col="0" pos="165,1313,1066,1418" s="1" row="0" columnSpan="1">
          <L pos="184,1356,380,1378" s="1">
          <W pos="184,1356,380,1378" v="COVERAGES" s="1" cn="9999999999">
1033
1034
1035
1036

```

You compare the locations of these occurrences with the VPB layout file that is the subset of this page.

- 5. In the Windows Explorer, in your batch folder: C:\Datacap\MyVPBProject\batches\<batch ID>, open the TM000002\_0\_layout.xml file in Notepad++.

The layout XML file is for the virtual page block (VPB).

- 6. Search for the word, **VIN**, and notice that it is listed on **Line 8** for the first time.

The line number on your file could be different.

```

6      <Para pos="167,1228,626,1254">
7
8      <L pos="167,1228,626,1254">
          <W pos="167,1228,238,1254" v="VIN:" s="2" cn="9999">
          <C pos="167,1228,192,1254" v="V" s="2" cn="10" />
9

```

The VPB page begins from the middle of the original **General Insurance Company.pdf** file (CoveragePage). The word, **VIN**, occurs at the beginning of the layout XML file for VPB, unlike the full page that you checked in the previous step.

- \_\_\_ 7. Search for the word, **Table**, and notice that it is listed on **Line 38** for the first time.

The line number on your file could be different.

```
37      </Block>
38      <Table columns="3" pos="165,1313,2071,2182" rows="8">
39          <Row pos="165,1313,2071,1418" s="1">
40              <Cell col="0" pos="165,1313,1066,1418" s="1" row="0" columnSpan="1">
41                  <L pos="184,1356,380,1378" s="1">
42                      <W pos="184,1356,380,1378" v="COVERAGES" s="1" cn="999999999">
```

Similar to the word **VIN**, the word **Table** also occurs earlier in the layout file. It is because the VPB, a subset of the actual page, starts from the middle of the complete page.

**Question:** How do the number of lines in the two layout files (`TM000002_layout.xml` of full page and `TM000002_0_layout.xml` of the VPB) compare?

Answer: The layout file for the VPB is shorter as compared to the layout file for the full page.

The total number of lines in VPB layout file is approximately **500** as compared to the full page that contains approximately **2455**

- \_\_\_ 8. Close the layout XML files.

## Section 3. Configure data extraction from tables

In this section, you configure the rulesets for extracting data from tables.

### 3.1. Configure the Build VPB CCO ruleset

The ruleset creates the CCO file from the layout file for VPB. You need the CCO file to run Locate actions.

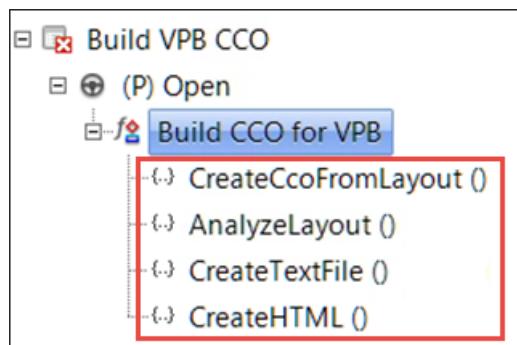
Use the following table to configure the Build VPB CCO ruleset.

If you need step by step details, refer to the instructions after the table.

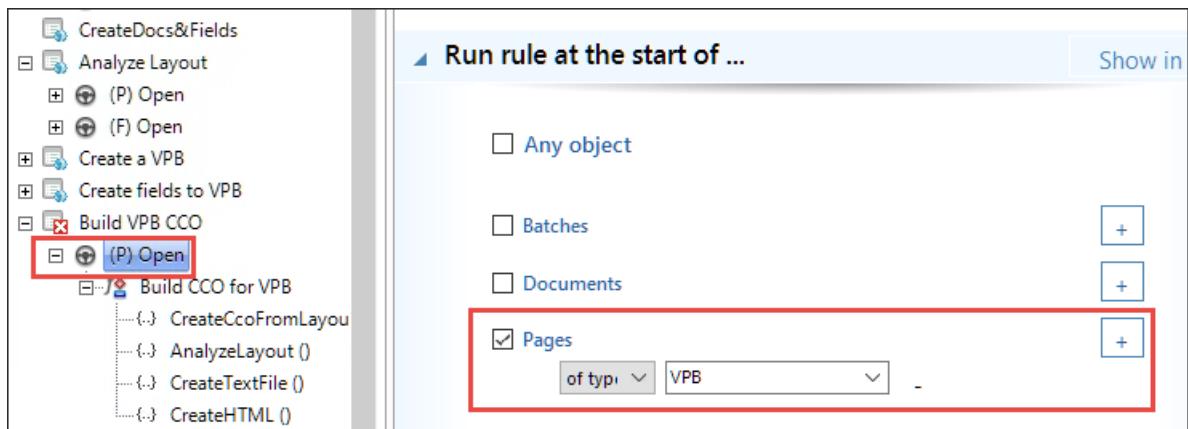
Item	Value
Ruleset name	Build VPB CCO
Rule name	(P) Open
Run rule at start of	Pages of type: VPB
Function name	Build CCO for VPB
Actions	<p>SharedRecognitionTools &gt; SharedRecognitionTools.Actions</p> <ul style="list-style-type: none"> <li>• CreateCcoFromLayout ()</li> </ul> <p>DocumentAnalytics &gt; DocumentAnalytics.Actions</p> <ul style="list-style-type: none"> <li>• AnalyzeLayout()</li> </ul> <p>SharedRecognitionTools &gt; SharedRecognitionTools.Actions</p> <ul style="list-style-type: none"> <li>• CreateTextFile()</li> </ul> <p>DocumentAnalytics &gt; DocumentAnalytics.Action</p> <ul style="list-style-type: none"> <li>• CreateHTML()</li> </ul>

- \_\_\_ 1. In the Datacap Studio, click the **Rulemanager** tab.
- \_\_\_ 2. Select **DataExtraction** and click **Add child object** (plus) icon.
- \_\_\_ 3. Select and click each object and rename it with the following values:
  - **Ruleset1**: Build VPB CCO
  - **Rule1**: (P) Open
  - **Function1**: Build CCO for VPB
- \_\_\_ 4. Assign the (P) Open rule to Pages.
  - \_\_\_ a. In the **Rulesets** pane, select the **Build VPB CCO > (P) Open** rule.
  - \_\_\_ b. In the **Properties** tab > **Run rule at start of** node, select **Pages** and click the plus sign next to it.
  - \_\_\_ c. For the **of type** field, select **VPB** from the list.

- 5. Select the **Build VPB CCO > (P) Open > Build CCO for VPB** function, and add the following four actions:
- **CreateCcoFromLayout ()** from the SharedRecognitionTools > SharedRecognitionTools.Actions action library
  - **AnalyzeLayout ()** from the DocumentAnalytics > DocumentAnalytics.Actions action library
  - **CreateTextFile ()** from the SharedRecognitionTools > SharedRecognitionTools.Actions action library
  - **CreateHTML ()** from the DocumentAnalytics > DocumentAnalytics.Actions action library
- 6. Click the **Save changes**  icon in the toolbar and
- 7. Verify that the completed **Build VPB CCO** ruleset contains the actions that you defined.



- 8. Verify that the **Build VPB CCO > (P) Open** rule is associated with the Run rule at the start of > **Pages** and of type > **VPB** is selected.



- 9. Select the **Build VPB CCO** ruleset, on the toolbar, click the lock icon, and click **Publish ruleset** from the list.
- 10. Select the **Build VPB CCO** ruleset, click the **Move up**  icon to place it between the **Create fields to VPB** and **Validate Fields** rulesets.

## 3.2. Configure the Table Extraction ruleset

Use the following table to configure the Table Extraction ruleset.

If you need step by step details, refer to the instructions after the table.

Item	Value
Ruleset name	Table Extraction
Rule name	(P) Open
Run rule at start of	Pages of type: VPB
Function name	Find table value
Actions	<p>DocumentAnalytics &gt; DocumentAnalytics.Actions</p> <ul style="list-style-type: none"> <li>• FindTableValueRegEx ("COVERAGES", "BODILY INJURY LIABILITY", "LIMITS", "Bodily Injury", "Limits", "Table", "False")</li> <li>• FindTableValueRegEx ("COVERAGES", "PROPERTY DAMAGE LIABILITY", "LIMITS", "Property Damage", "Limits", "Table", "False")</li> <li>• FindTableValueRegEx ("COVERAGES", "UNINSURED &amp; UNDERINSURED MOTORISTS", "LIMITS", "UnInsured", "Limits", "Table", "False")</li> <li>• FindTableValueRegEx ("COVERAGES", "COMPREHENSIVE", "DEDUCTIBLES", "Comprehensive", "Deductibles", "Table", "False")</li> </ul>



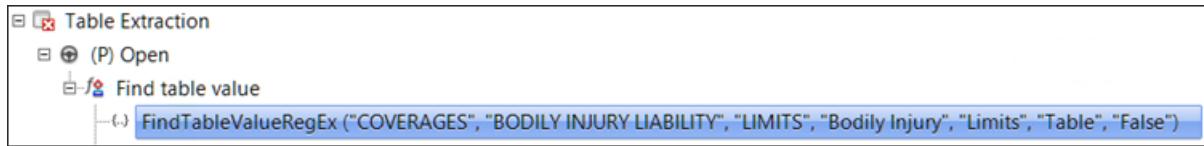
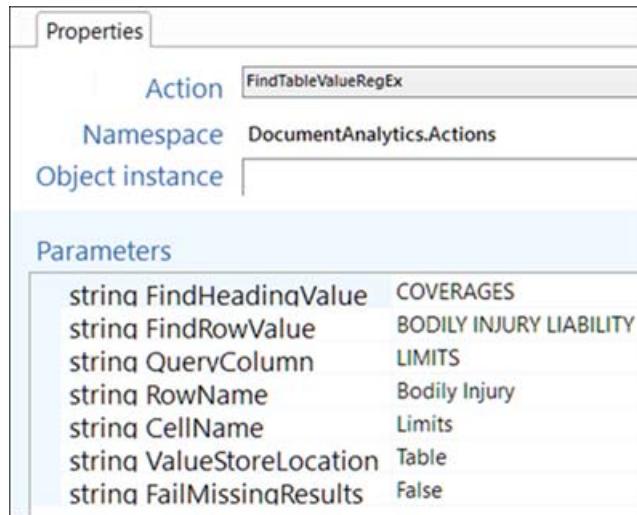
### Important

When you set the parameter values for an action, the system automatically adds quotation marks. Do not add the quotation marks.

After you set the values, ensure that the action on the Ruleset pane have only one set of quotation marks.

- 
- \_\_\_ 1. Select **DataExtraction** and click **Add child object** (plus) icon.
  - \_\_\_ 2. Select and click each object and rename it with the following values:
    - **Ruleset1**: Table Extraction
    - **Rule1**: (P) Open
    - **Function1**: Find table value
  - \_\_\_ 3. Assign the (P) Open rule to the VPB page.
    - \_\_\_ a. In the **Rulesets** pane, select the **Table Extraction > (P) Open** rule.

- \_\_\_ b. In the **Properties** tab > **Run rule at start of** node, select **Pages** and click the plus sign next to it.
- \_\_\_ c. For the **of type** field, from the list, select **VPB**.
- \_\_\_ 4. Add an action to the Find table value function.
  - \_\_\_ a. Select the Table Extraction > (P) Open > **Find table value** function.
  - \_\_\_ b. From the DocumentAnalytics > DocumentAnalytics.Actions action library, add the **FindTableValueRegEx** action to the function.
  - \_\_\_ c. Verify that the action is added to the Find table value function and select the **FindTableValueRegEx** action.
  - \_\_\_ d. In the **Properties** tab, set the values for the following parameters:
    - string **FindHeadingValue**: COVERAGES
    - string **FindRowValue**: BODILY INJURY LIABILITY
    - string **QueryColumn**: LIMITS
    - string **RowName**: Bodily Injury
    - string **CellName**: Limits
    - string **ValueStoreLocation**: Table
    - string **FailMissingResults**: False
  - \_\_\_ e. Verify that the completed action contains all the parameter values.



**Hint**

Compare the action parameter values with the table values in the original scanned PDF document (SomeInsurance Company.pdf) in the C:\Datacap\DataExtraction\images folder.

The table on the PDF file is also shown in the following screen capture.

COVERAGES	LIMITS	DEDUCTIBLES
BODILY INJURY LIABILITY	\$100,000 / \$300,000	
PROPERTY DAMAGE LIABILITY	\$50,000	
UNINSURED & UNDERINSURED MOTORISTS	\$30,000 / 360,000	
COMPREHENSIVE		ACV LESS \$1,000
COLLISION		\$1000 DED/WAIVER
EMERGENCY ROAD SERVICE	FULL	NON-DED
RENTAL REIMBURSEMENT	\$35/DAY - \$1,050 MAX	

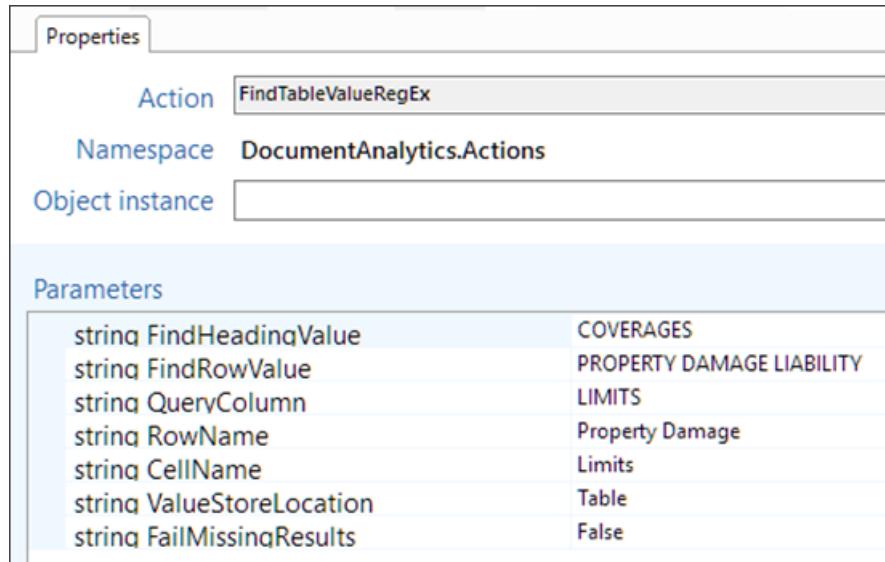
**Note**

What does each parameter value mean?

- The first 3 parameters specify to look for the heading line (**COVERAGES**) in the table to start, the wanted rows (**BODILY INJURY LIABILITY**), wanted column (**LIMITS**), and extract the cell data from the row and column intersections.
- Fourth and fifth parameters specify the field name for the row (**Bodily Injury**) and for the cell (**Limits**).
- The **ValueStoreLocation** parameter specifies the name of the DCO field (**Table**) that contains the extracted data.
- The last parameter controls the return value of the action if the search fails to find a result.

\_\_\_\_\_ f. Click the **Save changes** icon in the toolbar.

- 5. Add the **FindTableValueRegEx** action again three more times with the following set of parameter values to extract data for other rows.
- a. Complete the **FindTableValueRegEx** action second time with the following parameter values:
- string FindHeadingValue: COVERAGES
  - string FindRowValue: PROPERTY DAMAGE LIABILITY
  - string QueryColumn: LIMITS
  - string RowName: Property Damage
  - string CellName: Limits
  - string ValueStoreLocation: Table
  - string FailMissingResults: False
- b. Verify that the completed action contains all the parameter values.



- c. Complete the **FindTableValueRegEx** action third time with the following parameter values:
- string FindHeadingValue: COVERAGES
  - string FindRowValue: UNINSURED & UNDERINSURED MOTORISTS
  - string QueryColumn: LIMITS
  - string RowName: UnInsured
  - string CellName: Limits
  - string ValueStoreLocation: Table
  - string FailMissingResults: False

- \_\_\_ d. Verify that the completed action contains all the parameter values.

Parameters	
string FindHeadingValue	COVERAGES
string FindRowValue	UNINSURED & UNDERINSURED MOTORISTS
string QueryColumn	LIMITS
string RowName	UnInsured
string CellName	Limits
string ValueStoreLocation	Table
string FailMissingResults	False

- \_\_\_ e. Complete the **FindTableValueRegEx** action fourth time with the following parameter values:

Notice that this action uses **DEDUCTIBLES** instead of **LIMITS** column in the table.

- string FindHeadingValue: COVERAGES
- string FindRowValue: COMPREHENSIVE
- string QueryColumn: DEDUCTIBLES
- string RowName: Comprehensive
- string CellName: Deductibles
- string ValueStoreLocation: Table
- string FailMissingResults: False

- \_\_\_ f. Verify that the completed action contains all the parameter values.

Parameters	
string FindHeadingValue	COVERAGES
string FindRowValue	COMPREHENSIVE
string QueryColumn	DEDUCTIBLES
string RowName	Comprehensive
string CellName	Deductibles
string ValueStoreLocation	Table
string FailMissingResults	False

- \_\_\_ 6. Click the **Save changes** icon.
- \_\_\_ 7. Verify that the completed **Table Extraction** ruleset contains the actions and parameters as shown in the following screen capture.



### Hint

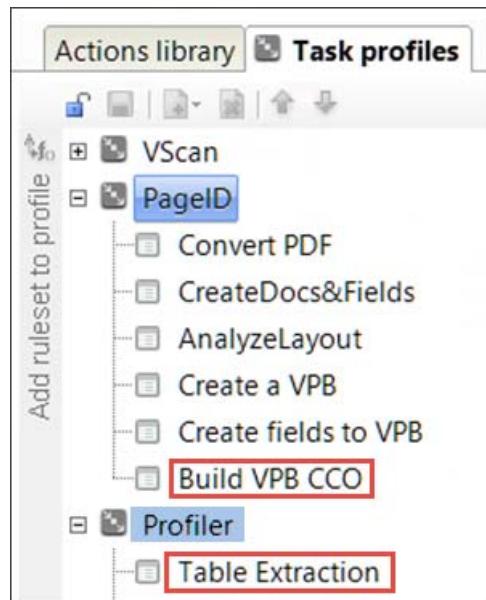
On the **Rulesets** tab, in the **Solution** section, you can refer to the completed Table Extraction ruleset: **Solution - Table Extraction**

- \_\_\_ 8. Verify that the Table Extraction > **(P) Open** rule is associated with the Run rule at the start of > **Pages** and of type > **VPB** is selected.

- \_\_\_ 9. Select the **Table Extraction** ruleset, on the toolbar, click the lock icon, and click **Publish ruleset**.
- \_\_\_ 10. Select the **Table Extraction** ruleset, click the **Move up** icon to place it between **Build VPB CCO** and **Validate Fields** in the list.

### 3.3. Add the Build VPB CCO and Table Extraction ruleset to task profile

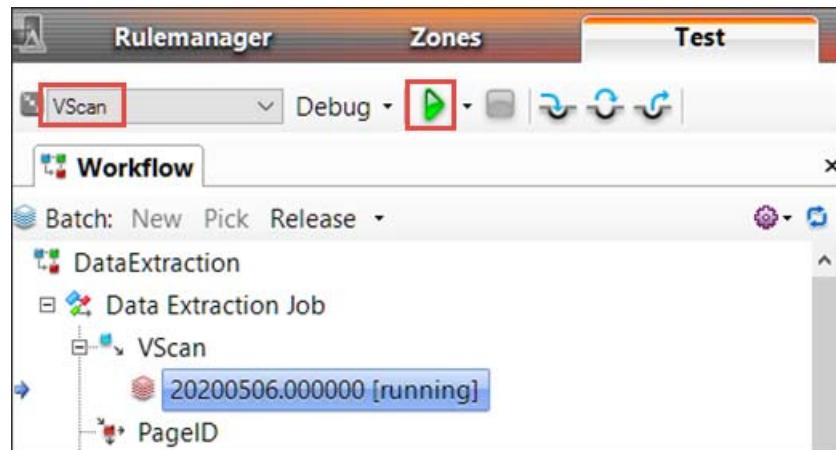
- 1. Click the **Task profiles** tab, lock it by clicking the lock  icon.
- 2. Add Build VPB CCO.
  - a. Select the **PageID** task profile.
  - b. On the **Rulesets** tab, select the **Build VPB CCO** ruleset and click the **Add ruleset to profile** vertical bar.
  - c. On the **Task profiles** tab, expand the **PageID** task profile and verify that it contains the **Build VPB CCO** ruleset.
- 3. Add the **Table Extraction** ruleset to the **Profiler** task profile.
- 4. On the **Task profiles** tab, expand the **Profiler** task profile and verify that it contains the **Table Extraction** ruleset.
- 5. On the **Task profiles** tab, click the **Save changes**  icon and click the lock icon to unlock it.
- 6. Verify that the rulesets are added to the task profiles in the order shown in the following screen capture.



## Section 4. Test the application

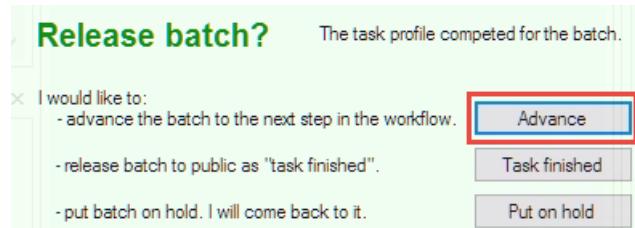
### 4.1. Start a batch to test the application

- \_\_\_ 1. In Datacap Studio, click the **Test** tab.
- \_\_\_ 2. On the **Workflow** tab, expand **DataExtraction Job** right-click **VScan**, and select **New**.
- \_\_\_ 3. In the top toolbar, ensure that **VScan** is selected.
- \_\_\_ 4. If it is not already selected, select it and click the **Process rules for target object** (green arrow icon).



### 4.2. Run the Page ID task

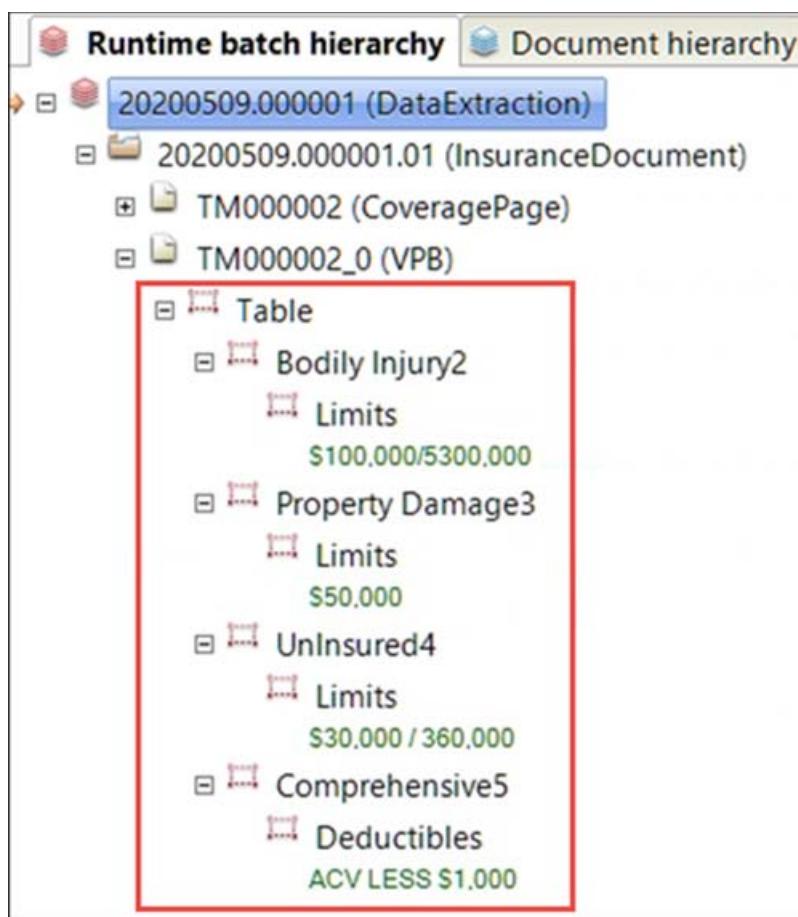
- \_\_\_ 1. In the **Release batch** dialog box, click **Advance** to move the batch to the **PageID** task.



- \_\_\_ 2. In the top toolbar, ensure that **PageID** is selected and click the green arrow icon to run the **PageID** task.  
Wait for the process to complete. It takes a few moments to complete.
- \_\_\_ 3. In the **Release batch** dialog box, click **Advance**.

### 4.3. Run the Profiler task

- 1. In the top toolbar, ensure that **Profiler** is selected and click the green arrow icon to run the **Profiler** task.  
This step extracts the data from the table.
- 2. In the **Release batch** dialog box, click **Keep running**.
- 3. In the **Runtime batch hierarchy tab**, verify that the document type is listed: **InsuranceDocument**
- 4. Expand **InsuranceDocument > VPB** and notice that it has the **Table** field.
- 5. Expand the **Table** node and sub-nodes and verify that the fields have values (shown in green text) that matches the data in the table in the scanned document.

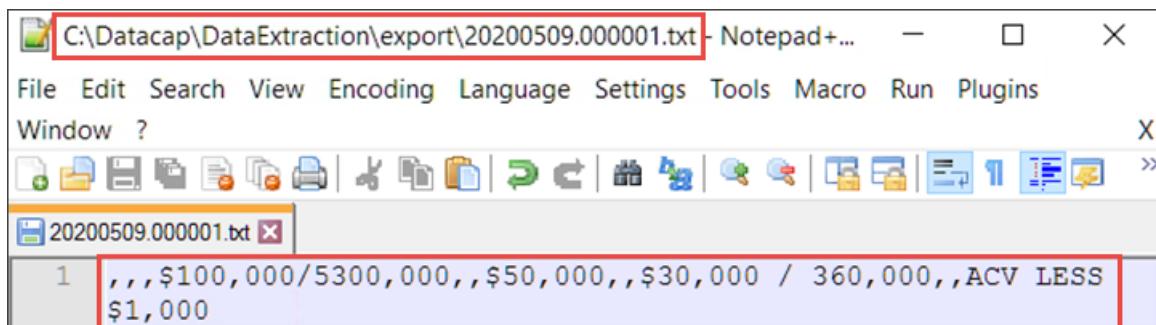


### 4.4. Advance the batch to Verify task and process the batch

- 1. On the **Workflow** tab, in the **DataExtraction Job > Profiler** node, right-click the batch (**running**), and select **Advance**.  
The workflow moves to the Verify step.
- 2. In the top toolbar, ensure that **Verify** is selected and click the green arrow icon to run the **Verify** task.

## 4.5. Process the Export task and verify the export file

- \_\_\_ 1. In the **Release batch** dialog box, click **Advance** to move the batch to the **Export** task.
- \_\_\_ 2. In Datacap Studio **Test** tab, on the **Runtime batch hierarchy** pane, take a note of the batch name: <Date>.<batch number> (Example: 20200509.000001).  
You use this value to identify the export file in the following steps.
- \_\_\_ 3. In the top toolbar, ensure that **Export** is selected and click the green arrow icon to run the task.
- \_\_\_ 4. In the **Release batch** dialog box, click **Advance** to complete the batch.
- \_\_\_ 5. Check the export text file.
  - \_\_\_ a. In Windows Explorer, browse to the C:\Datacap\DataExtraction\export folder and verify that a text file with your batch name (that you noted in step 2) is created with today's date.
  - \_\_\_ b. Open the file in Notepad++ and check that the file contains the comma-separated values for the table cells.



- \_\_\_ c. Compare the values in the file with the values in the table.

COVERAGES	LIMITS	DEDUCTIBLES
BODILY INJURY LIABILITY	\$100,000 / \$300,000	
PROPERTY DAMAGE LIABILITY	\$50,000	
UNINSURED & UNDERINSURED MOTORISTS	\$30,000 / 360,000	
COMPREHENSIVE		ACV LESS \$1,000

- \_\_\_ d. Close the file.
- \_\_\_ 6. Leave Datacap Studio open for the next exercise.



## Troubleshooting

For this exercise, on the **Rulesets** tab, in the **Solution** section, the completed rulesets are provided.

If the testing didn't work, you can copy the rules, functions, or actions from the solution and paste them to your ruleset for troubleshooting.

A solution of the Datacap application for this exercise is also provided in C:\Datacap folder with the application name, **SolutionDataExtraction**. You can access this application like any other Datacap application and refer to it.

---

## End of exercise

## Exercise review and wrap-up

This exercise showed how to configure rulesets to create a Virtual Page Block and extract values from the table.

# Exercise 10.Extracting data from label-value pairs

## Estimated time

01:30

## Overview

In this exercise, you configure rulesets to extract values from label-value pairs, test the application, and verify the results.

## Objectives

After completing this exercise, you should be able to:

- Extract data from label-value pairs
- Test the application and verify the extracted data

## Introduction

The label and value pairs are determined by using font attributes and relative positioning of labels to values to identify what is a label and what is a value.

In this exercise, you configure rulesets to extract data from label-value pairs and test the application.

This exercise includes the following sections:

- [Section 1, "Configure the Find Label Value Pairs function"](#)
- [Section 2, "Test the Find Label Value Pairs function"](#)
- [Section 3, "Configure data extraction from label-value pairs"](#)
- [Section 4, "Test the application"](#)

## Requirements

The exercise builds on [Exercise 8, "Creating page layouts"](#). Before starting this exercise, you must complete the [Exercise 8, "Creating page layouts"](#).

A solution of the Datacap application for the exercise is provided in C:\Datacap folder with the application name, **SolutionDataExtraction**. You can access this application like any other Datacap application.

## Section 1. Configure the Find Label Value Pairs function

In this section, you configure the rulesets for data extraction from the label-value pairs.



### Note

A solution of the Datacap application for this exercise is provided in C:\Datacap folder with the application name: **SolutionDataExtraction**

You can access this application like any other Datacap application.

### 1.1. Check the scanning document for label-value pairs

In this section, you check the document that is used for scanning and identify the label-value pairs.

- 1. Before you begin the exercise, ensure that the components are running. If you did not start the components already, refer to [Section 1, "Prepare your system: Start the components"](#) to start them.  
If you restart your system, you must start all the components.
- 2. In Windows Explorer, browse to the C:\Datacap\DataExtraction\images folder and open the General Insurance Company.pdf file.

<b>GENERAL INSURANCE COMPANY</b> VERIFICATION OF COVERAGE (SEE BELOW UNDER CAUTIONARY NOTE)							
<b>INSURED</b>  ROBERT HANSON 1234 Oak St. McKinney, CA, 75070	Policy Number: 902-319-444712 Policy Date: From 04/08/2017 To 10/08/2017 Registered State: TEXAS						
To whom it may concern: This letter is to verify that we have issued the policyholder coverage under the above policy number for the dates indicated in the effective and expiration date fields for the vehicle listed. This should serve as proof that the below mentioned vehicle meets or exceeds the financial responsibility requirement for your state.							
This verification of coverage does not amend, extend or alter the coverage afforded by this policy.							
Vehicle Year: 2015 Make: BMW Model: 528i VIN: WBSFH4419F83241810							
<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="width: 33%;">COVERAGES</th> <th style="width: 33%;">LIMITS</th> <th style="width: 33%;">DEDUCTIBLES</th> </tr> </thead> <tbody> <tr> <td>BODILY INJURY LIABILITY</td> <td>\$100,000 / \$300,000</td> <td></td> </tr> </tbody> </table>		COVERAGES	LIMITS	DEDUCTIBLES	BODILY INJURY LIABILITY	\$100,000 / \$300,000	
COVERAGES	LIMITS	DEDUCTIBLES					
BODILY INJURY LIABILITY	\$100,000 / \$300,000						

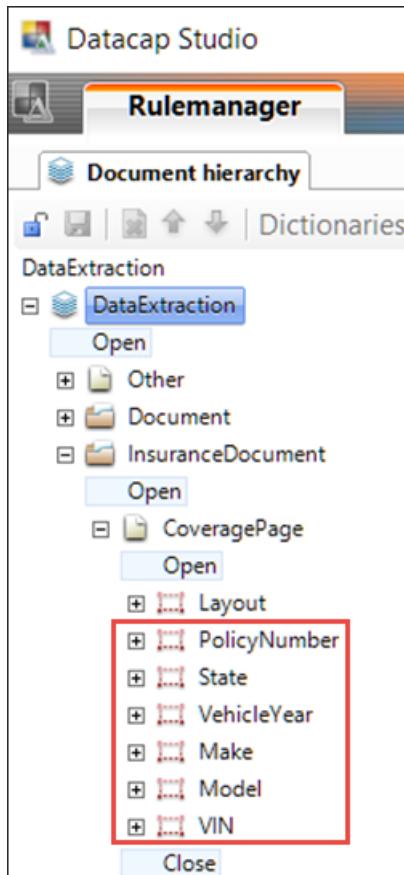
- 3. Refer to the PDF file and the screen capture of the page. The highlighted text shows label-value pairs:

- Upper part of the page: **Policy Number**, **Policy Date**, and **Registered State**
- Lower part of the page: **Vehicle Year**, **Make**, **Model**, and **VIN**



### Note

The fields for the label-value pairs are already defined in the Document hierarchy in Datacap Studio.



## 1.2. Start Datacap Studio

If the Datacap Studio is not already open, log in to Datacap Studio.

- 1. On the Windows desktop, double-click the **Datacap Studio** shortcut.
- 2. In the **Applications** dialog box, select **DataExtraction** and click **Next**.
- 3. On the **TaskMaster Login** page, enter the following values:
  - **User ID:** admin
  - **Password:** admin
  - **Station ID:** 1
- 4. Click **Finish**.

### 1.3. Configure the Label Value extraction ruleset

Use the following table to configure the Label Value extraction ruleset.

If you need step by step details, refer to the instructions after the table..

Item	Value
Ruleset name	Label Value extraction
Rule name	(P) Open
Run rule at start of	Pages of type: CoveragePage
Function name	Find Label Value Pairs
Actions	DocumentAnalytics > DocumentAnalytics.Actions <ul style="list-style-type: none"> <li>• FindLabelValuePairs ()</li> <li>• CopyLabelValuePairs ()</li> </ul>



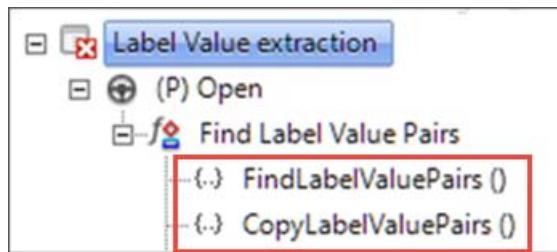
#### Information

The FindLabelValuePairs () action searches the layout XML file for labels and their associated values and returns the results.

The CopyLabelValuePairs () action creates fields in the batch hierarchy for each label-value pair that the FindLabelValuePairs () action finds. The action is helpful during the development to check that the FindLabelValuePairs () action returned values. The action requires a previously created layout XML file.

- 
- \_\_\_ 1. In the **Rulemanager** tab, select **DataExtraction** and click **Add child object** (plus) icon.
  - \_\_\_ 2. Select and click each object and rename it with the following values:
    - **Ruleset1**: Label Value extraction
    - **Rule1**: (P) Open
    - **Function1**: Find Label Value Pairs
  - \_\_\_ 3. Assign the (P) Open rule to Pages.
    - \_\_\_ a. In the **Rulesets** pane, select the **Label Value extraction > (P) Open** rule.
    - \_\_\_ b. In the **Properties** tab > **Run rule at start of** node, select **Pages** and click the plus sign.
    - \_\_\_ c. For the **of type** field, select **CoveragePage** from the list.

- 4. Select the **Find Label Value Pairs** function and from the DocumentAnalytics > DocumentAnalytics.Actions action library, add the following two actions:
- **FindLabelValuePairs ()**  
Notice that a similar action is listed (without an “s”: FindLabelValuePair). Ensure that correct action is selected.
  - **CopyLabelValuePairs ()**
- 5. Verify that the completed **Label Value extraction** ruleset contains the two actions that you defined in the order they are shown in the following screen capture.



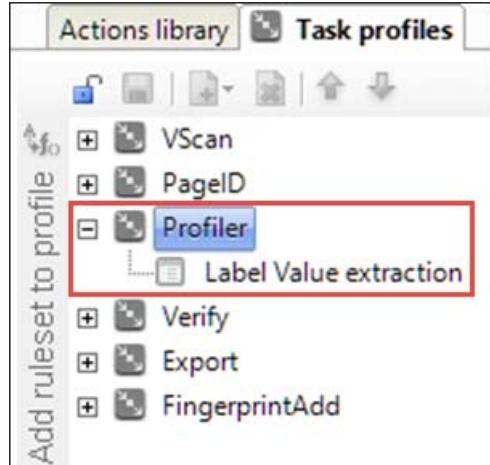
- 6. Verify that the Label Value extraction > **(P) Open** rule is associated with the Run rule at the start of > **Pages** and of type > **CoveragePage** is selected.

- 7. Save, order, and publish the ruleset.
- a. In the toolbar, click the **Save changes** icon.
  - b. Select the **Label Value extraction** ruleset, click the lock icon, and click **Publish ruleset**.
  - c. Select the **Label Value extraction** ruleset, click the **Move up** icon to place it between the **Table Extraction** and **Validate Fields** rulesets.

## 1.4. Add the Label Value extraction ruleset to the Profiler task profile

In this section, from the Profiler task profile, you remove the Table Extraction ruleset that you already tested it in the previous exercise. This step minimizes the process time. You add the Label Value extraction ruleset.

- \_\_\_ 1. Remove the Table Extraction ruleset from the Profiler task profile.
  - \_\_\_ a. Click the **Task profiles** tab, lock it by clicking the lock  icon.
  - \_\_\_ b. Expand the **Profiler** task profile and select the **Table Extraction** ruleset.
  - \_\_\_ c. Click the **Remove** (red x) icon in the toolbar to remove the ruleset.
- \_\_\_ 2. Add the Label Value extraction ruleset to the Profiler task profile.
  - \_\_\_ a. Select the **Profiler** task profile.
  - \_\_\_ b. On the **Rulesets** tab, select the **Label Value extraction** ruleset and on the right of the **Rulesets** pane, click the **Add ruleset to profile** vertical bar.
  - \_\_\_ c. On the **Task profiles** tab, expand the **Profiler** task profile and verify that it contains the **Label Value extraction** ruleset.

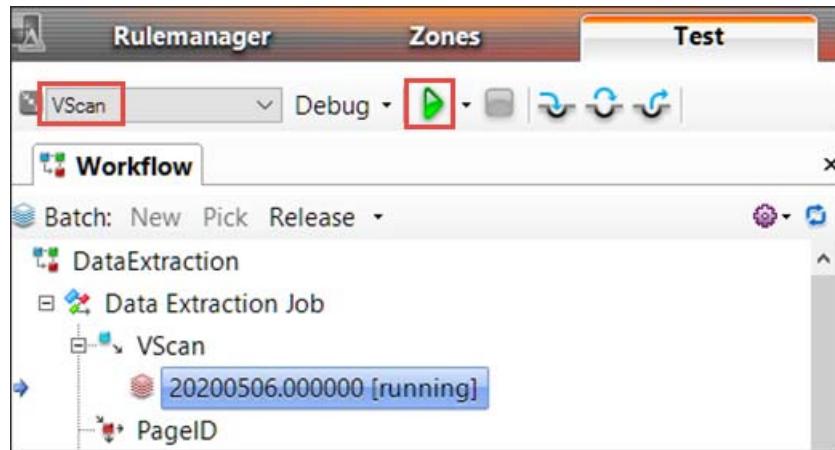


- \_\_\_ 3. Click the **Save changes**  icon on the **Task profiles** tab and unlock the Task profiles by clicking the lock icon.

## Section 2. Test the Find Label Value Pairs function

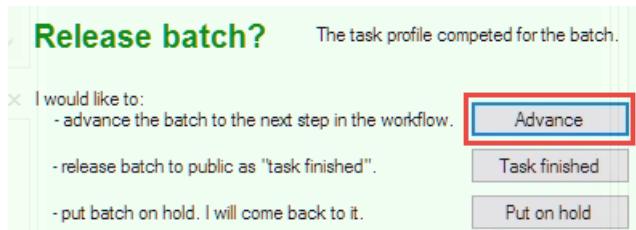
### 2.1. Start a batch to test the application

- \_\_\_ 1. In Datacap Studio, click the **Test** tab.
- \_\_\_ 2. On the **Workflow** tab, expand **DataExtraction Job** right-click **VScan**, and select **New**.
- \_\_\_ 3. In the top toolbar, ensure that **VScan** is selected.
- \_\_\_ 4. Click the **Process rules for target object** (green arrow icon) next to it.



### 2.2. Run the Page ID task

- \_\_\_ 1. In the **Release batch** dialog box, click **Advance** to move the batch to the **PageID** task.



- \_\_\_ 2. In the top toolbar, ensure that **PageID** is selected and click the green arrow icon again to run the **PageID** task.  
Wait for the process to complete. This step takes a few moments.
- \_\_\_ 3. In the **Release batch** dialog box, click **Advance**.

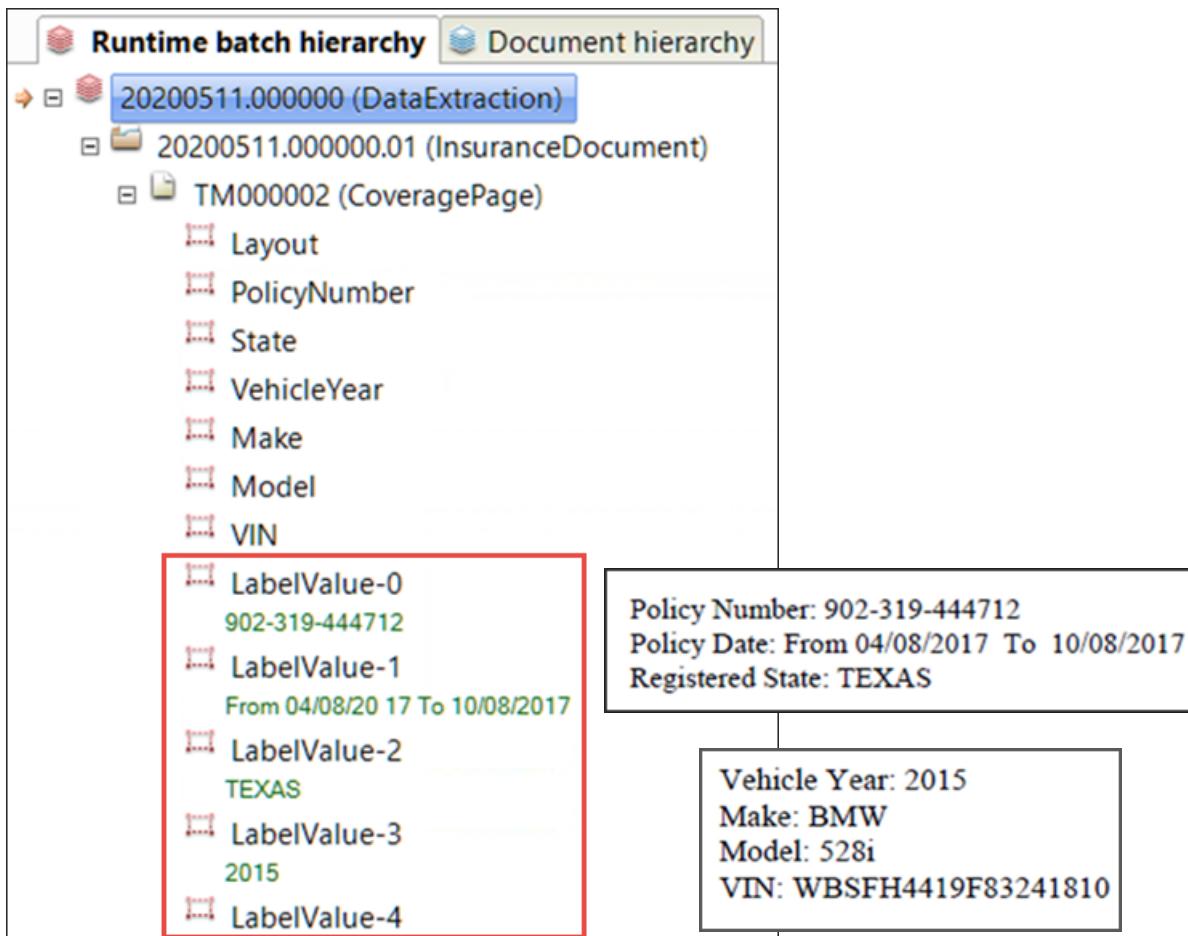
## 2.3. Run the Profiler task

1. In the top toolbar, ensure that **Profiler** is selected and click the green arrow icon to run the **Profiler** task.

This step searches the layout XML file for labels and their associated values and copies the results to the DCO.

The **CopyLabelValuePairs ()** action creates fields in the batch hierarchy for each label-value pair that the **FindLabelValuePairs ()** action finds.

2. In the **Release batch** dialog box, click **Keep running**.
3. In the **Runtime batch hierarchy tab**, expand the batch > **InsuranceDocument** > **CoveragePage** node.
4. Verify that the **LabelValue** fields have values (shown in green text) that matches the data in the scanned document (screen captures on the right).



5. On the **Workflow** tab, expand **DataExtraction Job** > **Profiler**, right-click the running batch item and click **Cancel** from the list.

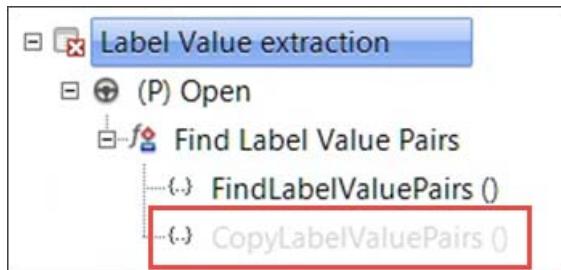
## Section 3. Configure data extraction from label-value pairs

In the previous section, the Datacap actions found all the label-value pairs on the scanned image and populated the data on Document hierarchy. In this section, you configure the rulesets to extract individual label-value pairs and populate specific fields. You also export the values to a text file.

### 3.1. Disable the **CopyLabelValuePairs ()** action

You used the **CopyLabelValuePairs ()** action to populate all the label-value pairs from the layout file data in the batch document hierarchy for testing. After you complete the testing, you disable the action. Disabling prevents large amount of unwanted data that is exported into the export file.

- 1. In Datacap Studio, click the **Rulemanager** tab.
- 1. Click the **Rulesets** pane and select the **Label Value extraction** ruleset.
- 2. Click the lock icon on the toolbar, and click **Lock ruleset for editing** from the list.
- 3. Expand the Label Value extraction > (P) Open > **Find Label Value Pairs** function.
- 4. Right-click the **CopyLabelValuePairs ()** action and select **Disabled** from the list.
- 5. Verify that the action is disabled.



- 6. Click the **Save changes** icon in the toolbar.

### 3.2. Configure the Field level rule for PolicyNumber

Use the following table to configure the (F) Open - PolicyNumber rule.

If you need step by step details, refer to the instructions after the table.

Item	Value
Ruleset name	Label Value extraction
Rule name	(F) Open - Policy Number
Run rule at start of	Fields of type: PolicyNumber like: PolicyNumber
Function name	Extract value for Policy Number
Action	DocumentAnalytics > DocumentAnalytics.Actions • FindLabelValuePair ("Policy Number", "@F")

- \_\_\_ 1. Right-click the **Label Value extraction** ruleset and click **Add Rule** from the list.
- \_\_\_ 2. Rename the following objects:
  - **Rule1:** (F) Open - Policy Number
  - **Function1:** Extract value for Policy Number
- \_\_\_ 3. Assign the (F) Open - Policy Number rule to a Field object.
  - \_\_\_ a. In the **Rulesets** pane, select the **Label Value extraction > (F) Open - Policy Number** rule.
  - \_\_\_ b. In the **Properties** tab > **Run rule at start of** node, select **Fields** and click the plus sign next to it.
  - \_\_\_ c. For the **of type** field, select **PolicyNumber** from the list.
  - \_\_\_ d. Click the plus sign next to **Fields** again.
  - \_\_\_ e. For the first field, from the list, select **like** and for the second field, select **PolicyNumber**.



- \_\_\_ 4. Add action to the Extract value for Policy Number function.
  - \_\_\_ a. Select the (F) Open - Policy Number > **Extract value for Policy Number** function.
  - \_\_\_ b. From the **Document Analytics > Document Analytics.Actions** action library, add the **FindLabelValuePair** action.
 

Notice that you used a similar action in the previous section (with an "s": **FindLabelValuePairs**). Ensure the correct action is selected.

The action populates a field using a value that is found by using the **FindLabelValuePairs** action.
  - \_\_\_ c. Select the action and in the **Properties** tab, set the values for the following parameters:
    - string Regex: Policy Number
    - string Target: @F

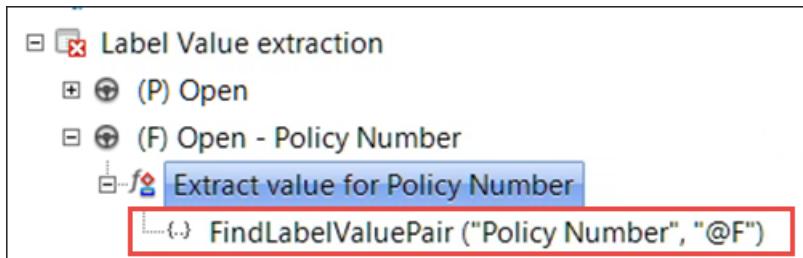


### Hint

The first parameter Policy Number is one of the labels in the original scanned PDF document (**SomeInsurance Company.pdf**) in the **C:\Datacap\DataExtraction\images** folder.

The second parameter identifies the target object (field) to save the value.

- 5. Verify that the completed **(F) Open - Policy Number** rule contains the function and action as shown in the following screen capture.



- 6. Verify that the **(F) Open - Policy Number** rule is associated with the Run rule at the start of > **Fields**. The of type > **PolicyNumber** and like > **PolicyNumber** values are selected.

- 7. Click the **Save changes** icon in the toolbar.

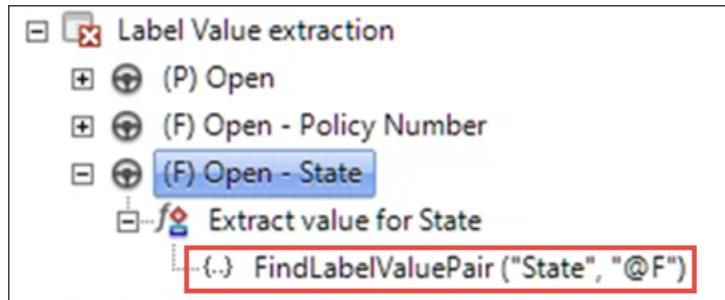
### 3.3. Configure the Field level rule for State

For step-by-step instructions, refer to the previous section (3.2).

- 1. Configure a rule to extract the value for the **Registered State** label on the scanned PDF document by using the data in the following table.

Item	Value
Ruleset name	Label Value extraction
Rule name	(F) Open - State
Run rule at start of	Fields of type: State like: State
Function name	Extract value for State
Action	DocumentAnalytics > DocumentAnalytics.Actions • FindLabelValuePair ("State", "@F")

- 2. Verify that the completed **(F) Open - State** rule contains the function and action as shown in the following screen capture.



- 3. Verify that the **(F) Open - State** rule is associated with the Run rule at the start of > **Fields**. The of type > **State** and like > **State** are selected.

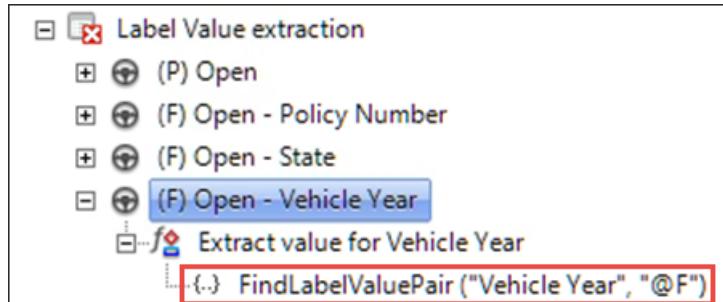
### 3.4. Configure the Field level rule for Vehicle Year

For step-by-step instructions, refer to the previous section (3.2).

- 1. Configure a rule to extract the value for the **Vehicle Year** label on the scanned PDF document by using the data in the following table.

Item	Value
Ruleset name	Label Value extraction
Rule name	(F) Open - Vehicle Year
Run rule at start of	Fields of type: VehicleYear like: VehicleYear
Function name	Extract value for Vehicle Year
Action	DocumentAnalytics > DocumentAnalytics.Actions • FindLabelValuePair ("VehicleYear", "@F")

- 2. Verify that the completed **(F) Open - Vehicle Year** rule contains the function and action as shown in the following screen capture.



- 3. Verify that the **(F) Open - Vehicle Year** rule is associated with the Run rule at the start of > **Fields**. The of type > **VehicleYear** and like > **VehicleYear** are selected.

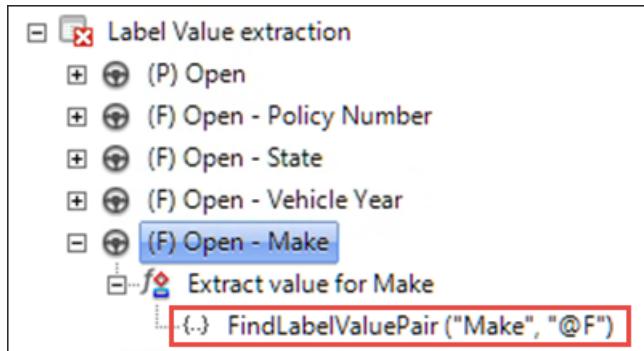
### 3.5. Configure the Field level rule for Make

For step-by-step instructions, refer to the previous section (3.2).

- 1. Configure a rule to extract the value for the **Make** label on the scanned PDF document by using the data in the following table.

Item	Value
Ruleset name	Label Value extraction
Rule name	(F) Open - Make
Run rule at start of	Fields of type: Make like: Make
Function name	Extract value for Make
Action	DocumentAnalytics > DocumentAnalytics.Actions • FindLabelValuePair ("Make", "@F")

- 2. Verify that the completed **(F) Open - Make** rule contains the function and action as shown in the following screen capture.



- 3. Verify that the **(F) Open - Make** rule is associated with the Run rule at the start of > **Fields**. The of type > **Make** and like > **Make** are selected.

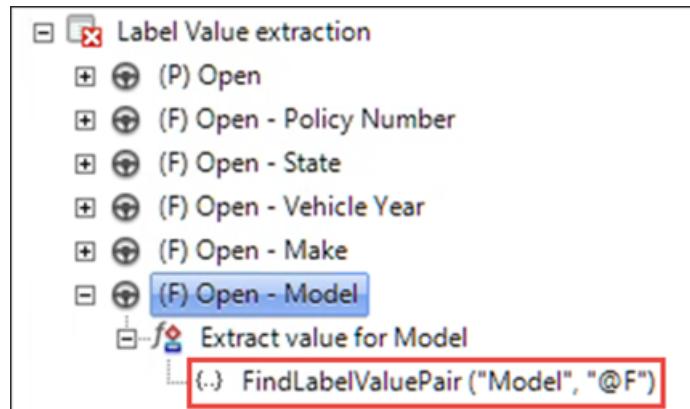
The screenshot shows the 'Run rule at the start of...' configuration. On the left, there is a list of rules, including '(F) Open - Make', which is highlighted with a red box. On the right, under 'Run rule at the start of...', there is a section for 'Fields' with two dropdown menus: 'of type' set to 'Make' and 'like' set to 'Make'. Both of these dropdowns are also highlighted with a red box.

### 3.6. Configure the Field level rule for Model

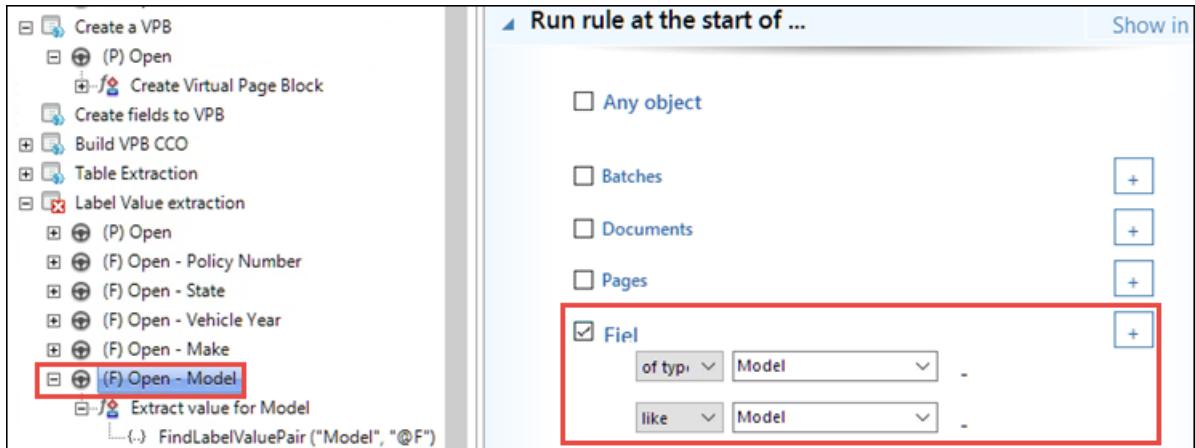
- 1. Configure a rule to extract the value for the **Model** label on the scanned PDF document by using the following table.

Item	Value
Ruleset name	Label Value extraction
Rule name	(F) Open - Model
Run rule at start of	Fields of type: Model like: Model
Function name	Extract value for Model
Action	DocumentAnalytics > DocumentAnalytics.Actions • FindLabelValuePair ("Model", "@F")

- 2. Verify that the completed **(F) Open - Model** rule contains the function and action as shown in the following screen capture.



- 3. Verify that the **(F) Open - Model** rule is associated with the Run rule at the start of > **Fields**. The of type > **Model** and like > **Model** are selected.



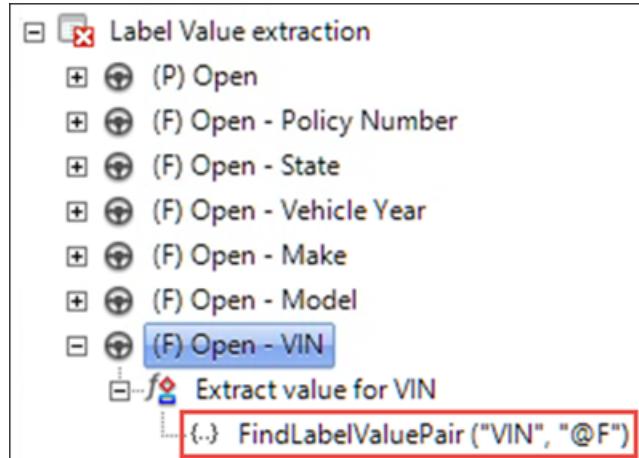
### 3.7. Configure the Field level rule for VIN

- 1. Configure a rule to extract the value for the **VIN** label on the scanned PDF document by using the following table.

Item	Value
Ruleset name	Label Value extraction
Rule name	(F) Open - VIN
Run rule at start of	Fields of type: VIN like: VIN

Function name	Extract value for VIN
Action	<p>DocumentAnalytics &gt; DocumentAnalytics.Actions</p> <ul style="list-style-type: none"> <li>• FindLabelValuePair ("VIN", "@F")</li> </ul>

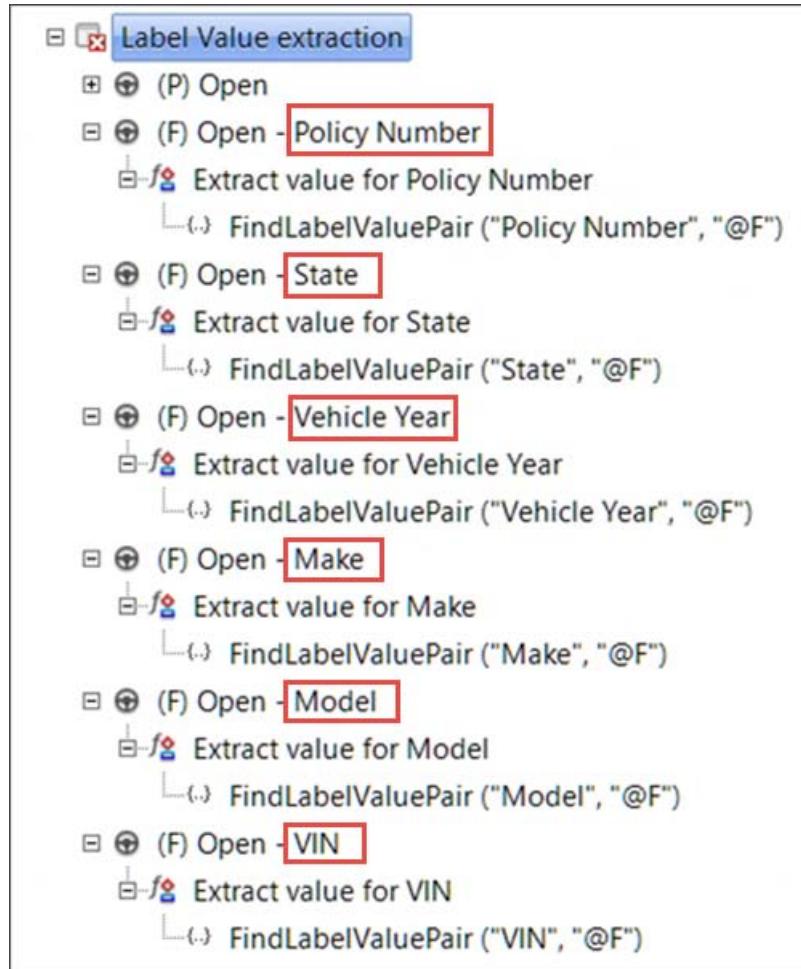
2. Verify that the completed **(F) Open - VIN** rule contains the function and action as shown in the following screen capture.



3. Verify that the **(F) Open - VIN** rule is associated with the Run rule at the start of > **Fields**. The of type > **VIN** and like > **VIN** are selected.

Run rule at the start of ...
<input type="checkbox"/> Any object <input type="checkbox"/> Batches <input type="checkbox"/> Documents <input type="checkbox"/> Pages <input checked="" type="checkbox"/> Fields of type: VIN like: VIN

The screen capture shows a list of all the rules for the Label Value extraction ruleset.



### Hint

On the **Rulesets** tab, in the **Solution** section, you can refer to the completed Label Value extraction ruleset: **Solution - Label Value extraction**

- 
- \_\_\_ 4. In the toolbar, click the **Save changes** icon.
  - \_\_\_ 5. Select the **Label Value extraction** ruleset, click the lock icon and click **Publish ruleset** from the list.

## Section 4. Test the application

In this section, you repeat the testing that you did in Section 2 and continue to process the batch to the Export step.

### 4.1. Run the VScan and Page ID tasks

- \_\_\_ 1. In Datacap Studio, click the **Test** tab.
  - \_\_\_ 2. Run the VScan task.
    - \_\_\_ a. On the **Workflow** tab, expand **DataExtraction Job** right-click **VScan**, and select **New**.
    - \_\_\_ b. In the top toolbar, ensure that **VScan** is selected.
    - \_\_\_ c. Click the **Process rules for target object** (green arrow icon).
  - \_\_\_ 3. Run the Page ID task.
    - \_\_\_ a. In the **Release batch** dialog box, click **Advance** to move the batch to the **PageID** task.
    - \_\_\_ b. In the top toolbar, ensure that **PageID** is selected and click the green arrow icon again to run the **PageID** task.
- Wait for the process to complete. It takes a few moments to complete.
- \_\_\_ 4. In the **Release batch** dialog box, click **Advance**.

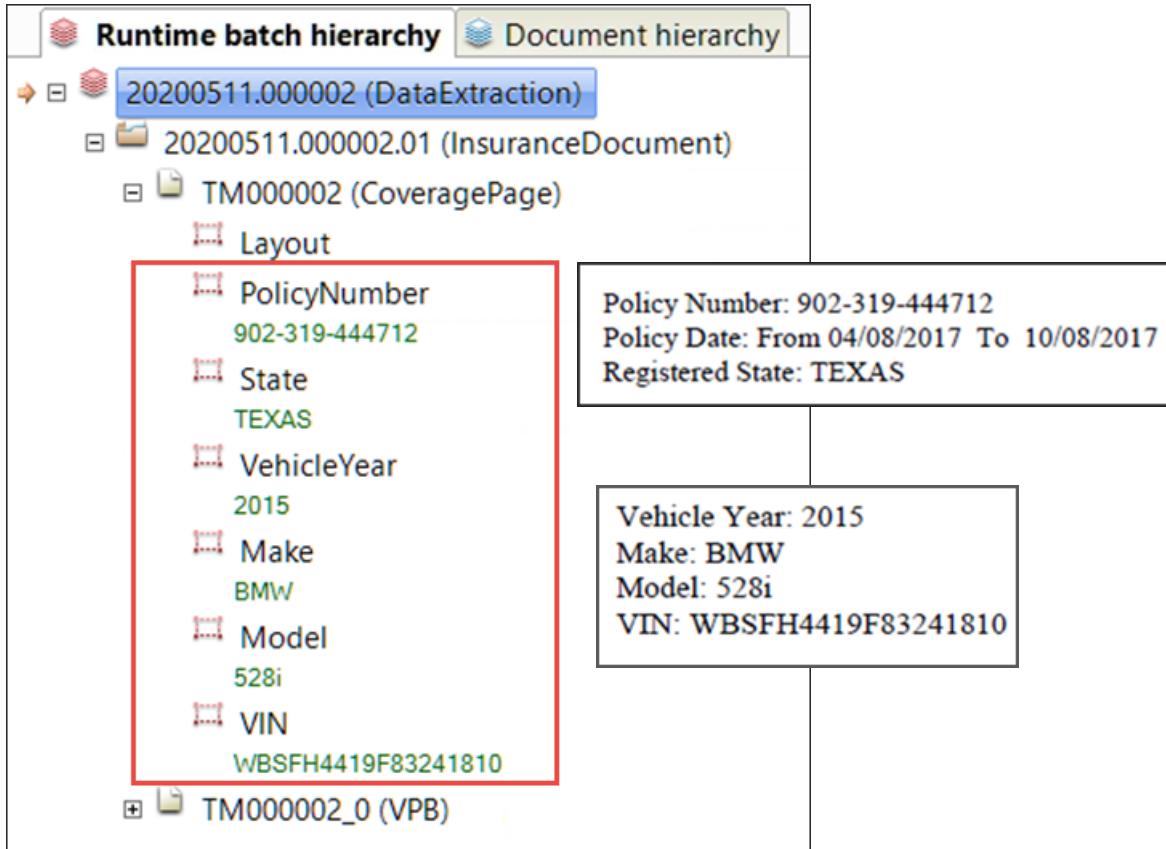
### 4.2. Run the Profiler task

- \_\_\_ 1. In the top toolbar, ensure that **Profiler** is selected and click the green arrow icon to run the **Profiler** task.

The step extracts the data from the label-value pairs and populates the fields.
- \_\_\_ 2. In the **Release batch** dialog box, click **Keep running**.
- \_\_\_ 3. In the **Runtime batch hierarchy tab**, expand the batch > **InsuranceDocument** > **CoveragePage** node.

- 4. Verify that the **LabelValue** fields have values (shown in green text) that matches the data in the scanned document (screen captures on the right).

Labels: **Policy Number**, **State**, **VehicleYear**, **Make**, **Model**, and **VIN**



In the following sections, you extract the data and export it to a file.

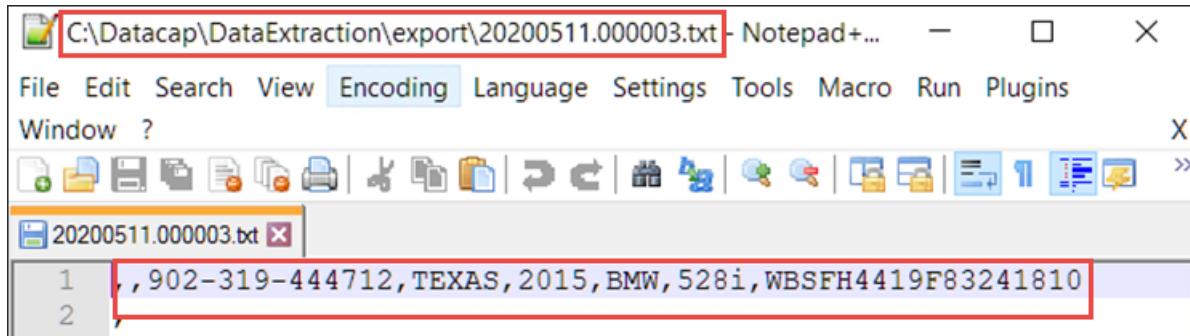
#### 4.3. Advance the batch to Verify task and process the batch

- 1. On the **Workflow** tab, in the **DataExtraction Job > Profiler** node, right-click the batch (**running**), and select **Advance**.  
The workflow moves to the Verify step.
- 2. In the top toolbar, ensure that **Verify** is selected and click the green arrow icon to run the **Verify** task.

#### 4.4. Process the Export task and verify the export file

- 1. In the **Release batch** dialog box, click **Advance** to move the batch to the **Export** task.
- 2. On the **Runtime batch hierarchy** pane, take a note of the batch name: <Date>.<batch number> (Example: 20200511.000002).  
You use this value to identify the export file in the following steps.
- 3. In the top toolbar, ensure that **Export** is selected and click the green arrow icon to run the task.

- \_\_\_ 4. In the **Release batch** dialog box, click **Advance** to complete the batch.
- \_\_\_ 5. Check the export text file.
  - \_\_\_ a. In Windows Explorer, browse to the C:\Datacap\DataExtraction\export folder and verify that a text file with your batch name (that you noted in step 2) is created with today's date.
  - \_\_\_ b. Open the file in Notepad++ and check that the file contains the comma-separated values for the label-value pairs.



- \_\_\_ c. Compare the values in the export file with the values in the scanned document.

Policy Number: 902-319-444712 Policy Date: From 04/08/2017 To 10/08/2017 Registered State: TEXAS
Vehicle Year: 2015 Make: BMW Model: 528i VIN: WBSFH4419F83241810

- \_\_\_ d. Close the file.
- \_\_\_ 6. Click **Exit** at the upper right corner to close Datacap Studio.



### Troubleshooting

On the **Rulesets** tab, in the **Solution** section, you can refer to the completed Label Value extraction ruleset: **Solution - Label Value extraction**

A solution of the Datacap application for this exercise is also provided in C:\Datacap folder with the application name, **SolutionDataExtraction**. You can access this application like any other Datacap application and refer to it.

---

## End of exercise

## Exercise review and wrap-up

This exercise showed how to configure rulesets to extract values from the label-value pairs.



IBM Training



© Copyright International Business Machines Corporation 2020.