



Processor



Extraction



PROCESSOR > GUIDES > EXTRACTION

Extract key-value pairs from bank statements



PSPDFKit Processor has been deprecated and replaced by [Document Engine](#). To migrate to Document Engine and unlock advanced document processing capabilities, refer to our migration guide. Learn more about these enhancements on our [blog](#).

This guide explains how to extract key-value pairs (KVPs) from bank statements using Processor. For example, this enables you to extract IBANs or account numbers. For more information, refer to the guide on [how key-value pair extraction works](#).

Before you get started, make sure [Processor is up and running](#).

You can download and use either of the following sample documents for the examples in this guide:

- ✧ [Example eight-page PDF](#)
- ✧ [Example four-page PDF](#)

You'll be sending [multipart POST requests with instructions](#) to Processor's `/build` endpoint. To learn more about multipart requests, refer to our blog post on the topic, [A Brief Tour of Multipart Requests](#).

Check out the [API Reference](#) to learn more about the `/build` endpoint and all the actions you can perform on PDFs with PSPDFKit Processor.

Sending the Request to Extract Data



To extract key-value pairs from a bank statement, post a multipart request to the `/build` API endpoint. In the instructions, specify the following output parameters:

- ❖ `type` specifies the output type. Set this to `json-content`.
- ❖ `keyValuePairs` is a Boolean value that determines whether to extract key-value pairs.
- ❖ `language` specifies the language used for recognizing text with optical character recognition (OCR). Sometimes, text is stored in a PDF or an image in a way that makes it so you cannot search or copy it. PSPDFKit's OCR engine allows you to recognize text and save it in a separate file where you can both search and copy and paste the text. For more information, refer to the list of supported languages.

SHELL

HTTP

```
1 curl -X POST http://localhost:5000/api/build \  
2   -F document=@/path/to/example-document.pdf \  
3   -F instructions='{  
4     "parts": [  
5       {  
6         "file": "document"  
7       }  
8     ],  
9     "output": {  
10      "type": "json-content",  
11      "keyValuePairs": true,  
12      "language": "english"  
13    }  
14  }' \  
15    -o result.pdf
```



For more information on the `/build` instructions, refer to the [API Reference](#).

Example Data Extraction Response

```
1 {  
2   "pages": [  
3     {  
4       "pageIndex": 0,  
5       "keyValuePairs": [  
6         {  
7           "confidence": 95.4,  
8           "key": {  
9             "bbox": {
```



```
10         "left": 0,
11         "top": 0,
12         "width": 100,
13         "height": 100
14     },
15     "content": "IBAN"
16 },
17 "value": {
18     "bbox": {
19         "left": 0,
20         "top": 0,
21         "width": 100,
22         "height": 100
23     },
24     "content": "FR7611808009101234567890147",
25     "dataType": "String"
26 }
27 }
28 ]
29 }
30 ]
31 }
```

Was this helpful?

 YES

 NO

Questions? [Contact us](#)

