### **Story Map Description**

### Configure input sources from provided file

* connect RDBMS/S3 storages - When I need to configure data input sources, I want to connect RDBMS/S3 storages, so I can use multiple types of data storage for my ETL process.
* change input sources - When I need to adjust the data inputs, I want to change the input sources, so I can work with the most relevant and up-to-date data.
* add sources to existing - When I have new data sources, I want to add sources to the existing configuration, so I can expand the scope of my ETL process.

### Configure ETL process from provided file

* exclude stage from process - When a specific stage in the ETL process is unnecessary, I want to exclude that stage, so I can optimize the processing time.
* define data structure - When I set up the ETL process, I want to define the data structure, so the system knows how to process and store the data accurately.
* define stages sequence - When I organize the ETL pipeline, I want to define the sequence of stages, so data flows in the correct order through the process.

### Extract data

* compare data from multiple sources - When I extract data from various sources, I want to compare data from multiple sources, so I can identify inconsistencies and ensure data accuracy.
* bring data to a unified format - When the data is extracted, I want to bring it to a unified format, so it can be processed and analyzed consistently.

### Cleanse data

* remove nullable - When I cleanse the data, I want to remove nullable data, so I eliminate incomplete records and improve data quality.
* apply cleansing rules - When I need to prepare the data, I want to apply predefined cleansing rules, so the data meets the required standards for processing.
* remove emissions - When the dataset includes anomalies, I want to remove emissions, so the processed data reflects accurate information.

### Transform data

* apply transformation rules - When I need to adjust data, I want to apply transformation rules, so the data fits the target format for the next stages of processing.
* edit transformation rules - When transformation requirements change, I want to edit transformation rules, so the system applies the correct data modifications.

### Anonymize data

* replace personally identifiable information - When I need to ensure data privacy, I want to replace personally identifiable information, so sensitive data is protected during analysis.
* replace personally identifiable data with pseudonyms - When the data must be anonymized, I want to replace personally identifiable data with pseudonyms, so I can work with the data without compromising privacy.

### Merge data

* put data to module storage - When I need to store processed data, I want to put data into module storage, so it is available for further analysis and processing.
* ensure consistency in the merged data - When merging data from multiple sources, I want to ensure consistency in the merged data, so the data remains accurate and reliable.

### Exporting dataset

* choose dataset format extraction - When I export a dataset, I want to choose the format for extraction, so the data is compatible with external systems.
* load final dataset - When I complete the ETL process, I want to load the final dataset, so the data can be used for reporting and further analysis.

### Include stages description

* logging for each process stage - When the ETL process runs, I want to log each process stage, so I have a detailed record for debugging and auditing purposes.
* show stage progress - When the ETL process is running, I want to show the stage progress, so I can monitor the system's status in real-time.

### Load previous dataset

* choose dataset to load - When I need to work with historical data, I want to choose the dataset to load, so I can access and analyze past information.
* get dataset info - When I load a dataset, I want to get dataset info, so I can understand the structure and contents of the data.

### Form an overall system report

* get a summary report - When I want a quick overview of the ETL process, I want to get a summary report, so I can understand system performance at a glance.
* get a detailed report - When I need deeper insights, I want to get a detailed report, so I can review the system’s operations and identify any issues.

### Manage permissions

* restrict access - When I need to control data access, I want to restrict access, so only authorized users can view or modify the data.
* view possible system permissions - When I manage user roles, I want to view possible system permissions, so I can assign the correct level of access to each user.