**Documentação do Agente OCR para Conversão de PDFs em Word**

**1. Objetivo**

O Agente OCR foi desenvolvido para converter arquivos PDF baseados em imagens em documentos Word editáveis, utilizando OCR para extrair texto.

**2. Requisitos do Sistema**

**Softwares Necessários**

* **Python 3.10+** (<https://www.python.org/>)
* **Tesseract OCR** ([Baixar aqui](https://github.com/UB-Mannheim/tesseract/wiki))
* **Poppler** ([Baixar aqui](https://github.com/oschwartz10612/poppler-windows/releases/))

**Bibliotecas Python Necessárias**

Antes de rodar o agente, instale as dependências com o seguinte comando:

pip install pytesseract pdf2image python-docx tkinter

**3. Como Usar o Agente?**

**Passo a Passo**

1. **Abra o Prompt de Comando (CMD)**
   * Pressione Win + R, digite cmd e pressione Enter.
2. **Acesse a pasta onde o script está salvo**
3. cd C:\Projetos
4. **Execute o script**
5. python converter.py
6. **Selecione o PDF e defina onde salvar o arquivo Word.**
7. O agente converterá o PDF em Word e limpará automaticamente a pasta temp\_images.

**4. O que o Agente Faz?**

✅ Converte **PDFs baseados em imagem** para texto.  
✅ Mantém a **formatação básica** do documento.  
✅ Limpa automaticamente a pasta **temp\_images** após a conversão.  
✅ Gera um **arquivo Word editável**.

**5. Possíveis Erros e Soluções**

**Erro: "Tesseract not found"**

* **Solução**: Verifique se o **Tesseract está instalado e no PATH**.
* Teste no CMD:
* tesseract --version
* Se o comando não for reconhecido, adicione o caminho do Tesseract ao PATH do Windows:
* C:\Arquivos de Programas\Tesseract-OCR\

**Erro: "Poppler not installed"**

* **Solução**: Baixe e instale o Poppler e adicione ao PATH.

**Erro: "PDF não encontrado"**

* **Solução**: Certifique-se de que o arquivo existe e que o caminho foi digitado corretamente.

**6. Conclusão**

Este agente está pronto para uso e pode ser utilizado para conversão rápida e eficiente de documentos PDF para Word. Caso deseje escalar o serviço para atender terceiros, podemos explorar soluções em nuvem.

🚀 **Agora você tem um agente OCR funcional!** 😃