

# Converting a PDF to a series of images with Python

Asked 16 years ago   Modified 6 years, 2 months ago   Viewed 49k times



51



I'm attempting to use Python to convert a multi-page PDF into a series of JPEGs. I can split the PDF up into individual pages easily enough with available tools, but I haven't been able to find anything that can covert PDFs to images.

PIL does not work, as it can't read PDFs. The two options I've found are using either GhostScript or ImageMagick through the shell. This is not a viable option for me, since this program needs to be cross-platform, and I can't be sure either of those programs will be available on the machines it will be installed and used on.

Are there any Python libraries out there that can do this?

python

pdf

imagemagick

jpeg

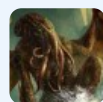
python-imaging-library

Share

Improve this question

Follow

asked Dec 1, 2008 at 19:31



Jaearess

675 ● 1 ● 7 ● 10

Did you ever find a solution? I got converting to work, but can't seem to get ImageMagick to respect pages.

– [stormlifter](#) Aug 1, 2012 at 23:53

Hope this helps -

[\[stackoverflow.com/questions/34129995/pdf-to-raw-image\]](https://stackoverflow.com/questions/34129995/pdf-to-raw-image)  
([stackoverflow.com/questions/34129995/pdf-to-raw-image](https://stackoverflow.com/questions/34129995/pdf-to-raw-image))

– [SystemOverflow LLC](#) Dec 7, 2015 at 16:09

5 Answers

Sorted by:

Highest score (default)



20

Share Improve this answer

edited Oct 20, 2018 at 17:18

Follow



[Ketul](#)

146 ● 7

answered Dec 1, 2008 at 19:33



[Adam Rosenfield](#)

399k ● 101 ● 522 ● 597



2 I'd like to kick in that yeah, you can just include the ImageMagick library with your project -- just make sure you review the license terms so you can put the appropriate stuff in your readme – [Coderer](#) Dec 18, 2008 at 15:43

28 Care to elaborate on what bindings might be useful? – [Ivo Flipse](#) Feb 12, 2013 at 12:49



Here's whats worked for me using the python ghostscript module (installed by '\$ pip install ghostscript'):

9



```
import ghostscript

def pdf2jpeg(pdf_input_path, jpeg_output_path):
    args = ["pdf2jpeg", # actual value doesn't matter
            "-dNOPAUSE",
            "-sDEVICE=jpeg",
            "-r144",
            "-sOutputFile=" + jpeg_output_path,
            pdf_input_path]
    ghostscript.Ghostscript(*args)
```

I also installed Ghostscript 9.18 on my computer and it probably wouldn't have worked otherwise.

Share Improve this answer

answered Mar 20, 2016 at 11:13

Follow



**Idan Yacobi**

91 ● 1 ● 3

---

This seems to create the new image, and I can see the pdf thumbnail, but for some reason it doesn't close the file of the new jpeg created. I got this output from python #####  
246643328 c\_void\_p(246643328L) any ideas? – [Jed](#) Jun 5, 2017 at 22:28

---

can you help solve my issue on this case?  
[stackoverflow.com/questions/44448552/...](https://stackoverflow.com/questions/44448552/) – [Jed](#) Jun 15, 2017 at 18:05

---

3 ghostscript does not seem to have support for python3  
– [unlockme](#) Nov 15, 2017 at 16:48

---

1 @Idan Yacobi, thank you! this is the best solution for my python27. – [Mark K](#) Aug 24, 2018 at 2:19

---



4



You can't avoid the Ghostscript dependency. Even Imagemagick relies on Ghostscript for its PDF reading functions. The reason for this is the complexity of the PDF format: a PDF doesn't just contain bitmap information, but mostly vector shapes, transparencies etc. Furthermore it is quite complex to figure out which of these objects appear on which page.

So the correct rendering of a PDF Page is clearly out of scope for a pure Python library.

The good news is that Ghostscript is pre-installed on many windows and Linux systems, because it is also needed by all those PDF Printers (except Adobe Acrobat).

Share Improve this answer

Follow

answered Mar 18, 2009 at 10:27



Franz

1,344 ● 8 ● 2



1



If you're using linux some versions come with a command line utility called 'pdftopbm' out of the box. Check out [netpbm](#)

Share Improve this answer

Follow

answered Jan 4, 2010 at 20:58



Jay

14.4k ● 5 ● 46 ● 74



Perhaps relevant:

[http://www.swftools.org/gfx\\_tutorial.html](http://www.swftools.org/gfx_tutorial.html)

1



Share Improve this answer

answered Dec 13, 2011 at 23:09

Follow



[mattbasta](#)

13.7k ● 9 ● 50 ● 70



---

11 If it's relevant, then supply the relevant parts in your answer  
– [Rebs](#) Feb 27, 2015 at 6:10

---