

Story map

ETL servise



Needs

- Easy configure ETL system
- Fast integration of ETL Service into existing projects
- Fastly get Dataset
- Have a practical permission system
- Convenient Dataset Storage
- Simple monitoring system



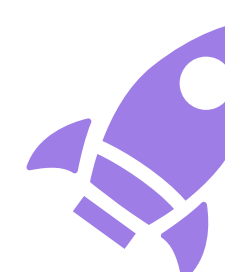
JOBS

- Cofnigure inputed sources
- Configure ETL process stages
- Execute ETL process
- Get final Dataset
- Manage role model of the project
- Access previous Datasets
- Conduct process description
- Store Dataset in module



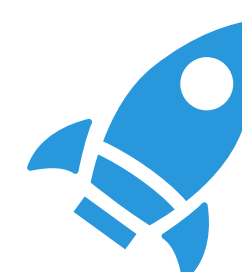
Features

- Configure input sources from provided file
 - add sources to existing
 - change input sources
 - connect RDBMS/S3 storages
- Configure ETL process from provided file
 - define data structure
 - exclude stage from process
 - define stages sequence



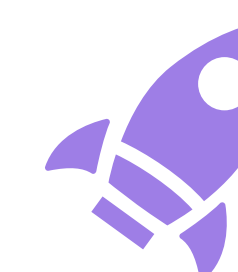
Jobs stories

- Cleanse data
 - remove nullubale
 - apply cleansing rules
 - remove emissions
- Extract data
 - compare data from multiple sources
 - bring data to a unified format



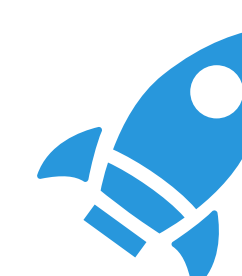
Features

- Transform data
 - apply transformation rules
 - edit transformation rules
- Anonymize data
 - replace personally identifiable information
 - replace personally identifiable data with pseudonyms



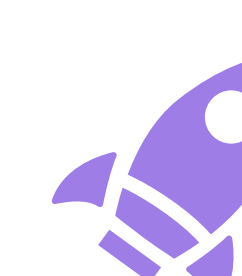
Jobs stories

- Merge data
 - put data to module storage
 - ensure consistency in the merged data
- Exporting dataset
 - choose dataset format extraction
 - load final dataset



Features

- Include stages description
 - logging for each process stage
 - show stage progress
- Load previous dataset
 - get dataset info
 - choose dataset to load



Jobs stories

- Form an overall system report
 - get a summary report
 - get a detailed report
- Manage permissions
 - view possible system permissions
 - restrict access