**Elasticsearch Data Upload**

**A screenshot of a computer

AI-generated content may be incorrect.**

Enter the elastic search credential to give access for data extraction

**User Guide: CSV to Elasticsearch Uploader**

**Overview**

**This application allows users to upload CSV data files to an Elasticsearch index using Streamlit. Users can define schema mappings, convert date columns to epoch format, and handle bulk uploads.**

**Features**

* **Upload schema and data CSV files**
* **Define or modify schema mappings**
* **Automatic date format detection and conversion to epoch**
* **Specify or generate unique document IDs**
* **Bulk upload data to Elasticsearch**
* **View failed records and errors**

**Step-by-Step Instructions**

**1. Launch the Application**

**Run the Streamlit application using the following command:**

**streamlit run csv\_to\_esindex.py**

**2. Enter Elasticsearch Credentials**

* **Elasticsearch Host: Provide the host URL (e.g., http://localhost:9200).**
* **Username & Password: Enter your Elasticsearch authentication credentials.**

**3. Upload Files**

* **Schema CSV: Upload a schema file defining column names, data types, and optional formats.**
* **Data CSV: Upload the actual data file to be indexed.**

**4. Verify and Edit Schema**

* **The extracted schema will be displayed for review.**
* **You can modify it manually in the provided text editor.**
* **Ensure all data types are valid for Elasticsearch.**

**5. Convert Date Columns to Epoch**

* **The application automatically detects date formats and converts them to epoch time.**
* **Converted columns are appended with \_asEpoch and added to the schema.**

**6. Define Index and ID Settings**

* **Enter the Elasticsearch index name where data will be stored.**
* **Choose an \_id column from the dataset or let the system generate unique IDs.**
* **If an ID column is selected, ensure it has no duplicates or null values.**

**7. Create Index in Elasticsearch**

* **Click the "Create Index in Elasticsearch" button.**
* **The system will create the index with the defined schema mappings.**

**8. Upload Data to Elasticsearch**

* **Click "Upload Data to Elasticsearch".**
* **The application will process records and attempt bulk indexing.**
* **If any records fail, they will be displayed with error details.**

**9. View and Download Indexed Data**

* **Click "Download Data from Elasticsearch" to retrieve stored data as a CSV file.**

**Error Handling & Debugging**

* **The application will notify you of incorrect schema formats or invalid data types.**
* **If a date format is incorrect, check the inferred format and adjust accordingly.**
* **Failed document uploads will be listed along with the reason for failure.**
* **If a bulk upload fails, check the Elasticsearch connection and credentials.**

**Troubleshooting Tips**

* **Ensure Elasticsearch is running before using the application.**
* **If the connection times out, increase the timeout value in the code.**
* **If errors persist, check logs for detailed debugging information.**

**This guide ensures a smooth user experience for uploading CSV data to Elasticsearch using Streamlit.**

**Schema File Format for CSV to Elasticsearch Uploader**

**The schema file should be a CSV file containing the column definitions for your dataset. It should have the following structure:**

| **column name** | **data type** | **format** |
| --- | --- | --- |
| **ID** | **keyword** |  |
| **Name** | **text** |  |
| **Age** | **integer** |  |
| **Salary** | **float** |  |
| **JoinDate** | **date** | **yyyy-MM-dd** |
| **BirthDate** | **date** | **dd-MM-yyyy** |

**Explanation of Columns:**

* **column name: The name of the column in your dataset.**
* **data type: The corresponding Elasticsearch data type (must be one of: text, keyword, integer, long, float, double, boolean, date).**
* **format (optional): Required only for date columns to specify the date format (e.g., yyyy-MM-dd, dd-MM-yyyy, MM/dd/yyyy, etc.).**