Technical Assessment

Objective:
This assignment is designed to assess your ability to **extract structured data from documents using Python**, **organize it in JSON format**, and **develop a React.js application to present the extracted data as a structured webpage**.

We encourage you to use AI tools, online resources, and documentation. However, during the interview, you should be able to **explain your thought process, design choices, and implementation details.**

Deadline: Wednesday, 19 March 2025, 11:59 PM

Task 1: Data Extraction & JSON Conversion (Python)

Goal: Extract structured data from a sample Word document and convert it into a logically structured JSON format.

Requirements:

- Write a Python script to extract content from data_001.docx (a sample Word file).
- Extract and structure the following elements:
  - **Headers** (section titles, sub-titles)
  - **Text Content** (paragraphs and descriptions)
  - **Tables** (row & column values)
- Organize the extracted data into a **well-structured JSON format**.

Deliverables:

1. **Python script** (extract_data.py) that reads data_001.docx and generates a JSON file.
2. **JSON output file** (output.json) containing extracted and structured data.

Task 2: Build a Web Page to Display Extracted Data (React.js)

Goal: Develop a **React.js application** to present the extracted data in a structured and user-friendly web page.

Requirements:

- Create a **React.js** project.
- Load the extracted JSON file (output.json).
- **Design a webpage** that displays:
    - **Headers & Content** in an organized format.
    - **Tables** in a properly formatted layout for readability.
- Ensure a clean and structured layout (styling is up to you, using **Tailwind, Material-UI, or basic CSS**).
- Allow users to navigate or switch between different sections of the document dynamically.

Deliverables:

1. **A React.js project** that displays the extracted data as a well-structured webpage.
2. **Code hosted on GitHub** (or share as a zip file).

Evaluation Criteria

- Logical structuring of JSON output
- Code readability and organization
- Webpage layout, design, and content presentation
- Ability to explain the thought process and approach
- Efficient use of Python for data extraction and React for frontend implementation

---

Submission Instructions

1. Upload your **Python script and output JSON file** to a public GitHub repository (or share a zip file).
2. Upload your **React.js project** to GitHub (or share the source files).
3. Send your **GitHub repo link / Zip file** via email before the deadline.

Deadline: Wednesday, 19 March 2025, 11:59 PM

We look forward to reviewing your submission and discussing your approach in the interview.

Good luck!