www.bytescout.com

document parser API in PHP and PDF.co Web API

PDF.co Web API: the Rest API that provides set of data extraction functions, tools for documents manipulation, splitting and merging of pdf files. Includes built-in OCR, images recognition, can generate and read barcodes from images, scans and pdf.

FOR MORE INFORMATION AND FREE TRIAL:

Download Free Trial SDK (on-premise version)

Read more about PDF.co Web API

Explore API Documentation

Get Free Training for PDF.co Web API

Get Free API key for Web API

visit www.ByteScout.com

Source Code Files:

SampleBloodReport.yml

```
templateVersion: 3
templatePriority: 0
sourceId: BloodTestTemplate
detectionRules:
  keywords: []
fields:
  PatientName:
    type: rectangle
    rectangle:
    - 177.75
    - 123.75
    - 62.25
    - 12.75
    pageIndex: 0
  ReportName:
    type: rectangle
    expression: '{{SmartDate}}'
```

```
    dataType: date
    rectangle:
    - 335.25
    - 94.5
    - 65.25
    - 12
    pageIndex: 0
  TestResults:
    type: rectangle
    dataType: table
    rectangle:
    - 41.25
    - 261.75
    - 532.5
    - 450.75
    pageIndex: 0
    rowMergingRule: byBorders
```

program.php

```
<!DOCTYPE html>
<html lang="en">
<head>
    <meta charset="UTF-8">
    <title>Document Parse Results</title>
</head>
<body>

<?php


// Get submitted form data
$apiKey = $_POST["apiKey"]; // The authentication key (API Key). Get your own by registering at https://app.pdf.co/do


// 1. RETRIEVE THE PRESIGNED URL TO UPLOAD THE FILE.
// * If you already have the direct PDF file link, go to the step 3.

// Create URL
$url = "https://api.pdf.co/v1/file/upload/get-presigned-url" .
    "?name=" . $_FILES["file"]["tmp_name"] .
    "&contenttype=application/octet-stream";

// Create request
$curl = curl_init();
curl_setopt($curl, CURLOPT_HTTPHEADER, array("x-api-key: " . $apiKey));
curl_setopt($curl, CURLOPT_URL, $url);
curl_setopt($curl, CURLOPT_RETURNTRANSFER, 1);
// Execute request
$result = curl_exec($curl);

if (curl_errno($curl) == 0)
{
    $status_code = curl_getinfo($curl, CURLINFO_HTTP_CODE);

    if ($status_code == 200)
    {
        $json = json_decode($result, true);

        // Get URL to use for the file upload
        $uploadFileUrl = $json["presignedUrl"];
        // Get URL of uploaded file to use with later API calls
        $uploadedFileUrl = $json["url"];

        // 2. UPLOAD THE FILE TO CLOUD.
```

```
        $localFile = $_FILES["fileInput"]["tmp_name"];
        $fileHandle = fopen($localFile, "r");

        curl_setopt($curl, CURLOPT_URL, $uploadFileUrl);
        curl_setopt($curl, CURLOPT_HTTPHEADER, array("content-type: application/octet-stream"));
        curl_setopt($curl, CURLOPT_PUT, true);
        curl_setopt($curl, CURLOPT_INFILE, $fileHandle);
        curl_setopt($curl, CURLOPT_INFILESIZE, filesize($localFile));

        // Execute request
        curl_exec($curl);

        fclose($fileHandle);

        if (curl_errno($curl) == 0)
        {
            $status_code = curl_getinfo($curl, CURLINFO_HTTP_CODE);

            if ($status_code == 200)
            {
                // Read all template texts
                $templateText = file_get_contents($_FILES["fileTemplate"]["tmp_name"]);

                // 3. PARSE UPLOADED PDF DOCUMENT
                ParseDocument($apiKey, $uploadedFileUrl, $templateText);
            }
            else
            {
                // Display request error
                echo "<p>Status code: " . $status_code . "</p>";
                echo "<p>" . $result . "</p>";
            }
        }
        else
        {
            // Display CURL error
            echo "Error: " . curl_error($curl);
        }
    }
    else
    {
        // Display service reported error
        echo "<p>Status code: " . $status_code . "</p>";
        echo "<p>" . $result . "</p>";
    }

    curl_close($curl);
}
else
{
    // Display CURL error
    echo "Error: " . curl_error($curl);
}

function ParseDocument($apiKey, $uploadedFileUrl, $templateText)
{
    // (!) Make asynchronous job
    $async = TRUE;

    // Prepare URL for Document parser API call.
    // See documentation: https://apidocs.pdf.co/?#1-pdfdocumentparser
    $url = "https://api.pdf.co/v1/pdf/documentparser" .
        "?async=" . $async;

    // Post fields
    $data = array('url'=>$uploadedFileUrl, 'template'=>$templateText);

    // Create request
    $curl = curl_init();
    curl_setopt($curl, CURLOPT_HTTPHEADER, array("x-api-key: " . $apiKey));
    curl_setopt($curl, CURLOPT_URL, $url);
    curl_setopt($curl, CURLOPT_POST, true);
    curl_setopt($curl, CURLOPT_POSTFIELDS, $data);
    curl_setopt($curl, CURLOPT_RETURNTRANSFER, 1);

    // Execute request
    $result = curl_exec($curl);
```

```
    echo $result . "<br/>";

    if (curl_errno($curl) == 0)
    {
        $status_code = curl_getinfo($curl, CURLINFO_HTTP_CODE);

        if ($status_code == 200)
        {
            $json = json_decode($result, true);

            if ($json["error"] == false)
            {
                // URL of generated JSON file that will available after the job completion
                $resultFileUrl = $json["url"];
                // Asynchronous job ID
                $jobId = $json["jobId"];

                // Check the job status in a loop
                do
                {
                    $status = CheckJobStatus($jobId, $apiKey); // Possible statuses: "working", "failed", "aborted", "success"

                    // Display timestamp and status (for demo purposes)
                    echo "<p>" . date(DATE_RFC2822) . ": " . $status . "</p>";

                    if ($status == "success")
                    {
                        // Display link to JSON file with information about parsed fields
                        echo "<div><h2>Parsing Result:</h2><a href="" . $resultFileUrl . "' target='_blank'>" . $resultFileUrl . "<
                        break;
                    }
                    else if ($status == "working")
                    {
                        // Pause for a few seconds
                        sleep(3);
                    }
                    else
                    {
                        echo $status . "<br/>";
                        break;
                    }
                }
                while (true);
            }
            else
            {
                // Display service reported error
                echo "<p>Error: " . $json["message"] . "</p>";
            }
        }
        else
        {
            // Display request error
            echo "<p>Status code: " . $status_code . "</p>";
            echo "<p>" . $result . "</p>";
        }
    }
    else
    {
        // Display CURL error
        echo "Error: " . curl_error($curl);
    }
}

function CheckJobStatus($jobId, $apiKey)
{
    $status = null;

    // Create URL
    $url = "https://api.pdf.co/v1/job/check?jobid=" . $jobId;

    // Create request
    $curl = curl_init();
    curl_setopt($curl, CURLOPT_HTTPHEADER, array("x-api-key: " . $apiKey));
    curl_setopt($curl, CURLOPT_URL, $url);
    curl_setopt($curl, CURLOPT_RETURNTRANSFER, 1);

    // Execute request
```

```
    $result = curl_exec($curl);

    if (curl_errno($curl) == 0)
    {
        $status_code = curl_getinfo($curl, CURLINFO_HTTP_CODE);

        if ($status_code == 200)
        {
            $json = json_decode($result, true);

            if ($json["error"] == false)
            {
                $status = $json["status"];
            }
            else
            {
                // Display service reported error
                echo "<p>Error: " . $json["message"] . "</p>";
            }
        }
        else
        {
            // Display request error
            echo "<p>Status code: " . $status_code . "</p>";
            echo "<p>" . $result . "</p>";
        }
    }
    else
    {
        // Display CURL error
        echo "Error: " . curl_error($curl);
    }

    // Cleanup
    curl_close($curl);

    return $status;
}

?>

</body>
</html>
```

VIDEO

https://www.youtube.com/watch?v=NEwNs2b9YN8

ON-PREMISE OFFLINE SDK

60 Day Free Trial or Visit PDF.co Web API Home Page
Explore PDF.co Web API Documentation
Explore Samples
Sign Up for PDF.co Web API Online Training

ON-DEMAND REST WEB API

Get Your API Key
Explore Web API Docs
Explore Web API Samples

visit www.ByteScout.com

visit www.PDF.co