

ALB Data Dev for Materials

This page presents the data development documentation for Materials data for the sites of Alpharetta and Bollate.

Summary

- [Summary](#)
- [Description](#)
- [Data Ingestion](#)
- [Data Preparation or Parsing](#)
- [Data Integration or Computing](#)
- [Data Presentation](#)
- [Data Visualization](#)

Description

Data Ingestion

The data ingestion phase for Materials follows the standard approach describe on [ALB Data Dev Architecture - General](#) except for...

Data Sources

Labware

we retrieve data from an oracle database

Related documents

Document Name	Link
Data Model for Material	https://docs.google.com/document/d/1j3Fe-qBIRE9UDvkSdh9ly7BvRPWnR2PmSt96VZFKiUY/edit?usp=sharing

Instruments

The instrument files are, most of them, **manually** added to a Google Driver's folder or a folder in lab servers.

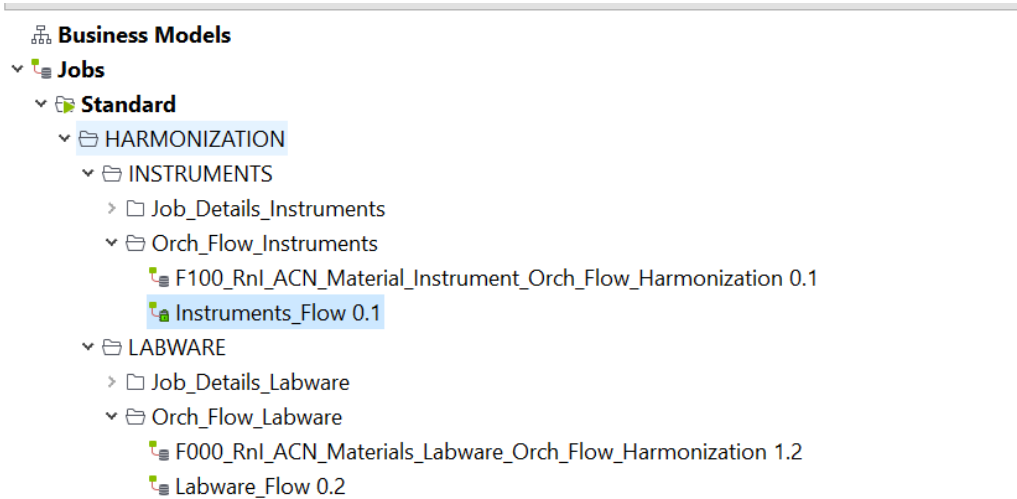
The spreadsheet **Where to find instrument files?** has the full list of instrument folders and location:

Data Mapping

No data mapping needed as the data is copied exactly as the source.

Talend Jobs

The jobs Instruments_Flow and Labware_Flow are responsables for orchestrating the Data Ingestion:



Data Preparation or Parsing

The data preparation phase for Materials follows the standard approach describe on [ALB Data Dev Architecture - General](#).

Data Integration or Computing

The data integration phase for Materials follows the standard approach describe on [ALB Data Dev Architecture - General](#).

Data Mapping

The spreadsheet below presents all data transformations between the tables on Staging and ODS. This steps aims to add some calculations and intelligence to the data. Not all tables will need passing through this step.

<https://docs.google.com/spreadsheets/d/1tjV8I42BkyFW48-SIUIQZQdDpnaYbtYq/edit?usp=sharing&ouid=117421811921505411556&rtfpof=true&sd=true>

Data Presentation

The data presentation phase for Materials follows the standard approach describe on [ALB Data Dev Architecture - General](#).

Talend Jobs

No need of jobs as there are no steps to load DW/DM. A priori, all data is presented as views.

Orchestrating Jobs

All the jobs are run in sequence under the follow job and project name on TAC/Talend Cloud:

Project	Job/Flow
RnI_ACN_Materials	F000_RnI_ACN_Materials_Orch_Flow

For scheduling details check the Operational documentation.

Folders

Nom	Modifié le	Type	Taille
Instruments	23/05/2022 17:51	Dossier de fichiers	
labware	31/05/2022 18:00	Dossier de fichiers	
Output	11/01/2022 15:58	Dossier de fichiers	
Working	23/05/2022 18:32	Dossier de fichiers	

Tables (Staging)

Staging	
labware_conso	
labware_delta	
labware_testcomponent_conso	
labware_testcomponent_delta	
method_description_conso	
method_description_delta	
raw_data_conso	
raw_data_delta	
sample_information_conso	
sample_information_delta	
summary_results_conso	
summary_results_delta	

Data Visualization