Capstone Project Proposal

Hierarchical Multi-Agent Reinforcement Learning Platform for Urban Synthetic Data Generation

January 19, 2026

1 Background & Motivation

1.1 Problem Statement

The training of modern Artificial Intelligence (AI) models relies heavily on massive, high-quality datasets. However, in domains involving complex real-world urban operations—such as transportation, economics, and public health—acquiring such data faces three critical bottlenecks:

- **Data Incompleteness and Latency:** Traditional census data or sensor readings often suffer from significant reporting delays and lack real-time granularity.
- **Privacy and Regulatory Constraints:** Real-world data containing Personally Identifiable Information (PII) cannot be directly released for training or testing due to strict privacy regulations (e.g., GDPR).
- **Scarcity of Extreme Scenario Data:** It is impossible to conduct destructive tests in the real world (e.g., inducing an economic crash, extreme policy shifts, or natural disasters). Consequently, models often lack robustness when facing "Long-tail events."

1.2 Proposed Solution

This project aims to build a **"Virtual Urban Simulation Environment."** By constructing a hierarchy of AI Agents—ranging from government decision-makers to individual citizens—we will simulate the intricate operating mechanisms of a real city.

Unlike recreational games (e.g., *SimCity*), the primary objective of this platform is to **generate realistic, safe, and structured Synthetic Data**. This platform will serve researchers and enterprises by providing a sandbox for model training, policy impact assessment, and stress testing.

2 Project Objectives

1. **Construct a Hierarchical Multi-Agent System (HMAS):** Implement agents across four distinct functional layers (L_0 to L_3) to simulate both vertical management hierarchies and horizontal social interactions.

2. **Implement Dynamic Strategic Interaction:** Utilize Reinforcement Learning (RL) to create a closed loop where high-level policies affect the lower-level environment, and bottom-up behavioral feedback corrects high-level decisions.

3. **Generate High-Quality Synthetic Data:** Produce structured datasets ranging from micro-level individual behaviors (Household level) to macro-level economic indicators, solving the issues of data privacy and scarcity.

3 System Architecture & Methodology

The system adopts a **Hierarchical Multi-Agent Reinforcement Learning (HMARL)** architecture, dividing urban operations into four strategic levels.

3.1 Agent Layer Design

L3: Strategic Layer (Government/Mayor)
Role: Formulate long-term strategies and resource allocation.
Decision Frequency: Monthly / Quarterly / Yearly.
Actions: Adjusting `tax_rate`, `budget_allocation`, and `zoning_rules`.
Algorithm: Planning-based decision making + Multi-objective RL.

L2: Tactical Layer (Department Heads)
Role: Translate high-level strategies into specific execution plans (e.g., Dept. of Transportation).
Decision Frequency: Weekly / Monthly.
Actions: Setting `signal_plan`, `tariffs`, and road right-of-way allocation.
Rewards: Departmental KPIs (e.g., traffic efficiency) minus operational costs.

L1: Operational Layer (Service Operators)
Role: Entities providing specific services (e.g., Bus Companies, Hospitals).
Decision Frequency: Daily / Weekly.
Actions: Optimization of `schedule`, workforce dispatching, and pricing.
Algorithm: Contextual RL or Rule-based fine-tuning.

L0: Behavioral Layer (Citizens/Households)
Role: The atomic units of the system that react to the environment.
Decision Frequency: Minute / Hour.
Actions: Mobility decisions, consumption choices, and residential relocation.
Algorithm: Primarily **Imitation Learning** (to mimic realistic human behavior).

3.2 Interaction Loop Mechanism

The system operates on a **Bi-level Feedback Loop** (Stackelberg Game structure):

1. **Top-Down (Policy Propagation):** Policies from L_3/L_2 alter environmental parameters (e.g., fuel tax hike), directly modifying the State/Action space for lower levels.

2. **Bottom-Up (Data Aggregation):** Behavioral outcomes of L_0/L_1 (e.g., congestion, unemployment) are aggregated into KPIs, serving as Reward signals for L_2/L_3.

4 Expected Contributions

- **Data Democratization:** Provide academia and industry with an open-source urban data generator, lowering the barrier for AI research.
- **Policy Sandbox:** Enable governments to test radical policies (e.g., congestion pricing) in a "zero-risk" virtual environment.
- **Technical Validation:** Validate the convergence of Hierarchical RL in a large-scale environment with heterogeneous agents.

5 Tentative Roadmap

Phase 1 (Months 1-2)	**Environment & Baseline.** Build the base map and simulation engine. Implement L_0 Citizen mobility models using Imitation Learning.
Phase 2 (Months 3-4)	**Operational & Tactical Layers.** Implement L_1 Operators and L_2 Department Agents. Test traffic signal control mechanisms.
Phase 3 (Months 5-6)	**Strategic Layer & Economic Loop.** Integrate L_3 Government Agents. Complete the closed economic loop (Tax $\rightarrow$ Budget $\rightarrow$ Service $\rightarrow$ Consumption).
Phase 4 (Months 7-8)	**Integration & Case Studies.** Full system integration testing. Simulate specific scenarios (e.g., fuel price shocks) and generate synthetic data reports.

6 Proposed Technology Stack

- **Simulation Engine:** CityFlow / SUMO / Custom Grid World.
- **RL Framework:** Ray RLLib / PettingZoo / Stable Baselines3.
- **Data Processing:** Pandas, GeoPandas, NumPy.
- **Visualization:** Matplotlib, React/D3.js Dashboard.

References

[1] Zheng, S., Trott, A., Srinivasan, S., et al. (2022). *The AI Economist: Taxation policy design via two-level deep reinforcement learning*. Science Advances, 8(18).

[2] Zheng, L., et al. (2022). *Stackelberg Actor-Critic: Game-Theoretic Reinforcement Learning Algorithms*. AAAI Conference on Artificial Intelligence.

[3] Vu, Q. L. H., et al. (2022). *Stackelberg Policy Gradient: Evaluating the Performance of Leaders and Followers*. ICLR Workshop on Gamification and Multiagent Solutions.

[4] Zhang, H., et al. (2019). *CityFlow: A Multi-Agent Reinforcement Learning Environment for Large Scale City Traffic Scenario*. The World Wide Web Conference (WWW).