

www.bytescout.com

How to split PDF by text from URL (node for PDF splitting API in JavaScript using PDF.co Web API)

PDF.co Web API: the flexible Web API that includes full set of functions from e-signature requests to data extraction, OCR, images recognition, pdf splitting and pdf splitting. Can also generate barcodes 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](http://www.ByteScout.com)

Source Code Files:

app.js

```
var https = require("https");  
var fs = require("fs");
```

```

// Use "npm install request" command to install.
var request = require("request");

// The authentication key (API Key).
// Get your own by registering at https://app.pdf.co/documentation/api
const API_KEY = "*****";

// Source PDF file to split
const SourceFileUrl = "https://bytescout-com.s3-us-west-2.amazonaws.com/files/demo-fil
// Split Search String
const SplitText = "invoice number";

// Prepare request to `Split PDF By Text` API endpoint
var queryPath = `/v1/pdf/split2`;

// JSON payload for api request
var jsonPayload = JSON.stringify({
  searchString: SplitText, url: SourceFileUrl, async: true
});

var reqOptions = {
  host: "api.pdf.co",
  method: "POST",
  path: queryPath,
  headers: {
    "x-api-key": API_KEY,
    "Content-Type": "application/json",
    "Content-Length": Buffer.byteLength(jsonPayload, 'utf8')
  }
};

// Send request
var postRequest = https.request(reqOptions, (response) => {
  response.on("data", (d) => {
    // Parse JSON response
    var data = JSON.parse(d);
    if (data.error == false) {
      console.log(`Job #${data.jobId} has been created!`);
      checkIfJobIsCompleted(data.jobId, data.url);
    }
    else {
      // Service reported error
      console.log(data.message);
    }
  });
}).on("error", (e) => {
  // Request error
  console.error(e);
});

// Write request data
postRequest.write(jsonPayload);
postRequest.end();

function checkIfJobIsCompleted(jobId, resultFileUrlJson) {
  let queryPath = `/v1/job/check`;

  // JSON payload for api request

```

```

let jsonPayload = JSON.stringify({
  jobId: jobId
});

let reqOptions = {
  host: "api.pdf.co",
  path: queryPath,
  method: "POST",
  headers: {
    "x-api-key": API_KEY,
    "Content-Type": "application/json",
    "Content-Length": Buffer.byteLength(jsonPayload, 'utf8')
  }
};

// Send request
var postRequest = https.request(reqOptions, (response) => {
  response.on("data", (d) => {
    response.setEncoding("utf8");

    // Parse JSON response
    let data = JSON.parse(d);
    console.log(`Checking Job #${jobId}, Status: ${data.status}, Time: ${new D

    if (data.status == "working") {
      // Check again after 3 seconds
      setTimeout(function () { checkIfJobIsCompleted(jobId, resultFileUrlJso
    }
    else if (data.status == "success") {

      request({ method: 'GET', uri: resultFileUrlJso, gzip: true },
        function (error, response, body) {

          // Parse JSON response
          let respJsonFileArray = JSON.parse(body);
          let part = 1;

          respJsonFileArray.forEach((url) => {
            var localFileName = `./part${part}.png`;
            var file = fs.createWriteStream(localFileName);
            https.get(url, (response2) => {
              response2.pipe(file)
                .on("close", () => {
                  console.log(`Generated PDF file saved as "${lo
                });
            });
            part++;
          }, this);

        });
    }
    else {
      console.log(`Operation ended with status: "${data.status}".`);
    }
  })
});

// Write request data
postRequest.write(jsonPayload);
postRequest.end();

```

```
}
```

package.json

```
{
  "name": "test",
  "version": "1.0.0",
  "description": "PDF.co",
  "main": "app.js",
  "scripts": {
  },
  "keywords": [
    "pdf.co",
    "web",
    "api",
    "bytescout",
    "api"
  ],
  "author": "ByteScout & PDF.co",
  "license": "ISC",
  "dependencies": {
    "request": "^2.88.2"
  }
}
```

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

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

www.bytescout.com