

www.bytescout.com

How to create custom template for document parser API in PHP with ByteScout Cloud API Server

ByteScout Cloud API Server: the ready to deploy Web API Server that can be deployed in less than thirty minutes into your own in-house Windows server (no Internet connection is required to process data!) or into private cloud server. Can store data on in-house local server based storage or in Amazon AWS S3 bucket. Processing data solely on the server using built-in ByteScout powered engine, no cloud services are used to process your data!.

FOR MORE INFORMATION AND FREE TRIAL:

[Download Free Trial SDK \(on-premise version\)](#)

[Read more about **ByteScout Cloud API Server**](#)

[Explore API Documentation](#)

[Get Free Training for ByteScout Cloud API Server](#)

[Get Free API key for Web API](#)

[visit www.Bytescout.com](http://www.Bytescout.com)

Source Code Files:

SampleTemplate.yml

```
templateName: My Custom Template
templateVersion: 4
templatePriority: 0
detectionRules:
  keywords:
    - Your Company Name
    - Invoice No\
    - TOTAL
objects:
- name: total
  objectType: field
  fieldProperties:
    fieldType: macros
    expression: TOTAL{{Spaces}}{{Number}}
    dataType: decimal
    pageIndex: 0
- name: dateIssued
  objectType: field
  fieldProperties:
    fieldType: macros
    expression: Invoice Date {{SmartDate}}
    dataType: date
    dateFormat: auto-mdy
    pageIndex: 0
- name: invoiceId
  objectType: field
  fieldProperties:
    fieldType: macros
    expression: Invoice No. {{Digits}}
    pageIndex: 0
- name: companyName
  objectType: field
  fieldProperties:
    fieldType: static
    expression: Vendor Company
    pageIndex: 0
- name: billTo
  objectType: field
  fieldProperties:
    fieldType: rectangle
    rectangle:
      - 32.25
      - 150
      - 348
      - 70.5
    pageIndex: 0
- name: notes
  objectType: field
  fieldProperties:
    fieldType: rectangle
    rectangle:
      - 32.25
      - 227.25
      - 531
      - 47.25
    pageIndex: 0
- name: table1
  objectType: table
  tableProperties:
```

```

start:
  expression: Item\s+Quantity\s+Price\s+Total
  regex: true
end:
  expression: TOTAL
  regex: true
row:
  expression: ^\s*(?<description>\w+.*)(?<quantity>\d+)\s+(?<unitPrice>\d+\.\d{2})
  regex: true

```

program.php

```

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

<?php

// Please NOTE: In this sample we're assuming Cloud Api Server is hosted at "https://l
// If it's not then please replace this with with your hosting url.

// Get submitted form data
$apiKey = $_POST["apiKey"]; // The authentication key (API Key). Get your own by regis

// 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://localhost/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-
    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://localhost/pdf/documentparser";

    // Prepare requests params
    $parameters = array();
    $parameters["url"] = $uploadedFileUrl;
    $parameters["template"] = $templateText;
    $parameters["async"] = $async;

    // Create Json payload
    $data = json_encode($parameters);

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

    // 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 complete
                $resultFileUrl = $json["url"];
                // Asynchronous job ID
                $jobId = $json["jobId"];

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

                    // 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 file
        echo "<div><h2>Parsing Result:</h2><a href='" . $resultFileUrl
        break;
    }
    else if ($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://localhost/job/check";

    // Prepare requests params
    $parameters = array();
    $parameters["jobid"] = $jobId;

    // Create Json payload
    $data = json_encode($parameters);

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

    // 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 ByteScout Cloud API Server Home Page](#)

[Explore ByteScout Cloud API Server Documentation](#)

[Explore Samples](#)

[Sign Up for ByteScout Cloud API Server Online Training](#)

ON-DEMAND REST WEB API

[Get Your API Key](#)

[Explore Web API Docs](#)

[Explore Web API Samples](#)

[visit www.ByteScout.com](http://www.ByteScout.com)

[visit www.PDF.co](http://www.PDF.co)

www.bytescout.com