

3. Chunking Function for Extracted Data

Task: Split extracted text into smaller overlapping chunks.

Explanation:

Chunking ensures large text data is broken into manageable parts for processing, maintaining context across boundaries.

- **Steps:**
 1. Divide the text into smaller overlapping sections.
 2. Use overlap to preserve context between chunks.
- **Tools:** Use `langchain.text_splitter.RecursiveCharacterTextSplitter`.
- **Sample Code Snippet:**

python

Copy code

```
from langchain.text_splitter import RecursiveCharacterTextSplitter
```

```
def chunk_text(text, chunk_size=500, overlap=50):
```

```
    splitter = RecursiveCharacterTextSplitter(chunk_size, overlap)
```

```
    chunks = splitter.split_text(text)
```

```
    return chunks
```