

# IDBS Access-Rights

ELN API • Oracle Views • Datalab Talend GCP BigQuery

<b>Project Name</b>	Each ELN product follows a distinct IDBS access rights workflow
<b>ETL</b>	Talend
<b>Status</b>	Deployed in PROD

## 1. Project Overview

### Objectives

The ELN system hosts multiple independent research projects. The Talend ETL flow enforces a project-level access policy: each collaborator is assigned to one or more specific ELN projects, and the pipeline only extracts and loads the data belonging to that collaborator's assigned projects.

### Scope

- All active Talend jobs connecting to ELN API, Oracle Views, and DataLab.
- Access management rules governing ELN products assignment per collaborator.
- GCP BigQuery datasets and tables that serve as the final landing zone.

## 2. Architecture & Data Flow

The ETL pipeline follows a classic ETL pattern orchestrated by Talend. The high-level data flow is described below:

[ELN API]

[Oracle Views ] [Talend ETL Jobs] [GCP BigQuery Tables]

[Datalab Platform ]

### Flow Steps

- **Extract:** Talend connects to each source (ELN REST API, Oracle JDBC, Datalab (GCP) ) and extracts raw data according to the job configuration.
- **Filter / Route:** For ELN sources, the flow applies a project-based access filter, ensuring each collaborator's data is scoped to their assigned ELN project(s) only.
- **Transform:** Data is cleaned, typed, and mapped to the target GCP schema within the Talend job components.
- **Load:** Transformed data is written to the corresponding BigQuery dataset and table using the GCP BigQuery connector.

## 3. Source Systems

Source System	Type	Description
ELN API	REST API	Electronic Lab Notebook – exposes project data per collaborator access level
Oracle Views	Relational DB Views	Pre-aggregated relational data extracted via JDBC connections
Datalab Platform	Internal Data Platform	Analytical datasets and processed outputs from the Datalab environment

### 3.1 ELN API

#### Authentication:

(TEST)

- **REST / HTTP:** Protocol
- **Authentication Type :** Basic HTTP
- **URL:** <https://mylab-test.syensqo.com:8443>
- **Login:** (Contact Admin)
- **Password:** (Contact Admin)

(PROD)

- **REST / HTTP:** Protocol
- **Authentication Type :** Basic HTTP
- **URL:** <https://mylab.syensqo.com:8443>
- **Login:** (Contact Admin)
- **Password:** (Contact Admin)

**Note :**

- Get the list of container IDs from the Oracle view, then iterate over it as an input variable in the URL.
- Container IDs are should be available in this view :IDBS\_EWB\_SEC.HISTORY\_ENTRIES.ENTITY\_PATH

**Example :**

- <https://mylab-test.syensqo.com:8443/ewb/services/1.0/feeds/entities/> (**ContainerID\_List**)

## 3.2 Oracle DB

**Authentication:**

(TEST)

- IDBS\_DB\_Database = [ELNTEST.eua.solvay.com](http://ELNTEST.eua.solvay.com)
- IDBS\_DB\_Host = [ACEW1TELNBDDB02.eua.solvay.com](http://ACEW1TELNBDDB02.eua.solvay.com)
- IDBS\_DB\_Port = 1521
- IDBS\_DB\_UserName = (Contact Admin)
- IDBS\_DB\_Password = (Contact Admin)

(PROD)

- IDBS\_DB\_Database = [ELNPROD.eua.solvay.com](http://ELNPROD.eua.solvay.com)
- IDBS\_DB\_Host = [ACEW1TELNBDDB02.eua.solvay.com](http://ACEW1TELNBDDB02.eua.solvay.com)
- IDBS\_DB\_Port = 1521
- IDBS\_DB\_UserName = (Contact Admin)
- IDBS\_DB\_Password = (Contact Admin)

**View:**

- IDBS\_EWB\_SEC.HISTORY\_ENTRIES.TARGET\_ENTITY\_ID AS Experiment\_ID
- IDBS\_EWB\_SEC.HISTORY\_ENTRIES.ENTITY\_PATH AS FOLDERPATH
- IDBS\_EWB\_SOLVAY\_CUSTOM.DM\_UserAccessPath\_Full AS PERMISSIONS
- DBS\_EWB\_SOLVAY\_CUSTOM.DM\_EntityRolePermissions\_Full AS ROLES

**Note:**

- The Oracle database connection is enabled for Talend Remote Engines. However, when connecting to the database via a VDI, the connection may be disabled.

## 3.3 Datalab

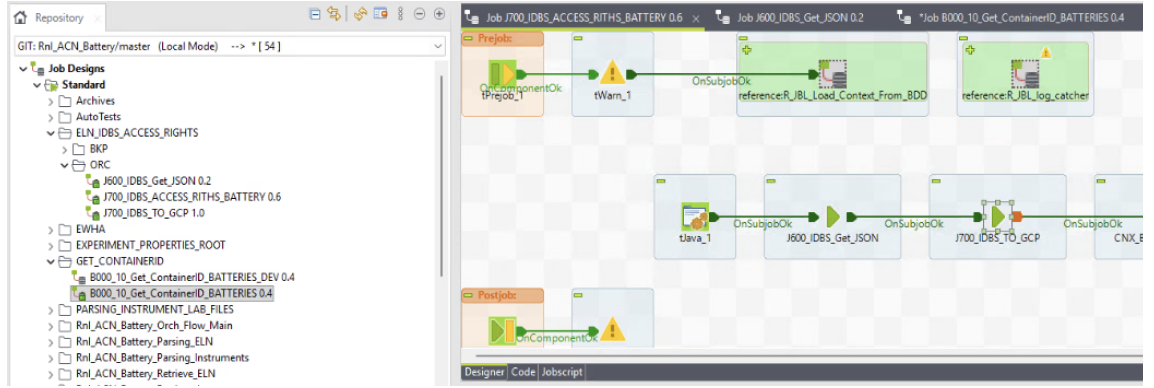
- **Source:** GCP Bigquery table
- **Project:** gcp-sqo-datalab-(\*)
- **Dataset:** bq\_ds\_datagrow\_dev\_ads\_static
- **Table:** application\_use
- **filter:** user\_status='active' AND ROLE IN ('admin', 'lab\_manager')

## 3.4 Target table:

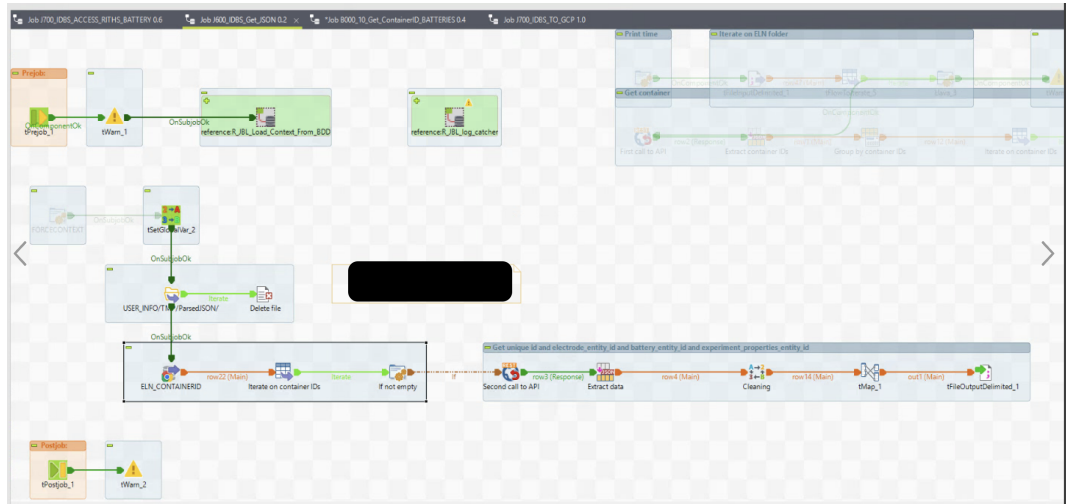
Bigquery: DM.ELN\_IDBS\_AccessRights

### Talend tasks:

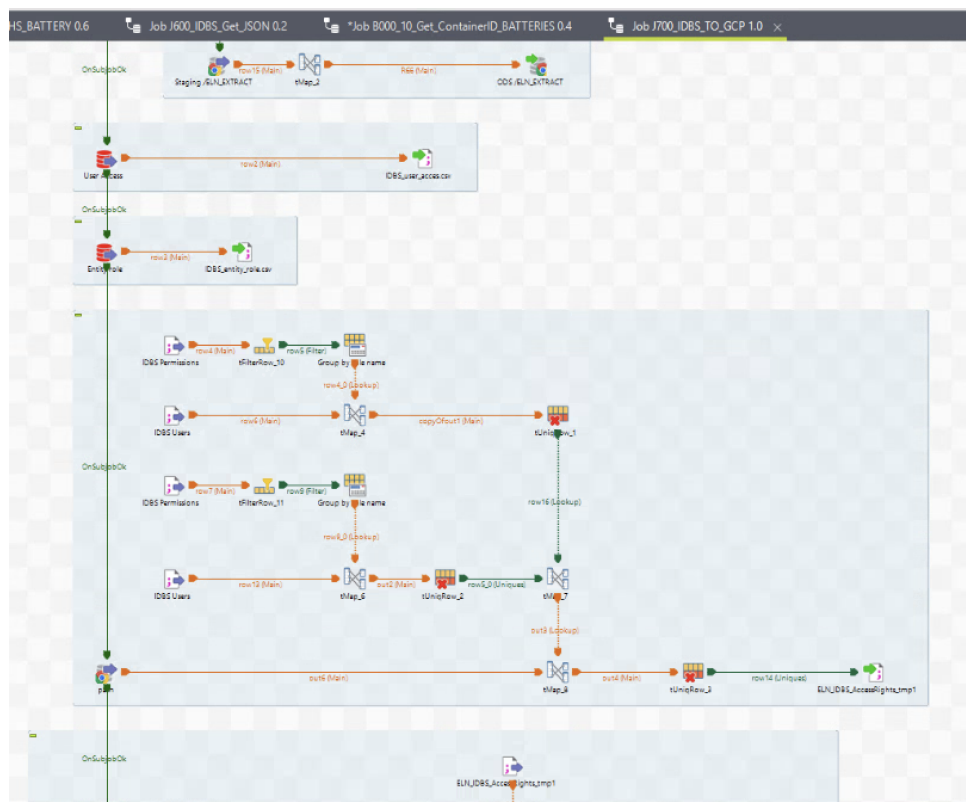
(Master)



(Get Experiments)



(Cross experiments with collaborators and roles )



## Error Handling Strategy

- All jobs implement a tDie component on the error branch to terminate gracefully and log a meaningful message.
- API connectivity failures trigger a tWarn followed by a retry loop.
- Job execution logs are visible in TMC.

## Scheduling

- ELN\_GCP job: Daily at 02:00 AM (server local time)
- Oracle\_GCP job: Daily at 02:00 AM
- Datalab\_GCP job: Daily at 02:00 AM

## 4. Contacts & responsibilities:

- [antoine.rey@syensqo.com](mailto:antoine.rey@syensqo.com) / ELN Administrator
- [prasanth.gnanasekar@syensqo.com](mailto:prasanth.gnanasekar@syensqo.com) / Data Engineering - Flow Maintenance

## 5. Annexe document:

This annex document could be helpful if you are looking for more details on access rights related to ELN. ([link](#))