



Processor



Conversion



PROCESSOR > GUIDES > CONVERSION

Convert PDF to PDF/A on Linux



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](#).

PDF/A is a document format intended for long-term preservation. PSPDFKit Processor supports converting source files into all PDF/A versions and conformance levels:

- ✧ PDF/A-1a, PDF/A-1b
- ✧ PDF/A-2a, PDF/A-2u, PDF/A-2b
- ✧ PDF/A-3a, PDF/A-3u, PDF/A-3b
- ✧ PDF/A-4, PDF/A-4e, PDF/A-4f

For more information on the long-term preservation of documents, check out our demo video below, or have a look at our [complete guide to PDF/A](#).



ASK AI



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.

Converting Documents to PDF/A

To generate a PDF/A document using `/build`, specify the relevant options in the `output` section of the build `instructions`, as shown in the following example:

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": "pdfa",
11     "conformance": "pdfa-2a",
12     "vectorization": true,
13     "rasterization": true
14 }
15 }' \
16 -o result.pdf
```

Configuring PDF/A Conversion

PDF/A documents are intended for long-term preservation, and their structure is different from PDF documents. To ensure compliance with your chosen conformance level, the conversion process may introduce changes to the document's content or appearance. This might change the document by adding, editing, or removing document structure elements, embedding fonts, etc.

In some cases, direct conversion isn't possible. Processor then uses other techniques such as vectorization and rasterization:

- ⌘ Vectorization means that if some document elements cannot be used directly in the PDF/A output, they're embedded in the output document as vector-based graphic elements. This technique is typically used for fonts and paths.
- ⌘ Rasterization means that if some document content cannot be used directly in the PDF/A output, it's embedded in the output document as raster images.

Both approaches result in the loss of fonts and text information because the text is converted into shapes and raster images. Text information can later be recovered using optical character recognition (OCR).

To control whether PSPDFKit Processor uses the vectorization and rasterization techniques if necessary, set the `vectorization` and `rasterization` options to `true`.

Licensing

To convert documents to PDF/A with PSPDFKit Processor, PDF/A needs to be included in your PSPDFKit Processor license. [Contact Sales](#) to add PDF/A to your license. After it's added to your license, update the offline `LICENSE_KEY` in your PSPDFKit Processor configuration.

PDF/A Validation

PSPDFKit Processor also supports validating the conformance level of existing PDF/A documents. To learn more about how to validate PDF/A conformance, refer to the [PDF/A validation guide](#).

Was this helpful?

☒ YES

☐ NO

Questions? [Contact us](#)

