

```
!pip install pytesseract opencv-python-headless easyocr
!apt-get install poppler-utils
```

```
from pdf2image import convert_from_path
# Upload the PDF (use your uploaded file)
pdf_path = '/content/ES-AHPHU - J-000312-0014 - 1579.p
# Convert PDF to image (use dpi=300 for better quality)
images = convert_from_path(pdf_path, dpi=300)
for idx, img in enumerate(images):
    img.save(f"page_{idx}.png")
```

```
from google.colab import files
uploaded = files.upload()
```

```
import easyocr
import matplotlib.pyplot as plt
from PIL import Image
reader = easyocr.Reader(['en'])
# Run OCR on the first page
img_path = "page_0.png"
bounds = reader.readtext(img_path, detail=0) # detail=
# Display extracted text
print("\n".join(bounds))
# Optional: View image
img = Image.open(img_path)
plt.imshow(img)
```