**Download Repertoires and Metadata:**

**Overview**

This Python script facilitates downloading repertoires and metadata from online repositories. It provides an interactive command-line interface for repository selection, download directory specification, and study ID entry. The script employs various Python libraries for asynchronous processing, networking, and data management. Once data is downloaded, **create\_projects\_structure.py** and **json\_to\_tsv.py** are called to structure the data into a coherent format and convert JSON formatted data into TSV files, respectively.

**Script Structure**

The script is composed of multiple functions and classes, each designed for specific tasks:

1. **Collecting Repertoires (collect\_repertoires)**: Queries repertoires from repositories using study IDs, sending requests to each repository and gathering results.
2. **Counting Rearrangements (count\_rearrangements)**: Asynchronously counts the rearrangements in the collected repertoires to improve performance.
3. **Main Function (collect\_repertoires\_and\_count\_rearrangements)**: Merges repertoire collection and rearrangement counting, offering summary statistics of the downloaded data.
4. **Downloading Repertoires (BatchDownloader)**: Manages the asynchronous download of repertoires, allowing multiple concurrent downloads with a set limit.
5. **Creating Project Structure (create\_projects\_structure.py)**: After downloading, this script organizes the genomic data files into a structured directory format.
6. **Converting JSON to TSV (json\_to\_tsv.py)**: Converts the downloaded JSON formatted data into TSV files, adhering to specific genomic data analysis formats.

**Usage**

1. Run the script: **python download\_repertoires\_and\_metadata.py**.
2. Input the download directory and study ID.
3. Select repositories for searching.
4. Initiate the download process.
5. Post-download, the **create\_projects\_structure.py** script will organize the data.
6. The **json\_to\_tsv.py** script will then convert the data into TSV format.
7. Repeat for additional study IDs or exit with "exit."

**Customization**

* **Repositories**: Customize the list of default repository URLs in the **default\_repository\_df** DataFrame to include personal repositories or modify existing ones.