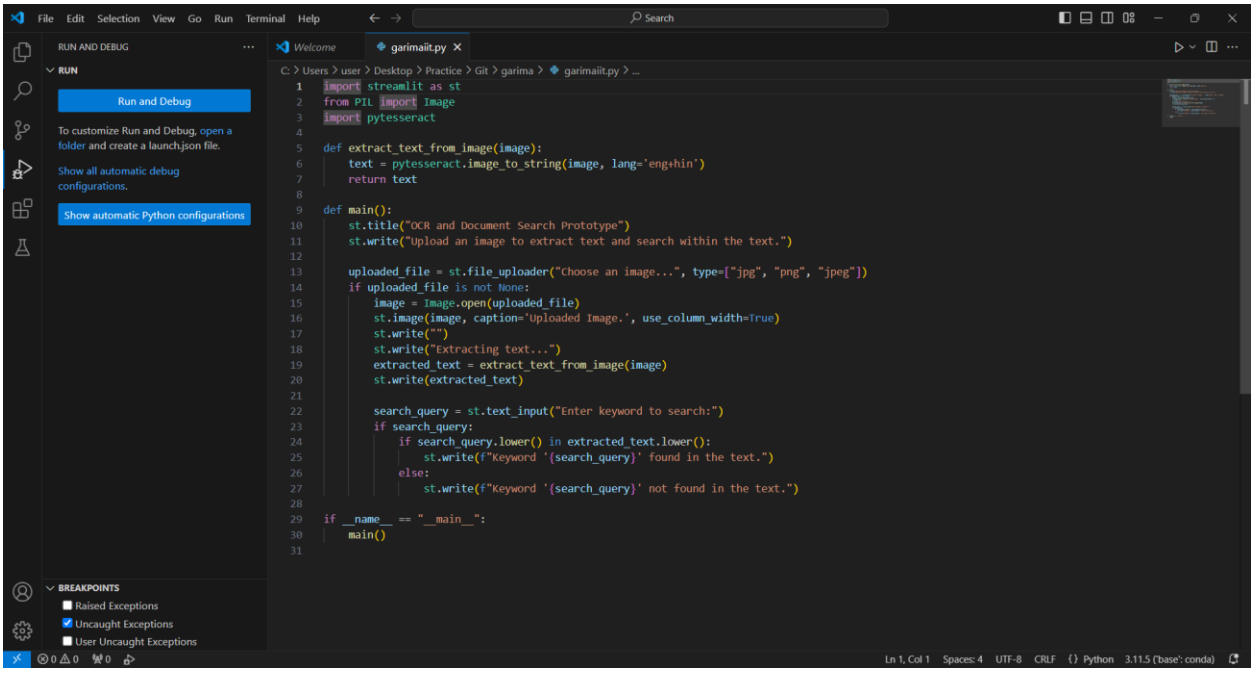
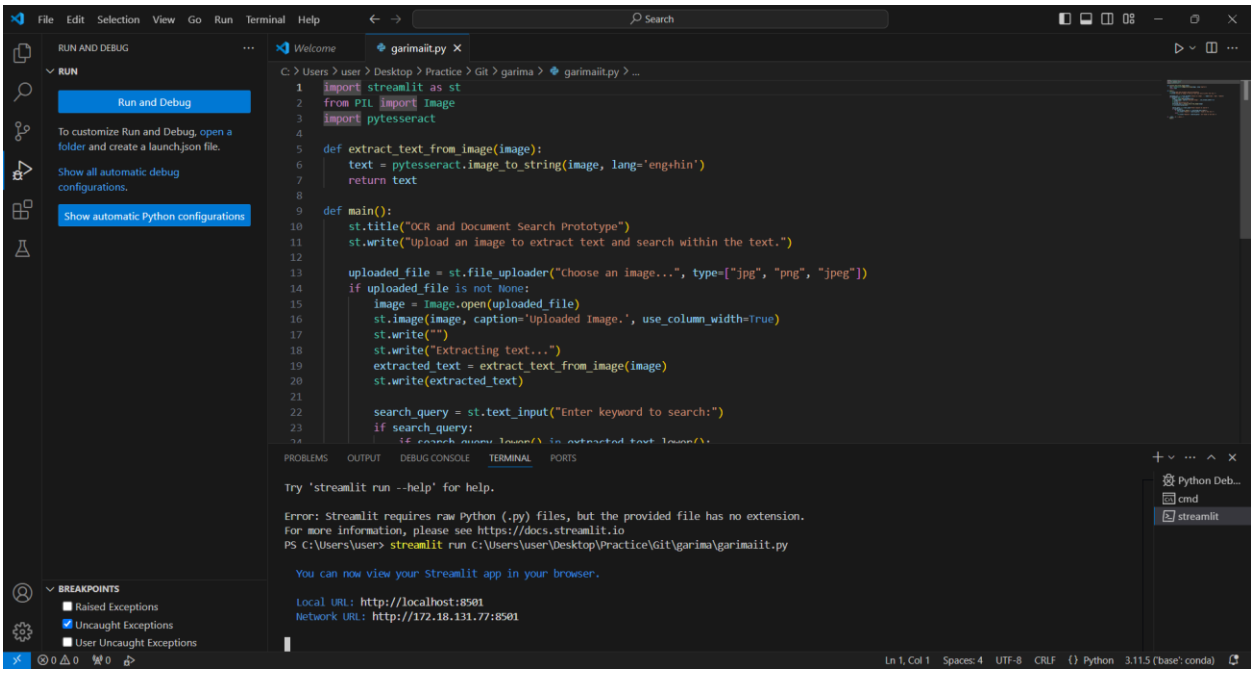
Code in Python



```python
import streamlit as st
from PIL import Image
import pytesseract

def extract_text_from_image(image):
    text = pytesseract.image_to_string(image, lang='eng+hin')
    return text

def main():
    st.title("OCR and Document Search Prototype")
    st.write("Upload an image to extract text and search within the text.")

    uploaded_file = st.file_uploader("Choose an image...", type=["jpg", "png", "jpeg"])
    if uploaded_file is not None:
        image = Image.open(uploaded_file)
        st.image(image, caption='Uploaded Image.', use_column_width=True)
        st.write("")
        st.write("Extracting text...")
        extracted_text = extract_text_from_image(image)
        st.write(extracted_text)

        search_query = st.text_input("Enter keyword to search:")
        if search_query:
            if search_query.lower() in extracted_text.lower():
                st.write(f"Keyword '{search_query}' found in the text.")
            else:
                st.write(f"Keyword '{search_query}' not found in the text.")

if __name__ == "__main__":
    main()
```

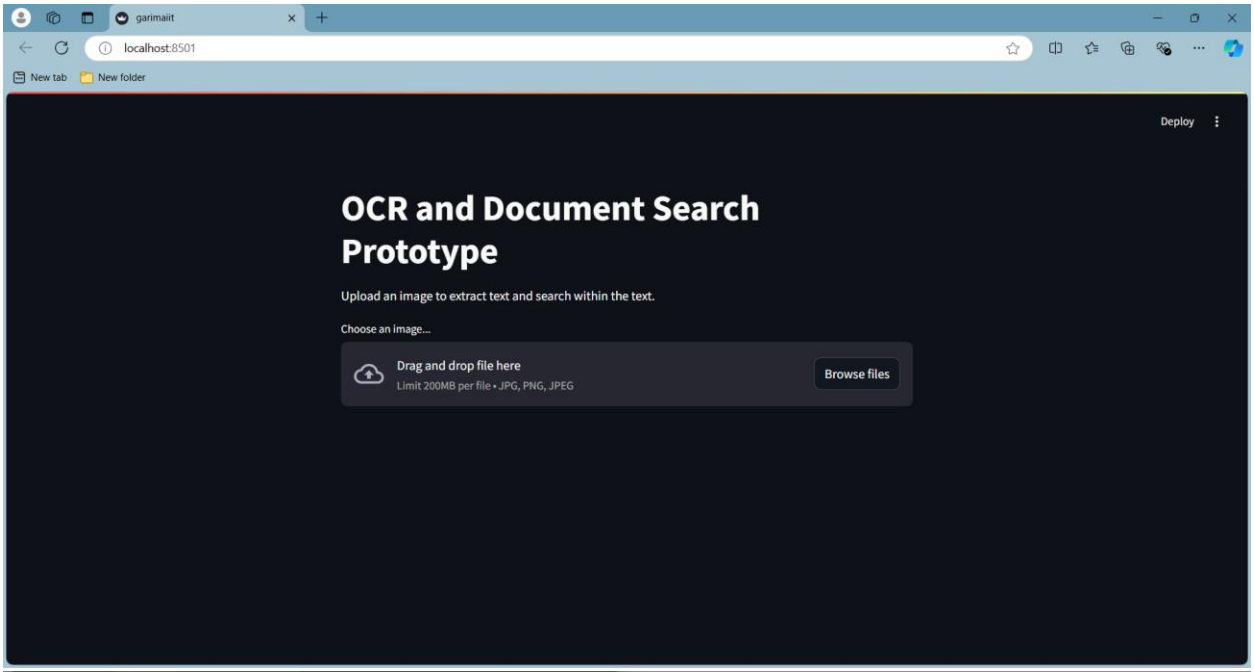
OUTPUT

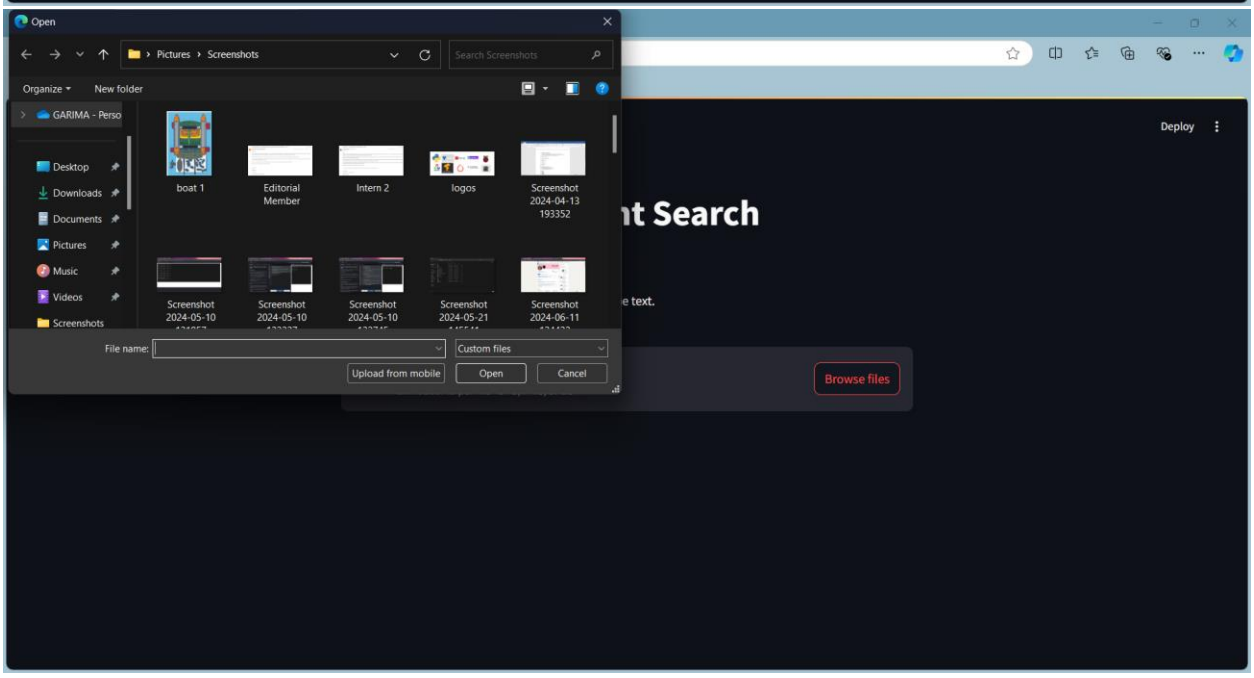


```
Try 'streamlit run --help' for help.

Error: Streamlit requires raw Python (.py) files, but the provided file has no extension.
For more information, please see https://docs.streamlit.io
PS C:\Users\user> streamlit run C:\Users\user\Desktop\Practice\Git\garima\garimaiit.py

  You can now view your Streamlit app in your browser.

  Local URL: http://localhost:8501
  Network URL: http://172.18.131.77:8501
```

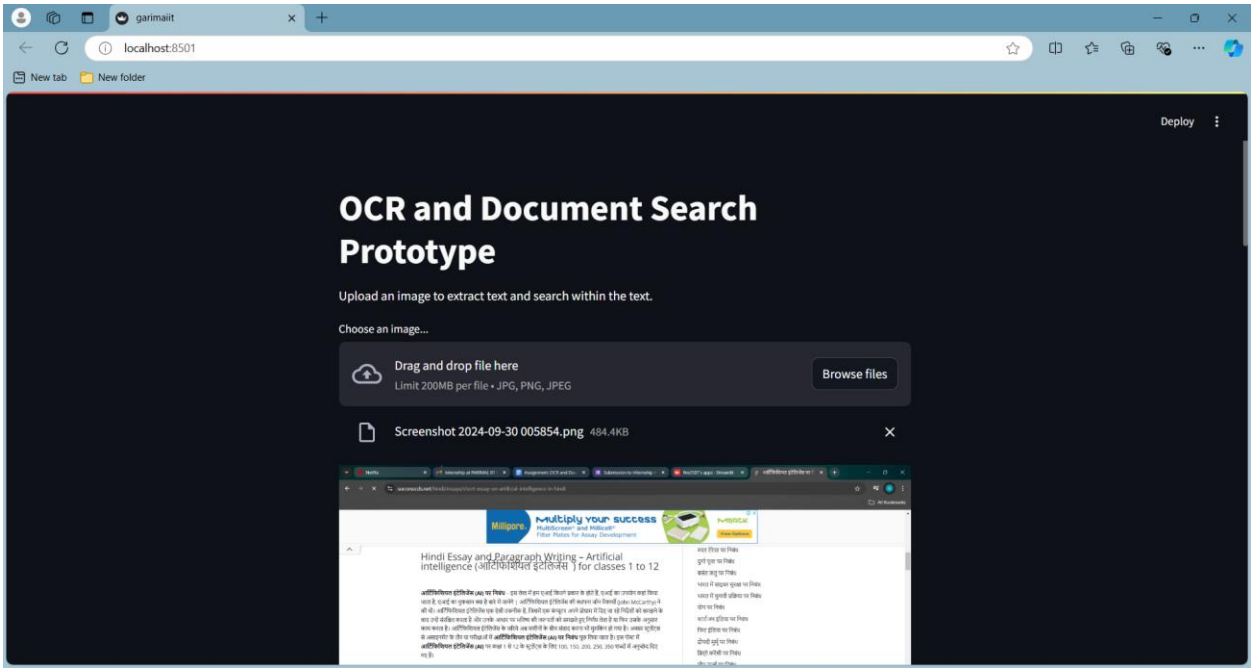
TEXT READER FROM UPLOADED IMAGES

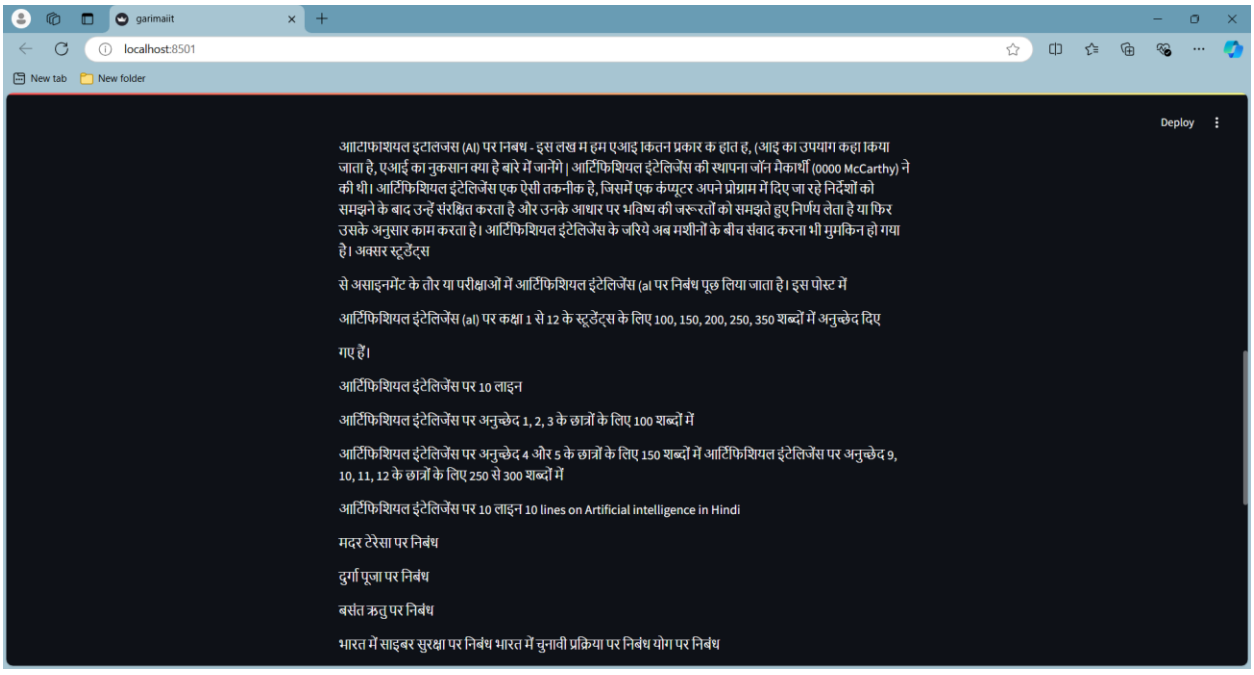
FRONTPAGE



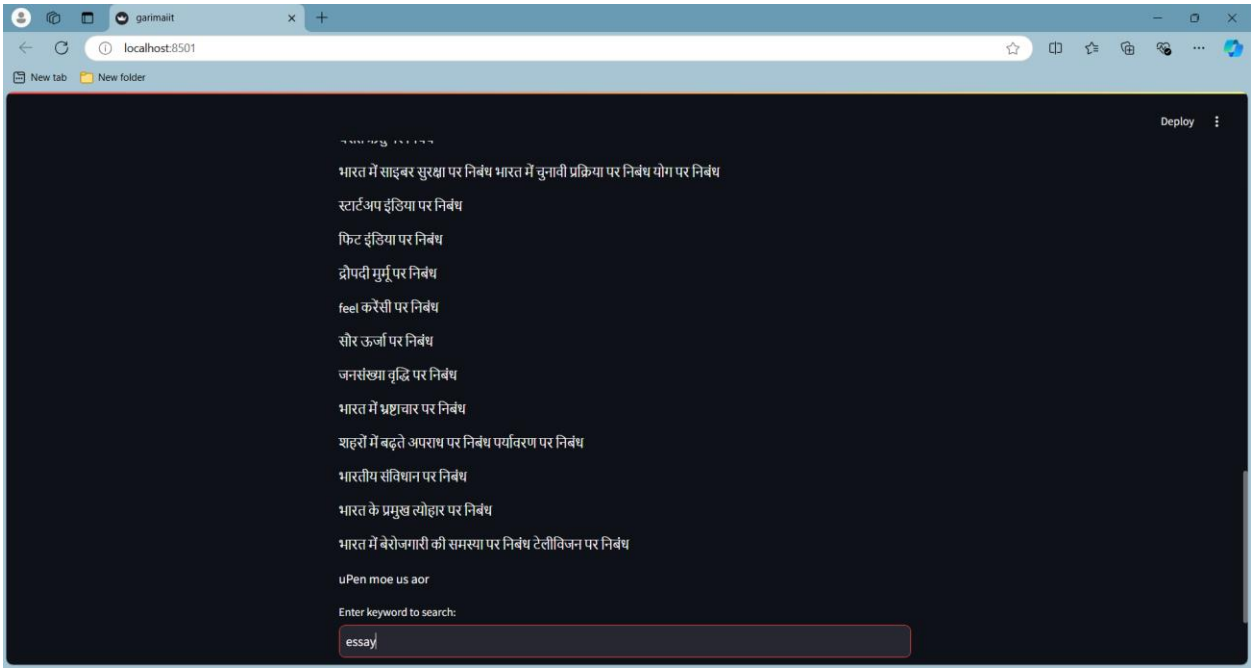


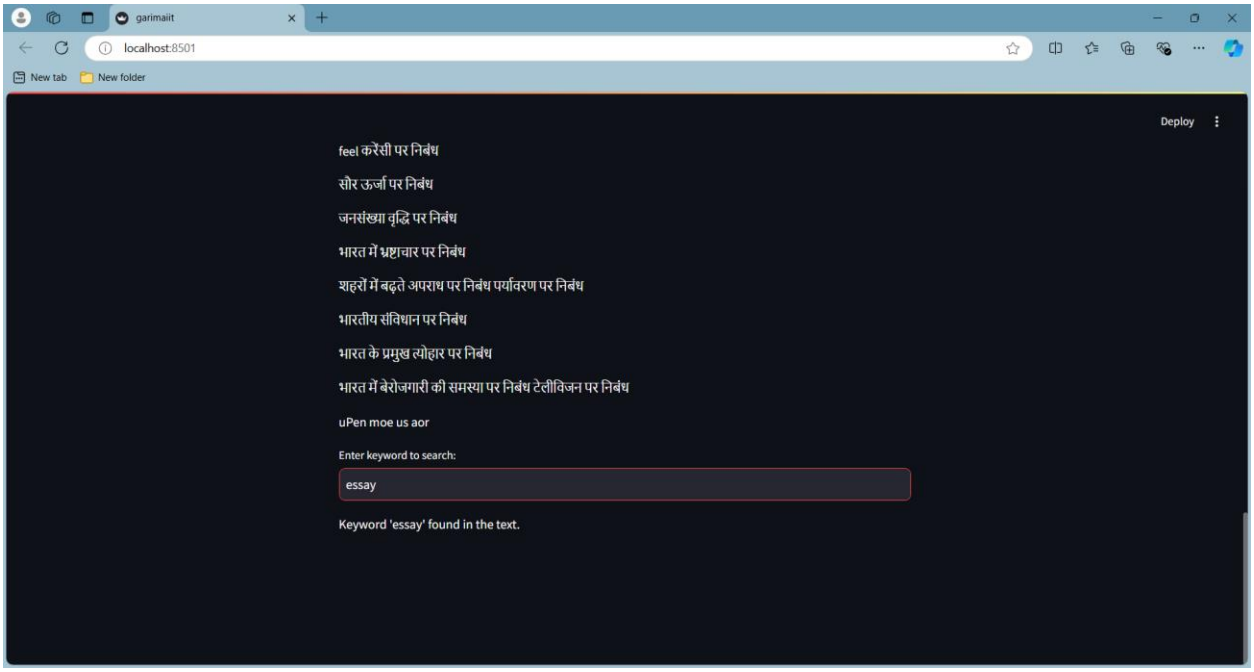
HINDI READING AND EXTRACTION



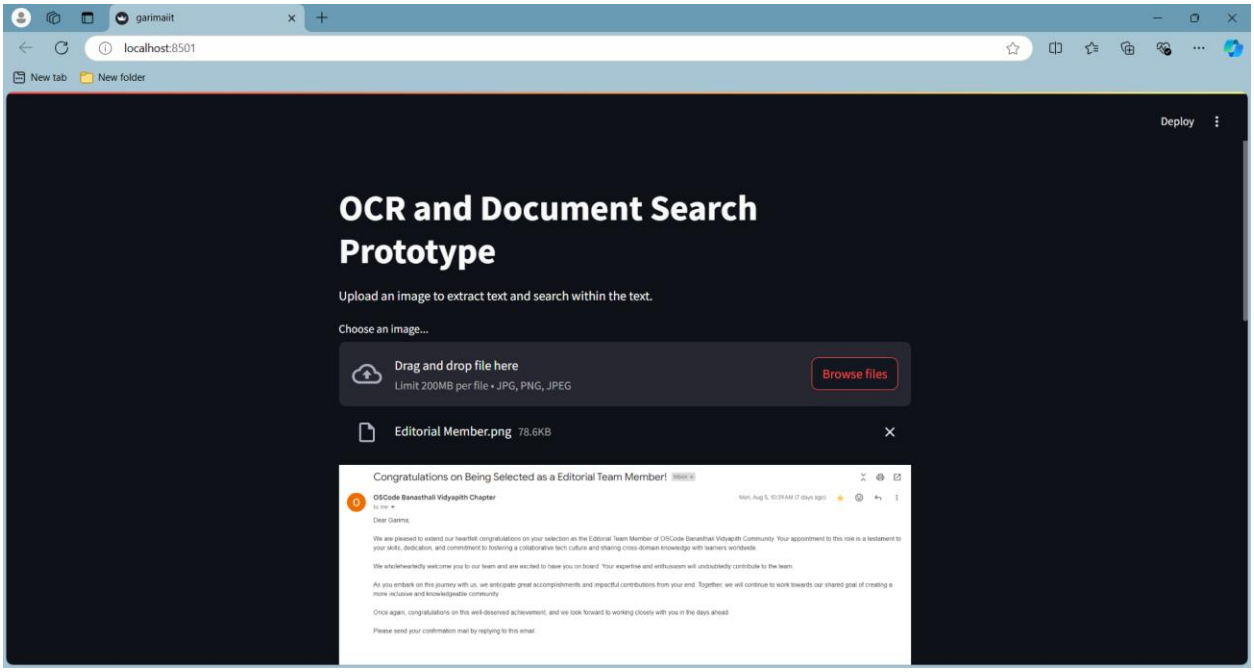


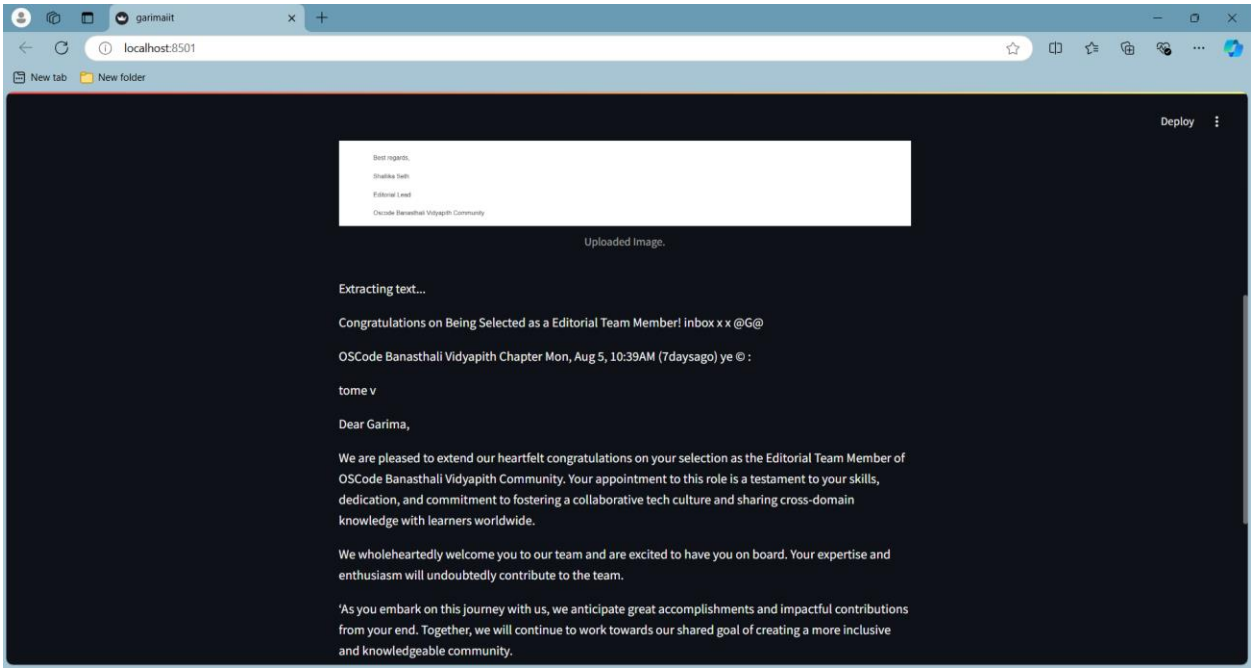
WORD SEARCHER IN HINDI



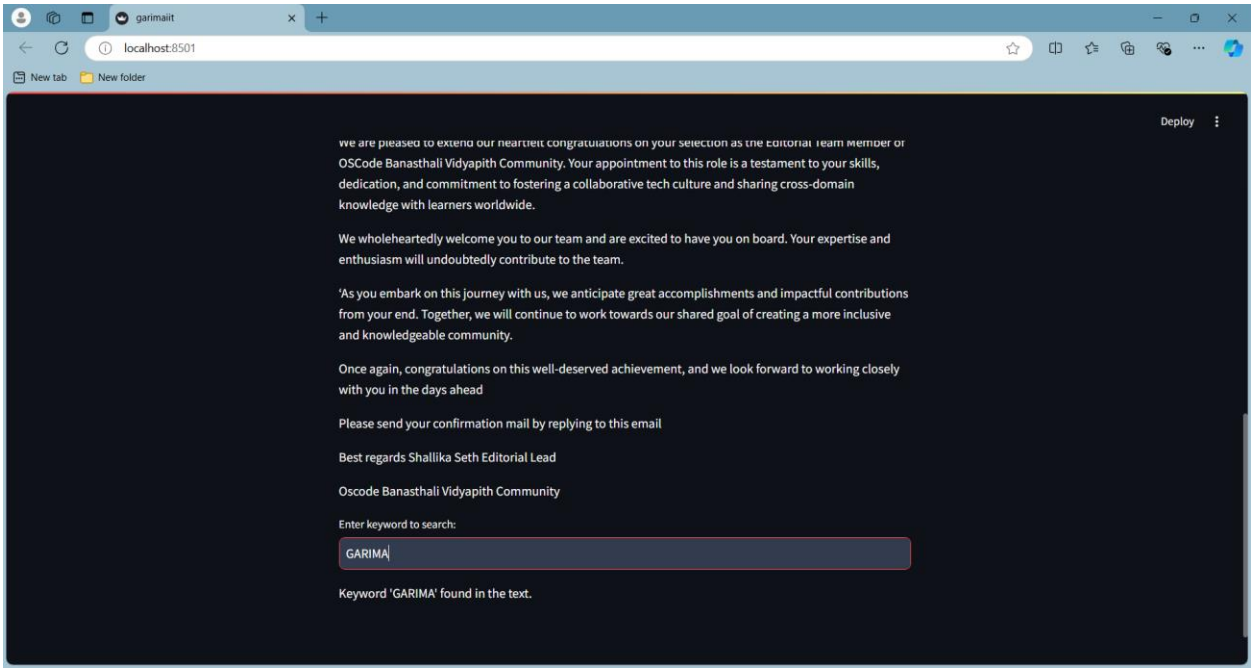


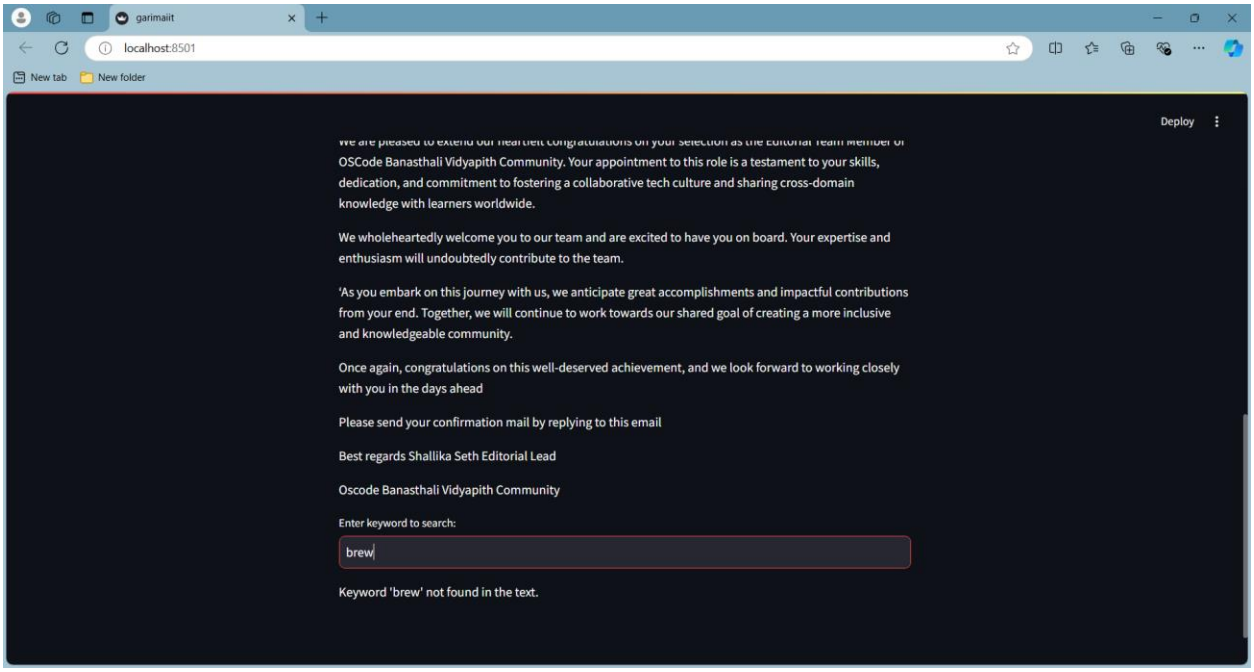
ENGLISH WORD EXTRACTOR





WORD FINDER IN ENGLISH





OCR web application is designed to read text from images in both Hindi and English. It allows users to upload an image, extracts the text using Optical Character Recognition (OCR), and provides a keyword search functionality to find specific words within the extracted text. Key Features:

Image Upload: Users can upload an image containing text.

Text Extraction: The app extracts text from the image using Tesseract OCR.

Keyword Search: Users can search for keywords within the extracted text, and the app highlights the matching sections.