

Demo: Lowering the Barrier to Traffic Simulation Through a Prompt-to-Config Agent



Scan Me

William Phong, Jianshu Liu, Patrick Simpauc, Audio Antuna, Logan Moreno
San Diego State University

<https://github.com/PiscesLab/PromptCARLA>

SDSU



Pisces Lab

Autonomous Driving Research Relies on Traffic Simulation

- Autonomous driving research aims to improve road safety, reduce traffic congestion, and expand mobility for increased independence and accessibility

- ❖ High-fidelity simulators such as SUMO, PTV ViSUM, and Ruth are very good simulation platforms that support controlled large-scale experimentation across diverse traffic environments.

- Among all the available simulation platforms, CARLA stands out!

- ❖ Fully customizable environment, including **map selection**, and **multiple weather conditions**



Town01	Town02	Town04	Town05	Town10HD
Size (km ²): 0.14	Size (km ²): 0.04	Size (km ²): 0.79	Size (km ²): 0.23	Size (km ²): 0.05
Junctions: 12	Junctions: 8	Junctions: 27	Junctions: 21	Junctions: 9
Road Crossing: No	Road Crossing: No	Road Crossing: Yes	Road Crossing: Yes	Road Crossing: Yes
Crosswalks: 0	Crosswalks: 27	Crosswalks: 66	Crosswalks: 16	Crosswalks: 15
Traffic Lights: 36	Traffic Lights: 24	Traffic Lights: 43	Traffic Lights: 53	Traffic Lights: 15



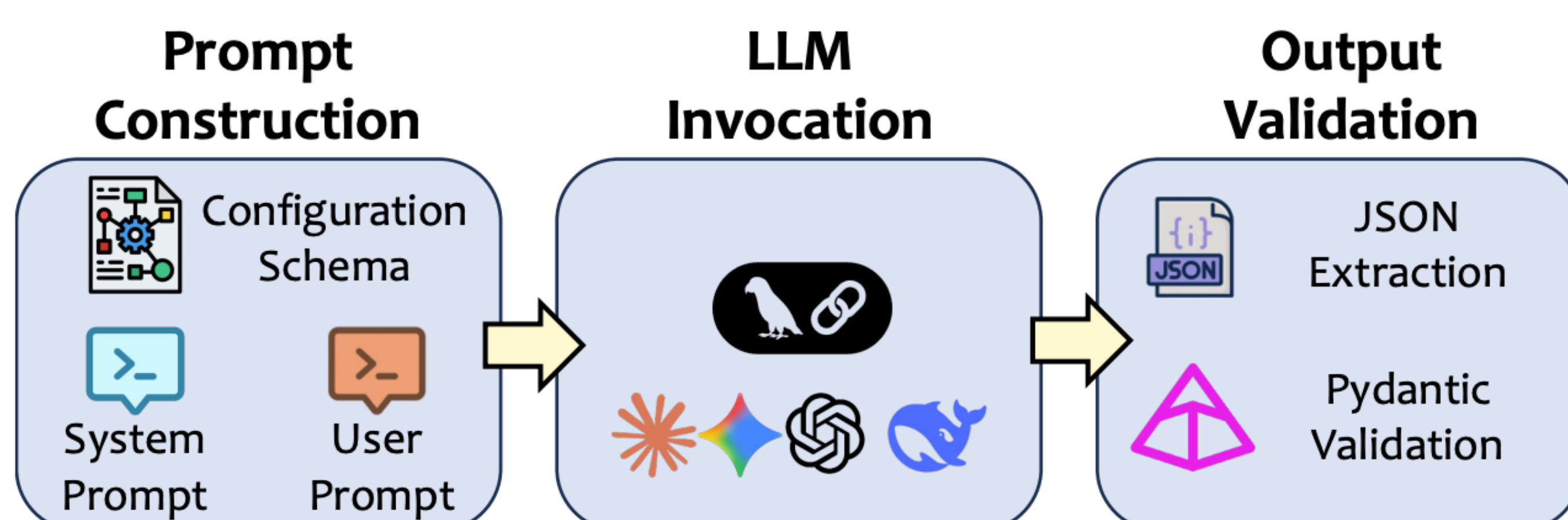
CARLA in Classrooms

- **CARLA is increasingly used in classrooms, but hard to start**

- ❖ Scenario generation requires domain knowledge
- ❖ Specialized syntax and tools increases the learning efforts



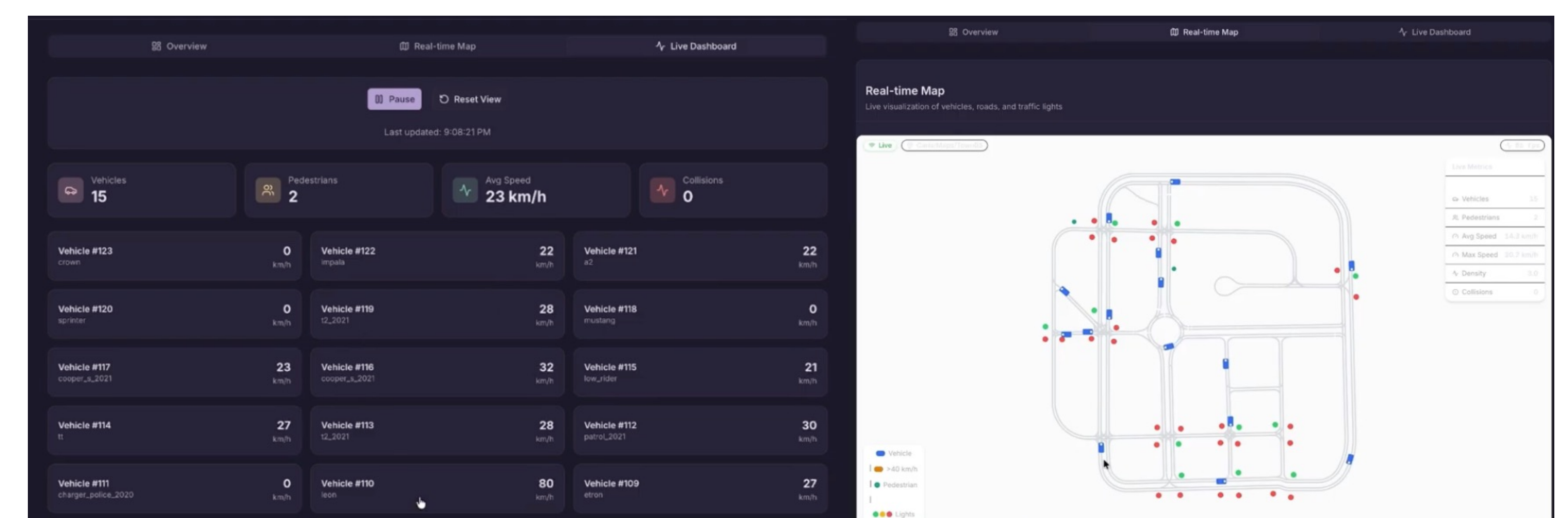
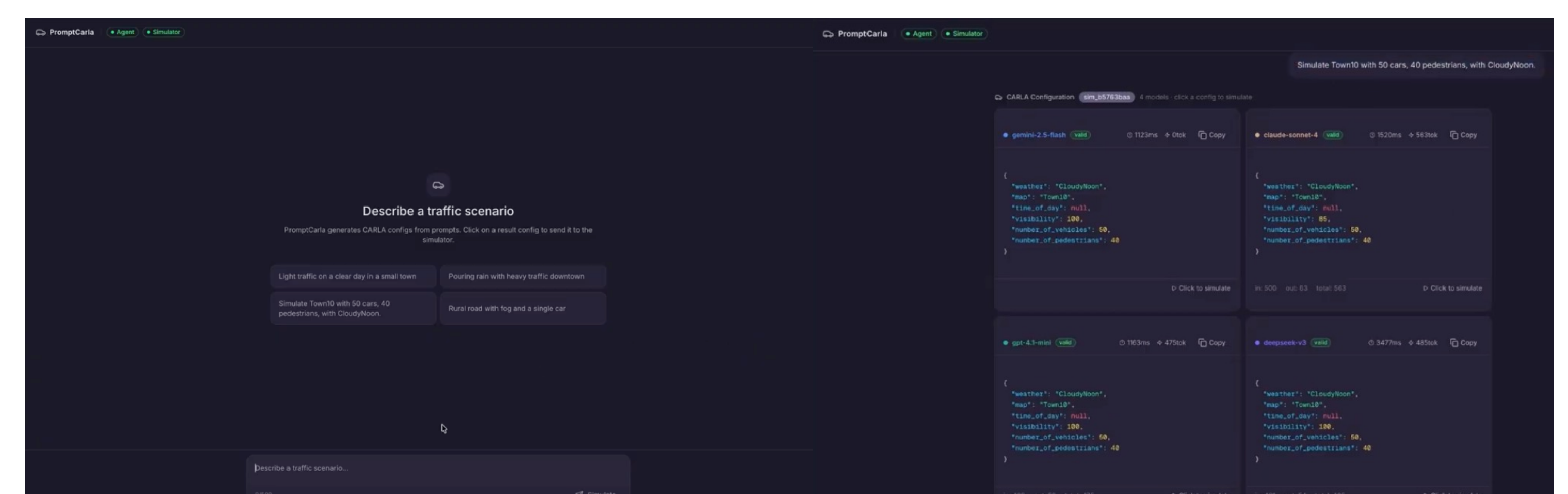
Our Solution: PromptCARLA



- **Key capabilities**

- ❖ Natural language to validated CARLA configuration via multi-model parallel inference
- ❖ Pydantic-based schema validation with automatic constraint correction
- ❖ Real-time 2D simulation map rendered with PixiJS v8
- ❖ Live actor telemetry streamed over WebSocket at 20Hz
- ❖ Docker Compose deployment across four isolated services

Demo Snapshots of PromptCARLA



Preliminary Results

	Tier	Model	Weather	Map	Vehicles	Peds.	InTok	OutTok	TotToks
Explicit parameter specification	T1	Claude	100%	100%	100%	100%	1,016	57	1,073
		DeepSeek	100%	100%	100%	100%	872	52	925
		Gemini	100%	100%	100%	100%	—	—	—
		GPT-4.1	100%	100%	100%	100%	861	55	916
Descriptive natural language	T2	Claude	65.0%	75.0%	—	—	1,006	73	1,079
		DeepSeek	50.0%	85.0%	—	—	866	61	927
		Gemini	60.0%	80.0%	—	—	—	—	—
		GPT-4.1	60.0%	75.0%	—	—	855	57	912
Adversarial and edge-case	T3	Claude	52.6%	21.1%	85.7%	33.3%	1,006	87	1,093
		DeepSeek	10.5%	21.1%	85.7%	33.3%	865	53	918
		Gemini	31.6%	47.4%	85.7%	33.3%	—	—	—
		GPT-4.1	10.5%	10.5%	83.3%	33.3%	854	82	936

Conclusion and Limitations

- Our preliminary results indicate all four LLM models achieved a 100% rate of producing parseable output
- However, the prompt dataset is relatively small, and we did not compare against alternative prompt-engineering baselines. In addition, no formal user study was conducted.
- Future work includes expanding support for additional traffic simulation platforms, conducting classroom usability studies, and extending the configuration schema to support richer scenario parameters.
- We also plan to investigate interactive multi-round scenario refinement, user feedback-driven correction mechanisms, and model combination strategies.

Contact

Dr. Jianshu Liu
Department of Computer Science
College of Science
Email: jliu11@sdsu.edu