

# ALB Data Dev for Coatings

This page presents the data development documentation for Coatings data for the sites of ...

## Summary

- [Summary](#)
- [Description](#)
- [ELN](#)
- [Data Ingestion](#)
- [Data Preparation or Parsing](#)
- [Data Presentation](#)
- [Orchestrating Jobs](#)
- [Specific Naming Conventions](#)
- [Data Visualization](#)

## Description

## ELN

## Data Ingestion

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

## Data Sources

### ELN

The list of the spreadsheets extracted from JSON files coming from ELN can be found in the Data Mapping of the next section.

### Related documents

#### Related documents

Document Name	Link
Data Model for Coatings	<a href="#">Data Model Link</a>
New Data Model for Coatings	<a href="#">Data Model Link</a>

## Instruments

The instrument files are, most of them, **manually** added to a folder in lab servers.

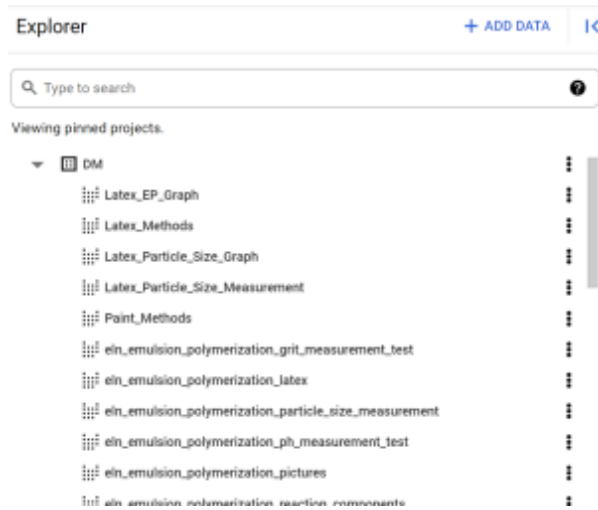
The spreadsheet **Where to find instrument files?** has the full list of instrument folders and location:

```
\\frph2-labpc-backup\labo\W-513825\DATAS\Experiment Results
\\frph2-labpc-backup\LABO\W-510315\Malvern Instruments\Zetasizer\Measurement Data\Coatings\Export Data - Coatings
\\USBRIFS51.ist-priv.rhodia.com\data$\USBRI30293\Malvern Instruments\Zetasizer\Export Data\Coatings
\\USBRIFS51.ist-priv.rhodia.com\data$\USBRI28321\Attension Instrument\Surface Tension\Export Data\Coatings
```

## Data Mapping

## Talend Jobs

The jobs F001, F000 and F100 are responsible for orchestrating the Data Ingestion:



## Data Preparation or Parsing

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

## Data Mapping

The spreadsheet below presents all data transformation between the raw files (extracted files) and a BigQuery delta table. Some files are unstructured and semi-structured. This steps aims to structure the files in the target table format and checking if the column's type (schema) are conformed expected.

## Data Presentation

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

## Data Mapping

No data mapping available, as there are no transformations for now. This must be created if there are requirements for that.

## Talend Jobs

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

## Data Model

The data model presents the tables/views presented on DW/DM dataset and the relation between them.

{Add model}

## 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_Coatings	F001_RnI_ACN_Coatings_ELN_EP_OrchFlow

Rnl_ACN_Coatings	F000_Rnl_ACN_Coatings_ELN_Paint_Formulation_Orch_Flow
Rnl_ACN_Coatings	F100_Rnl_ACN_Coatings_Instruments_Orch_Flow

For scheduling details check the Operational documentation.

## Specific Naming Conventions

### Table Names

### Tables (Staging)

The screenshot shows the Tableau Explorer interface. At the top, it says 'Explorer' with a '+ ADD DATA' button and a back arrow. Below that is a search bar with the placeholder text 'Type to search'. Underneath, it says 'Viewing pinned projects.' and shows a tree view under the 'DM' project. The tables listed are: Latex\_EP\_Graph, Latex\_Methods, Latex\_Particle\_Size\_Graph, Latex\_Particle\_Size\_Measurement, Paint\_Methods, ein\_emulsion\_polymerization\_grit\_measurement\_test, ein\_emulsion\_polymerization\_latex, ein\_emulsion\_polymerization\_particle\_size\_measurement, ein\_emulsion\_polymerization\_ph\_measurement\_test, ein\_emulsion\_polymerization\_pictures, and ein\_emulsion\_polymerization\_reaction\_components.

### Tables (DM)

This screenshot is identical to the one above, showing the same Tableau Explorer interface with the same list of tables under the 'DM' project.

## Data Visualization

Tableau workbook documentation : [Technical documentation](#)