



Extract data from document using PDF library

Package version: 1.1.0

Version	Date	Description
1	May 27, 2021	Document created

TABLE OF CONTENTS

INTRODUCTION	4
IMPORTANT RECOMMANDATION	5
General	5
Reuse the sample as a new project	5
DESCRIPTION	7
Settings	7
<i>Environment variables</i>	7
<i>Dependent packages</i>	7
Captures	7
Datatypes	7
<i>SaleOrderDetail</i>	7
Automations.....	7
<i>Extract Data from PDF</i>	7
Input/Output.....	7
Description.....	7
Merge PDF documents	8
Input/Output.....	8
Description.....	8
Split PDF document	8
Input/Output.....	8
Description.....	8
VERSION	9
SAP Build Process Automation	9
Target application	9
PREREQUISITES.....	10
Global setup	10
Specific steps to follow before launching the agent	10
EXPECTED OUTPUT	11

INTRODUCTION

This document describes the SAP Build Process Automation sample **How to use PDF module** and provides the following information:

- Description (functional and technical)
- Version used to generate this sample

It also contains information on prerequisites, such as the steps to follow before launching the agent.

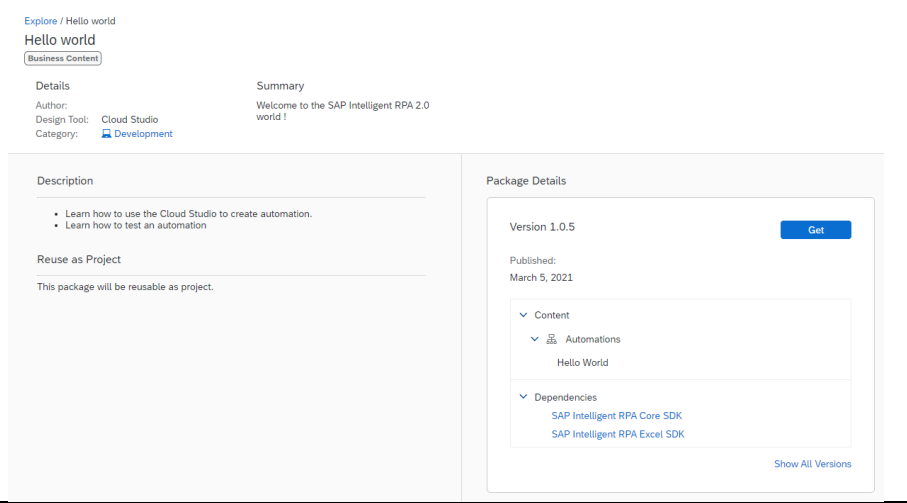
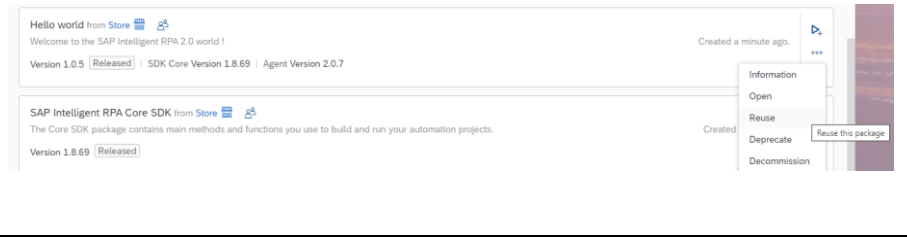
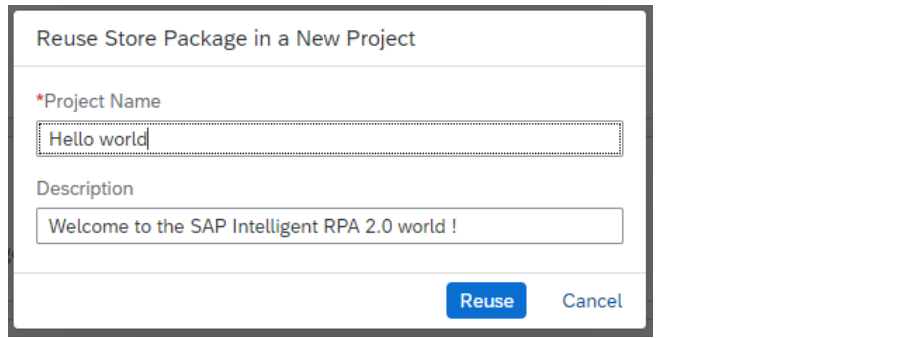
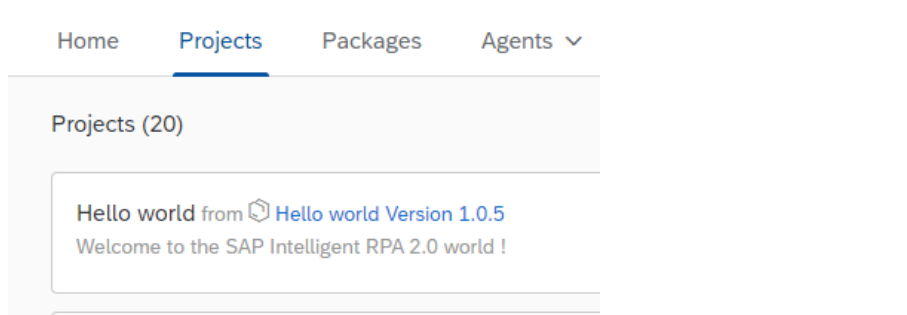
IMPORTANT RECOMMENDATION

General

To use this sample, you need to have a basic knowledge and understanding of SAP Build Process Automation tool. At the very least you need to know how to build an automation, add and modify activities and generate a package.

Reuse the sample as a new project

Note: screenshot might display a different name than the one of this sample.

<p>From the Cloud Factory, open the Store tab and select the sample you want to retrieve.</p> <p>Click on the Get button.</p>	 <p>Explore / Hello world Hello world (Business Content)</p> <p>Details Author: Cloud Studio Design Tool: Cloud Studio Category: Development</p> <p>Summary Welcome to the SAP Intelligent RPA 2.0 world !</p> <p>Description</p> <ul style="list-style-type: none"> Learn how to use the Cloud Studio to create automation. Learn how to test an automation <p>Reuse as Project This package will be reusable as project.</p> <p>Package Details</p> <p>Version 1.0.5 Get</p> <p>Published: March 5, 2021</p> <p>Content</p> <ul style="list-style-type: none"> Automations <ul style="list-style-type: none"> Hello World <p>Dependencies</p> <ul style="list-style-type: none"> SAP Intelligent RPA Core SDK SAP Intelligent RPA Excel SDK <p>Show All Versions</p>
<p>Once the package is retrieved, open the Packages tab of the Cloud Factory.</p> <p>Click on the Options button of the package you just retrieved and select the option Reuse.</p>	 <p>Hello world from Store Welcome to the SAP Intelligent RPA 2.0 world ! Version 1.0.5 (Released) SDK Core Version 1.8.69 Agent Version 2.0.7 Created a minute ago.</p> <p>SAP Intelligent RPA Core SDK from Store The Core SDK package contains main methods and functions you use to build and run your automation projects. Version 1.8.69 (Released) Created</p> <p>Information Open Reuse Deprecate Decommission</p> <p>Reuse this package</p>
<p>Set a name for the project to be created.</p>	 <p>Reuse Store Package in a New Project</p> <p>*Project Name Hello world</p> <p>Description Welcome to the SAP Intelligent RPA 2.0 world !</p> <p>Reuse Cancel</p>
<p>Open the project that has just been created.</p>	 <p>Home Projects Packages Agents</p> <p>Projects (20)</p> <p>Hello world from Hello world Version 1.0.5 Welcome to the SAP Intelligent RPA 2.0 world !</p>
<p>If needed, update the content of this project, and generate a new package from it.</p>	

You need to execute this procedure to be able to open the project and see all its content (the captured applications, the declared items, the automations, etc.).

DESCRIPTION

This package contains captures, datatype and automations that are described below. See chapter Version for more details about the version of the Desktop Agent and the SDK dependencies.

Settings

This section describes the settings of the project such as environment variables or dependent packages that are used in the automations.

Environment variables

N/A

Dependent packages

N/A

Captures

This section describes the captures which were made to pilot the application in this sample. It will also describe the different methods which were used to capture the pages and declare the items.

Datatypes

This section describes the datatype used in this sample. It describes the structure of the datatype and where it is used in the automations.

SaleOrderDetail

Name	Type	Description
orderType	String	
salesOrga	String	
distribChannel	String	
division	String	
PONumber	String	
soldToParty	String	
items	Array of String	
date	String	

Automations

Extract Data from PDF

Input/Output

Name	Type	Input/Output	Description
filePath	String	Input	
orderDetail	SaleOrderDetail	Output	

Description

The agent first opens a PDF document and retrieve all specific data using a dedicated activity “**Get Text (after)**”

Then the agent calculates the offset which is required, to read the data in the table. Left offset and right offset are calculated.

At the end of the automation, all data are inserted into an instance of a SaleOrderDetail datatype.

Note: The PDF library should not be used for scanned document. Otherwise the bot won't be able to retrieve any data from the document.

It can only be used with generated document

Merge PDF documents

Input/Output

Name	Type	Input/Output	Description
mergeOutputPath	String	Input	
inputMergeFolder	String	Input	Path of the folder containing the files to be merged
filesToMerge	String (List)	Input	Name of files to be merged

Description

The agent checks the existence of the input folder. It will produce a single PDF document using all the files provided as input parameters.

Split PDF document

Input/Output

Name	Type	Input/Output	Description
inputFile	String	Input	Path of the PDF document to be processed
Range	String	Input	Range of the input document to be extracted (ex: 1-3)
outputPath	String	Input	Name of the generated PDF document

Description

The agent checks the existence of the input folder then it will extract pages from the input file, using the *range* parameter.

VERSION

The product versions used to generate this sample are detailed below. This sample is provided “as is”, with no warranty that it will work correctly with other versions. If some versions of your software are different (such as the tool version or the target application version), you may need to recapture the application and/or update the workflow activities.

SAP Build Process Automation

This sample targets the Desktop Agent **2.0.13** or higher.

The following SDK dependencies were used to generate this sample: 1.28.47

See [documentation](#) for more details about the compatibility between SDK version and Desktop Agent.

Target application

N/A

PREREQUISITES

Global setup

SAP Build Process Automation must be installed in accordance with the installation guide available [here](#). An SAP Build Process Automation Factory must be available with a suitable environment (containing an agent). All information can be found in the “Getting Started” section accessible via the above link.

Specific steps to follow before launching the agent

1. Download the sample.zip archive on your machine and extract its content
2. When you launch the automation *Extract Data*, provide the path of the file that you just extracted

EXPECTED OUTPUT

The automation returns the content of the PDF document:

Output Parameters

```

  orderDetail
  orderType: "AVC"
  salesOrga: "0001"
  distribChannel: "01"
  division: "01"
  PONumber: "42"
  soldToParty: "ITAG01"
  items List, 3 elements
    0 List, 3 elements
      0: "HW-1761"
      1: "10"
      2: "0001"
    1 List, 3 elements
      0: "HW-1761"
      1: "20"
      2: "0001"
    2 List, 3 elements
      0: "HW-1761"
      1: "30"
      2: "0001"
  date: "17.11.2020"
```

www.sap.com/contactsap

© 2019 SAP SE or an SAP affiliate company. All rights reserved.

No part of this publication may be reproduced or transmitted in any form or for any purpose without the express permission of SAP SE or an SAP affiliate company.

The information contained herein may be changed without prior notice. Some software products marketed by SAP SE and its distributors contain proprietary software components of other software vendors. National product specifications may vary.

These materials are provided by SAP SE or an SAP affiliate company for informational purposes only, without representation or warranty of any kind, and SAP or its affiliated companies shall not be liable for errors or omissions with respect to the materials. The only warranties for SAP or SAP affiliate company products and services are those that are set forth in the express warranty statements accompanying such products and services, if any. Nothing herein should be construed as constituting an additional warranty.

In particular, SAP SE or its affiliated companies have no obligation to pursue any course of business outlined in this document or any related presentation, or to develop or release any functionality mentioned therein. This document, or any related presentation, and SAP SE's or its affiliated companies' strategy and possible future developments, products, and/or platform directions and functionality are all subject to change and may be changed by SAP SE or its affiliated companies at any time for any reason without notice. The information in this document is not a commitment, promise, or legal obligation to deliver any material, code, or functionality. All forward-looking statements are subject to various risks and uncertainties that could cause actual results to differ materially from expectations. Readers are cautioned not to place undue reliance on these forward-looking statements, and they should not be relied upon in making purchasing decisions.

SAP and other SAP products and services mentioned herein as well as their respective logos are trademarks or registered trademarks of SAP SE (or an SAP affiliate company) in Germany and other countries. All other product and service names mentioned are the trademarks of their respective companies. See www.sap.com/copyright for additional trademark information and notices.

THE BEST RUN

