Klippa

Klippa OCR API

Command Package

Readme

Version 1.1

12/10/2021

Table of Contents

1. Introduction

This document contains all essential information for the user to make full use of this A2019 *Klippa OCR API Command package*. It includes a description of the functions and capabilities and step-by-step procedures for setup & configuration of the *Klippa OCR API Command package*.

1.1 Overview

This command package can extract data from documents in a structured way. It leverages the Klippa OCR API for it. The Data is returned in JSON and as a key/value dictionary as returned by the API. The fields that are returned in a structured way can be viewed here: https://custom-ocr.klippa.com/docs

Submit your files to the API, convert them into structured data and continue with the desired output. All in a second's work.

The API supports formats such as JPG, PNG and PDF as input. Our system recognizes over 20 document types and 30 data fields.

1.2 Use cases

The key use cases include:

1. Extract data from invoices for further processing in bookkeeping or automatic booking
2. Extract data from identity cards to verify persons
3. Extract data from scanned documents to have the information in structured data

2. Requirements & Prerequisites

2.1 System Requirements

Enterprise A2019 (Cloud deployed) and Community Edition device requirements.
Review the machine hardware specifications, operating system versions, and browser types supported by Automation Anywhere Enterprise for creating and running bots and command packages as an Enterprise A2019 (Cloud deployed) or Community Edition user on your local machine.

2.2 Prerequisites

An API key for the Klippa OCR API. You can contact us here to get one:
https://www.klippa.com/en/ocr/ocr-api/request-a-key/

3. Getting Started

3.1 Quick Start

3.1.1 Setup

Make sure you have your Klippa OCR API key ready and also the template that you are going to use. When no template is given the default one is used, which extracts data for financial documents.

3.1.2 Configuration and Use

The command package provided a demo bot to show how the command package can be used. The package contains 1 command:

Parse document

Parses documents through the Klippa OCR API.

The command needs at least 2 values to be configured:

API Key: The Klippa OCR API key, can be credential, variable or insecure string

Document URL/Path: either an URL to a document or a file path, the file can be a control room file, desktop file or variable.

We also support some optional options:

Template: The template to use for the data extraction, every template has it's own purpose and output can differ from the default template.

PDF Text Extraction: Only used for PDF's. Full or Fast. Full always scans the whole document, Fast will try to extract text from the PDF for faster processing, if that fails it will use Full. Fast can sometime miss certain details when text is in images, for example the address of a company in a document footer.

User Data: Extra metadata in JSON format to give to the parser. Only works with templates that are configured to accept user data.

User Data Set External ID: The external ID of the user data set.

Hash Duplicate Group ID: An identifier to use when saving/detecting hash duplicates. This way you can allow to have the same document scanned more than once for multiple groups. When doing a scan, the combination of the Hash Group ID and the document Hash will be used to detect duplicates. This value is saved hashed on our side. Common use cases: Company ID, Campaign ID, User ID.

More documentation on the options can be found here: https://custom-ocr.klippa.com/docs#section/Introduction

The output of the document extraction can be put in a variable of type "Dictionary" and Subtype "Any". The key "response" will always contain a string with the plain JSON as replied by the API. The key "data" will contain a dictionary with the parsed data from the API.

4.   Support & FAQs

4.1   Support

Free bots are not officially supported.   You can get access to Community Support through the following channels:

- You can get access to Community Support, connecting with other Automation Anywhere customers and developers on APeople – the Bot Building Forum, the Bot Store Support Forum, or the Developers Everywhere Group.
- Automation Anywhere also provides a Product Documentation portal which can be accessed for more information about our products and guidance on Enterprise A2019.

4.2   FAQs

For questions relating to Enterprise A2019: See the Enterprise A2019 FAQs.

Appendix A: Record of Changes

No.	Version Number	Date of Change	Author	Notes
1	*1.0.0*	*21-01-2021*	*Jeroen Bobbeldijk*	*Initial version*

Appendix B: References

No.	Topic	Reference Link
1	Overview of Enterprise A2019	Click here
2	Guidance: Building basic A2019 bots	Click here
3	Guidance: Building A2019 action packages	Click here
4	APeople Community Forum	Click here
5	Automation Anywhere University	Click here