

```
from swarms.utils.try_except_wrapper import try_except_wrapper
```

```
try:
```

```
    import pypdf
```

```
except ImportError:
```

```
    import subprocess
```

```
    subprocess.check_call(["python", "-m", "pip", "install", "pypdf"])
```

```
    import pypdf
```

```
@try_except_wrapper
```

```
def pdf_to_text(pdf_path: str) -> str:
```

```
    """
```

```
    Converts a PDF file to a string of text.
```

```
    Args:
```

```
        pdf_path (str): The path to the PDF file to be converted.
```

```
    Returns:
```

```
        str: The text extracted from the PDF.
```

```
    Raises:
```

```
        FileNotFoundError: If the PDF file is not found at the specified path.
```

```
        Exception: If there is an error in reading the PDF file.
```

```
    """
```

try:

# Open the PDF file

with open(pdf\_path, "rb") as file:

pdf\_reader = pypdf.PdfReader(file)

text = ""

# Iterate through each page and extract text

for page in pdf\_reader.pages:

text += page.extract\_text() + "\n"

return text

except FileNotFoundError:

raise FileNotFoundError(

f"The file at {pdf\_path} was not found."

)

except Exception as e:

raise Exception(

f"An error occurred while reading the PDF file: {e}"

)

# Example usage

# text = pdf\_to\_text("test.pdf")

# print(text)