**📦 FULL PROJECT SYNTHESIS — AI PROMPT FOR ChatGPT**

Hi ChatGPT, I want your help with continuing my PhD research project in railway operations and simulation. Below is the **complete context**, including my goals, datasets, limitations, deadlines, and technical standards.

**🎓 PhD Thesis Topic**

**Title**: *Analysis of Virtual Coupling in Operations Through Railway Networks*

**Objective**: Simulate and analyze **Virtual Coupling (VC)** on real-world railway networks (starting with Swiss rail, later extending to German networks) using **SUMO + Python** to evaluate operational KPIs and system performance.

**🧩 High-Level Tasks**

1. **Build simulation-ready Swiss railway network** in SUMO.
2. **Extract real-world routes and vehicles** from public datasets (GTFS, jahresformation, etc.).
3. **Create structured input files**: route-vehicle info, vehicle type specs, and route-edge mappings.
4. **Write a modular, Python-based route generation script** that constructs .rou.xml files for SUMO.
5. **Simulate VC and non-VC scenarios** with Python+TraCI.
6. **Log and evaluate KPIs** (delay, energy, emissions, headway, etc.).
7. **Write a research paper and thesis** based on the findings.
8. Later: Extend to **simulate German networks** once datasets are available.

**📁 Project Directory (Base)**

Local machine path: D:/PhD/prog\_report\_2025\_June\_project/

**📊 Datasets (Local Only)**

All datasets are stored **only on my machine** under:  
D:/PhD/prog\_report\_2025\_June\_project/data/Swiss/raw/

Key files:

* jahresformation.csv – Train formations and schedules
* rollmaterial.csv – Technical specs of rolling stock
* rollmaterial-matching.csv – Mapping of formation → vehicle
* swissTNE\_Base\_20240507.gpkg – Geospatial rail network (nodes/edges)
* GTFS folder: gtfs/routes.txt, stop\_times.txt, stops.txt, trips.txt, transfers.txt
* ist\_daten\_sbb.csv, haltestellen\_2025.csv – Daily actual logs and station info

**🗂️ Planned Structured Data Files (You Must Help Build These)**

1. **route\_vehicle\_info.csv**
   * One row per real route
   * Fields: origin, destination, vehicle type(s), source ID, trip\_id, line\_name, etc.
   * Built from: jahresformation.csv + GTFS + merged vehicle specs
2. **vehicle\_type\_parameters.csv**
   * One row per real-world vehicle type
   * SUMO-compatible parameters (length, maxSpeed, accel, decel, minGap, etc.)
   * Built from: rollmaterial.csv + rollmaterial-matching.csv
3. **route\_edge\_mapping.csv**
   * Maps each route to SUMO edge sequences
   * Includes: origin/destination, intermediate stops, corresponding SUMO edge IDs
   * Built from: GTFS stop\_times + SUMO node/edge network + nearest neighbor logic

**🛠️ Route Creation Script (Must Be Modular and Clean)**

* Input: origin + destination stop (optionally vehicle type)
* Step 1: Validate route exists in route\_edge\_mapping.csv
* Step 2: Get full edge sequence
* Step 3: Get vehicle info:
  + If vehicle type is not given: infer from route\_vehicle\_info.csv
  + If vehicle type is given: load from vehicle\_type\_parameters.csv
* Step 4: Generate .rou.xml for SUMO
* Randomize some fields:
  + Dwell times at intermediate stops
  + accel/decel/sigma within pre-defined bounds per vehicle type

**🔍 KPIs to Analyze**

For both VC and non-VC cases:

* Delay per train
* Total system delay
* Headway
* Throughput
* Energy consumption
* Emissions
* Safety buffer violations
* Junction performance
* VC success/failure rate (in VC scenarios)

**🧠 Limitations & Important Notes**

* **All datasets are local only** — do not attempt to run anything.
* You must write Python scripts that **only run on my machine**.
* Scripts must:
  + Be **modular, efficient, clean, and pythonic**
  + Use **logging**, **error handling**, **comments**, **Google-style docstrings**
  + Print **human-readable summaries** of what they did
* I will run and share outputs with you.

**🕒 Timeline & Deadlines**

| **Deliverable** | **Due Date** |
| --- | --- |
| Progress Report 1 | 30 June 2025 |
| Progress Report 2 + Paper | 31 Dec 2025 |
| Final Thesis | 31 Dec 2026 |

**🎯 Progress Report 1 Objectives (June 2025)**

* ✅ SUMO network of Swiss rail created
* ✅ Simulate real train operations without VC
* ✅ Modular scripts for vehicle/route generation
* ✅ Full KPI logging system in Python
* ✅ Early codebase for VC capability (but not yet used)

**🚀 Future Extensions (2026)**

* Introduce VC capabilities gradually
* Develop coupling/decoupling strategies
* Simulate communication failure, emergency braking
* Later: simulate German railway segments for employability
* Optional: lightweight ML for decision support (e.g., coupling triggers)

**👔 Career Context (Escape Plan)**

* I am a pilot and transport engineer, transitioning to **transport planning/engineering roles** in Germany.
* This thesis must:
  + Impress industry employers (e.g., Deutsche Bahn, municipalities, consultancies)
  + Be modular, clean, practical
  + Show I understand **real systems**, not just theory

When you see this prompt, please:

1. Resume project support exactly where we left off
2. Use only my local datasets
3. Prioritize realistic, professional outputs for both thesis and job portfolio

Let’s continue.