2025-08-11 Meeting Agenda

📅 Date

11 Aug 2025

👥 Participants

- RAIL PG-2 project team
- Murtaza (Proxy Client)

🥅 Goals

- Establish a common understanding of the project deliverables
- Confirm communication and tools
- Discuss the initial backlog

🗣 Discussion topics

1. **Project Overview**
   - data pipeline + ML model to predict rail breaks using ARTC dataset
   - 4 phases? Data ingestion → Processing & Cleaning → Modeling → Output/report

2. **Data & Tools**
   - Data sources: rail break history + feature datasets → any data quality issues?
   - How to access the resources (Dataset & Insight Factory platform / Azure Databricks)?
   - Local hardware requirements? Can we suppose everything runs on cloud?
   - Recommendations on tools/libraries/ML methods?
3. **Roles & Responsibilities**
   - High-level roles (ingestion, processing, modeling, output/reporting)
4. **Architecture**
   - draft architecture (4 layers: ingestion → processing → ML → output)
   - preferred architecture pattern?
5. **Sprint 1 Goals**
   - short-term priorities (env setup, data access, documentation, backlog creation)