Leveraging Large Language Models (LLMs) for Interpreting Datasets of Autonomous and Semi-Autonomous Vehicles

Prof. Max Mauro Dias Santos, PhD
DAELE – UTFPR-PG
Paraná - Brazil

Toronto Metropolitan University

Max Mauro Dias Santos, PhD

Department of Electronics
Federal Technological University of Paraná – Ponta Grossa

Responsabilities

- **Associate Professor** – Teaching
- **Chair** – Management
- **Director** – Research and Development

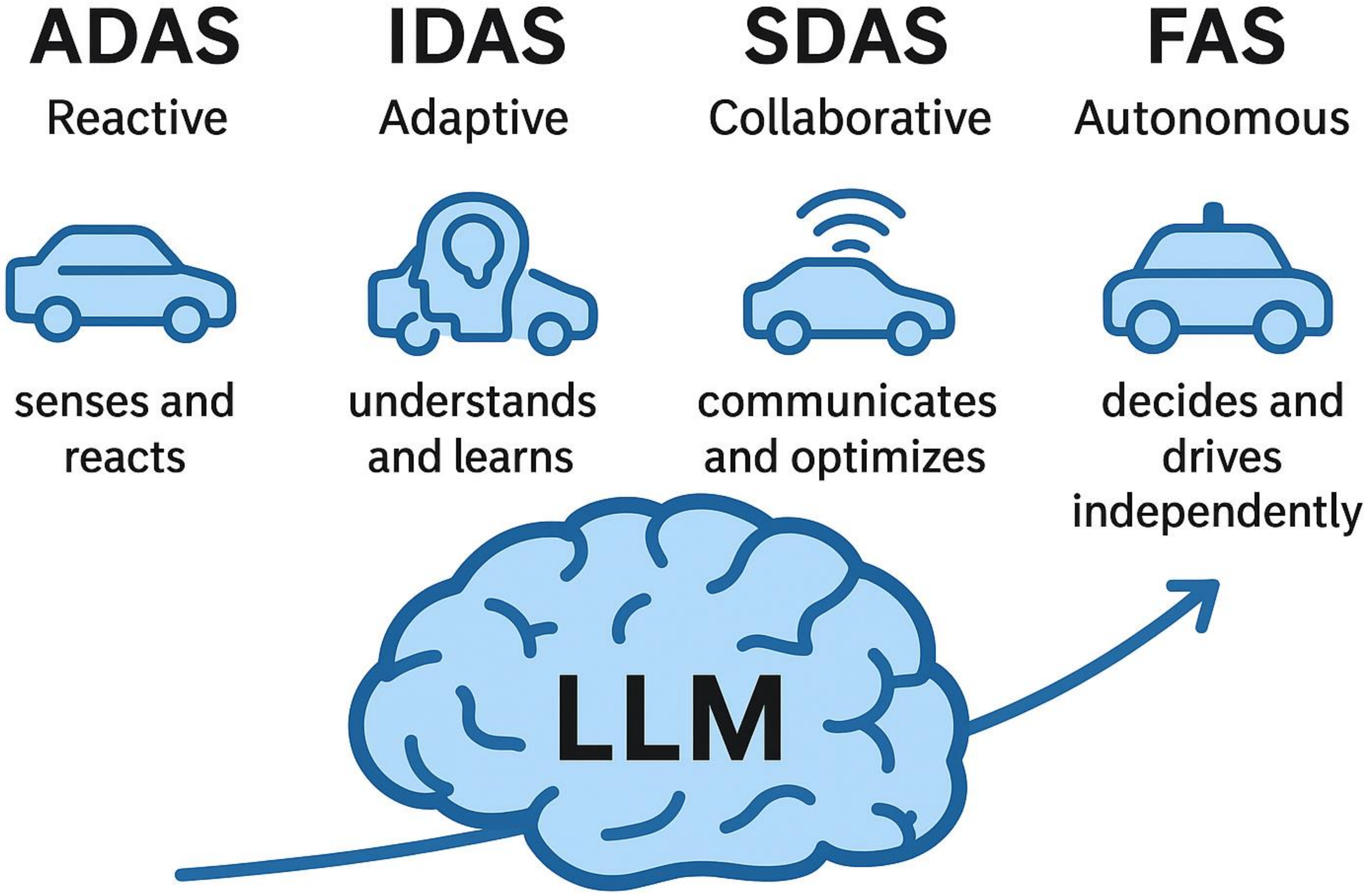
maxsantos@utfpr.edu.br

Toronto Metropolitan University

Summary

1. Introduction
2. Context and Motivation
3. Why are LLM and SLM so good for Mobility?
4. ADAS → IDAS → SDAS → FAS
5. Foundations of LLMs
6. LLMs for the Automotive Domain
7. SLMs for the Automotive Domain
8. Our Contributions in LLMs for the Automotive
9. Potential Use Cases
10. Expected Benefits
11. Challenges and Limitations
12. Future Perspectives
13. References

Toronto Metropolitan University

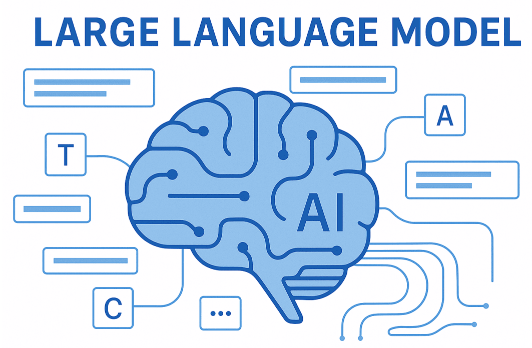


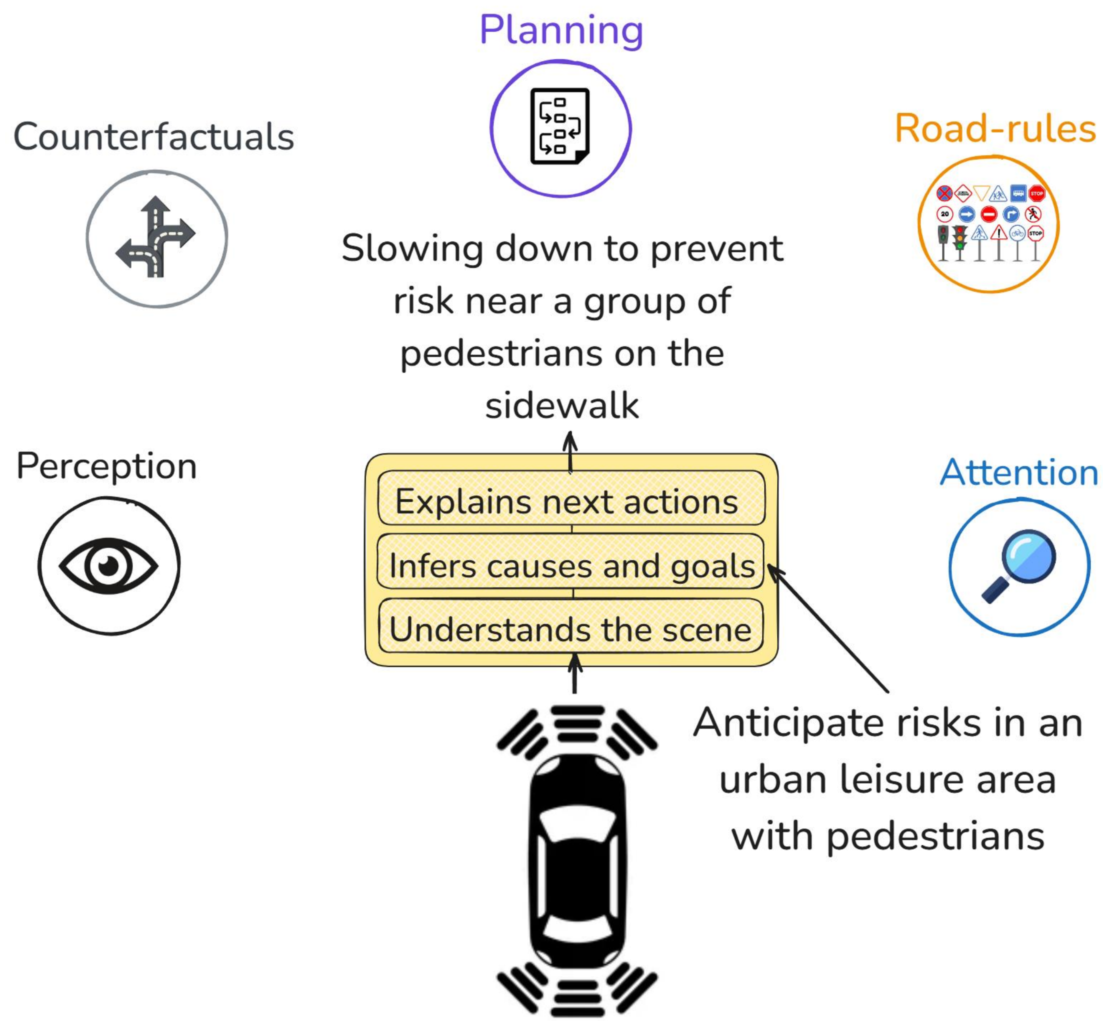
Toronto Metropolitan University

1. Introduction

LLM

ITS





Intersection of LLM and ITS

2. Context and Motivation

Autonomous and semi-autonomous vehicles must operate safely in dynamic, uncertain environments, where robust data interpretation is essential.

Datasets are too large and heterogeneous to analyze manually

Complex interactions between sensors create ambiguity or conflicting evidence

Safety validation requires **interpreting rare or edge-case scenarios** that are difficult to capture

Regulatory bodies demand **transparent and explainable justifications** for vehicle behavior

Toronto Metropolitan University

LLMs bring unique capabilities to address these gaps.

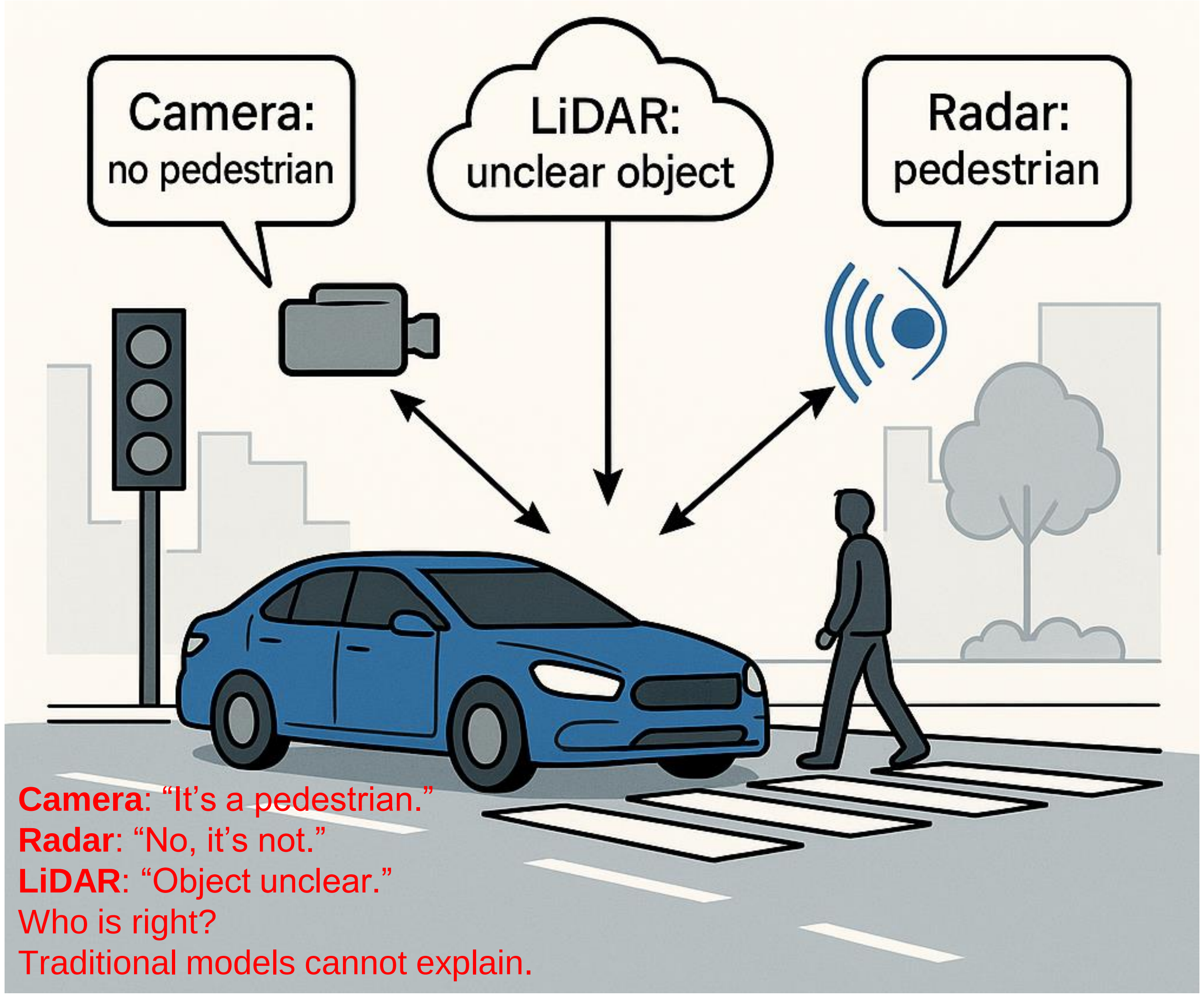
2. Context and Motivation

Datasets are too large and heterogeneous to analyze manually

Toronto Metropolitan University

2. Context and Motivation

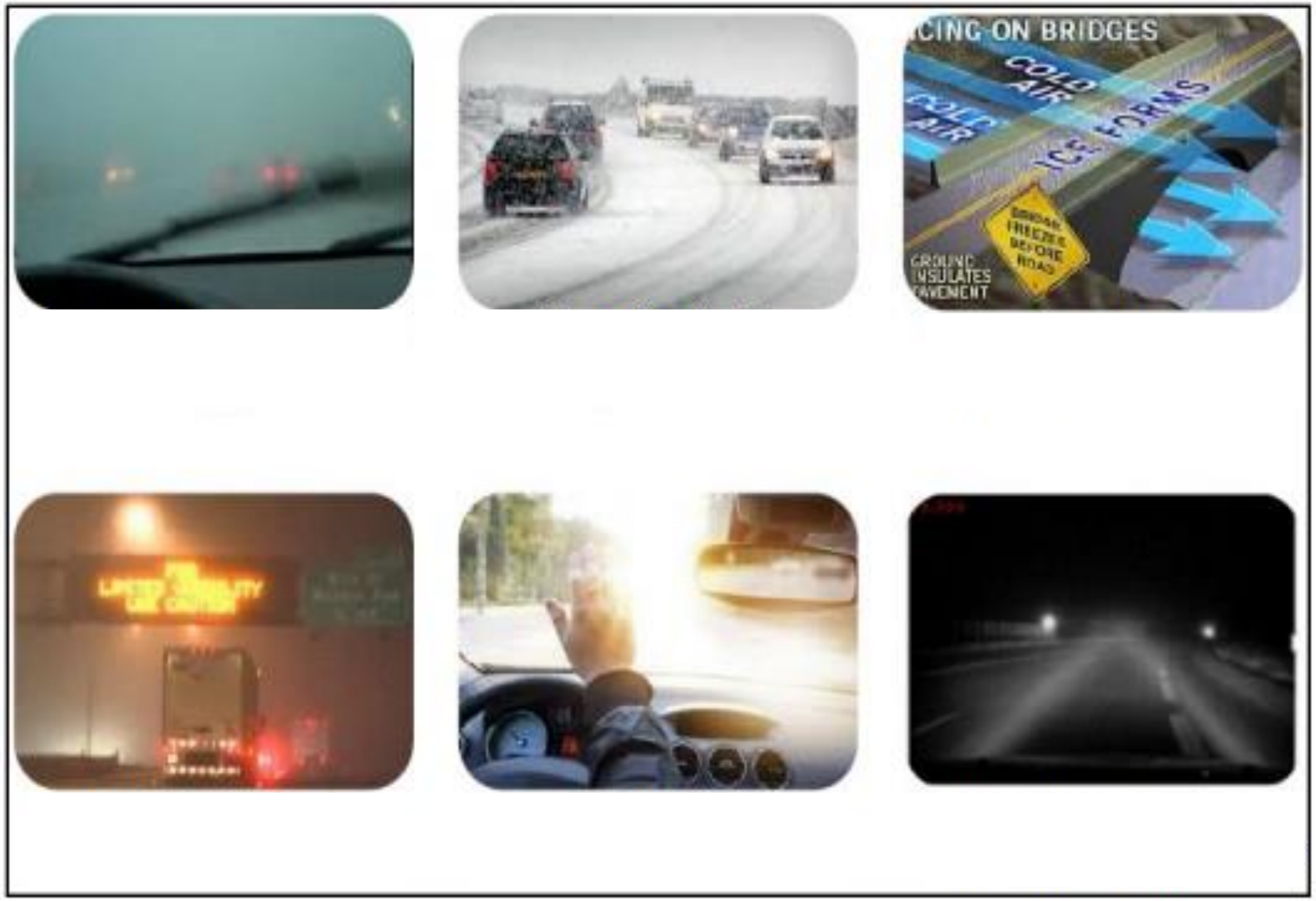
Complex interactions between sensors create ambiguity or conflicting evidence



Camera: “It’s a pedestrian.”
Radar: “No, it’s not.”
LiDAR: “Object unclear.”
Who is right?
Traditional models cannot explain.

2. Context and Motivation

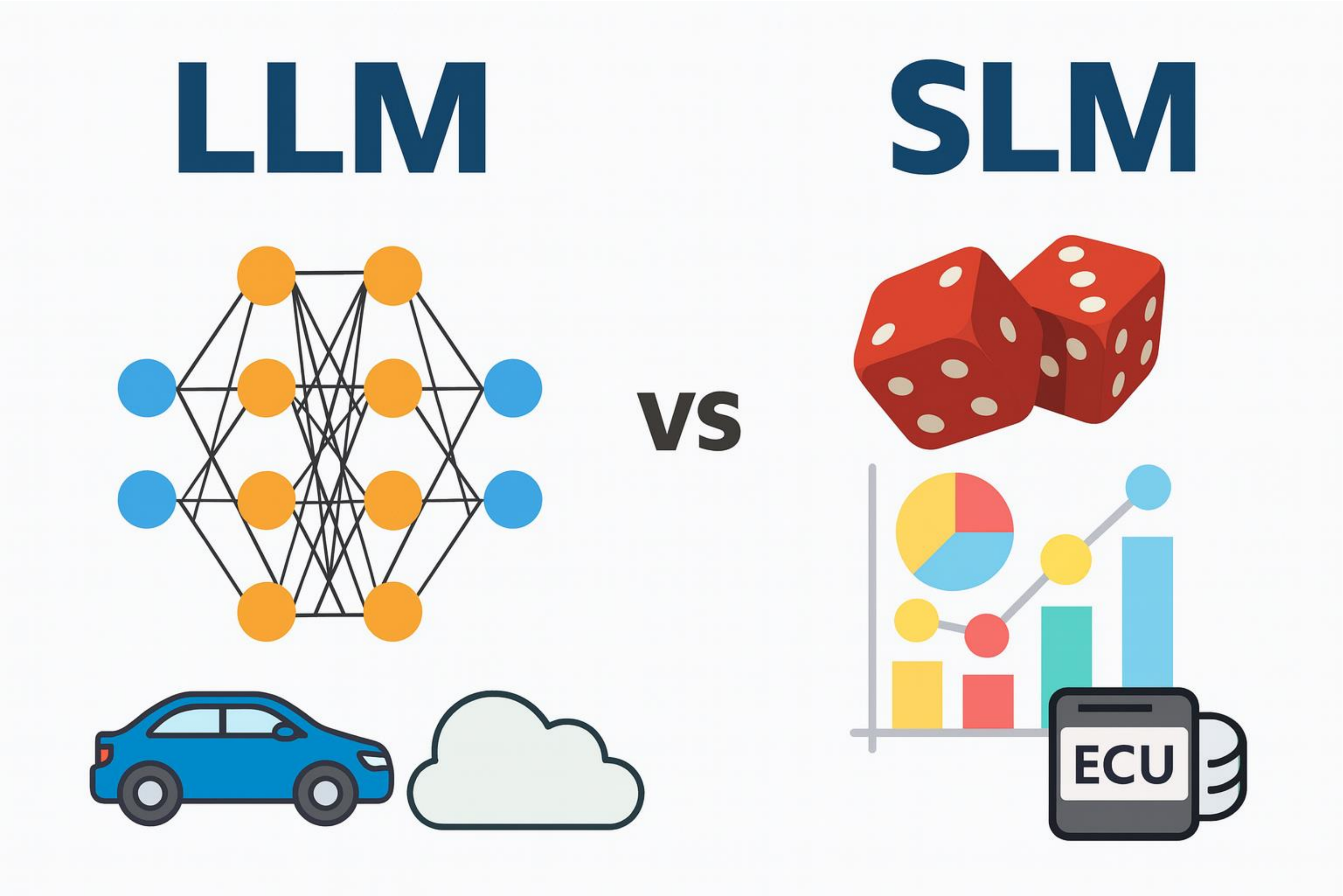
Safety validation requires **interpreting rare or edge-case scenarios** that are difficult to capture



2. Context and Motivation

Regulatory bodies demand **transparent and explainable justifications** for vehicle behavior

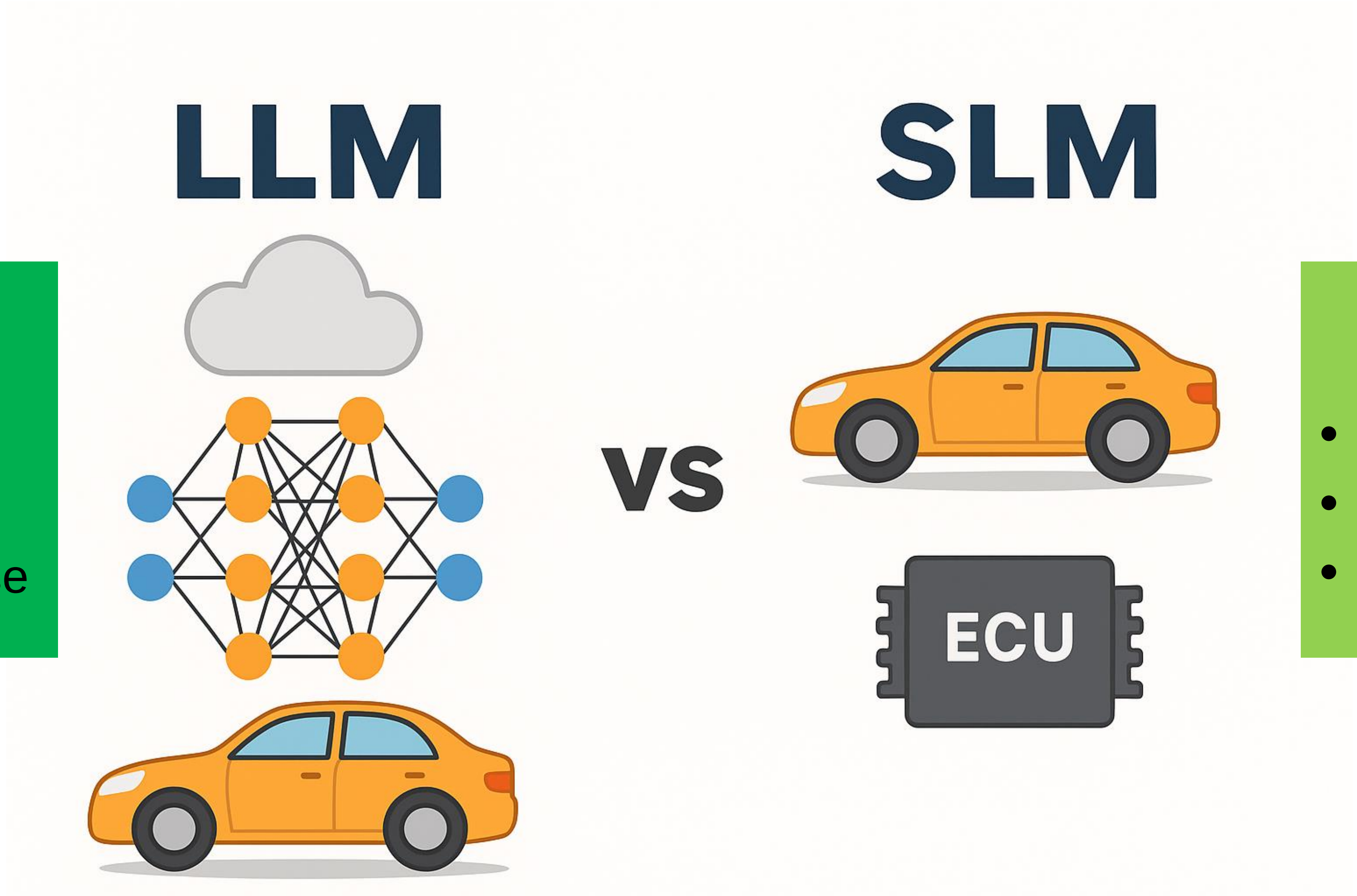
3. Why are LLM and SLM so good for Mobility?



3. Why are LLM and SLM so good for Mobility?

Cloud

- Large scale
- High latency
- General purpose

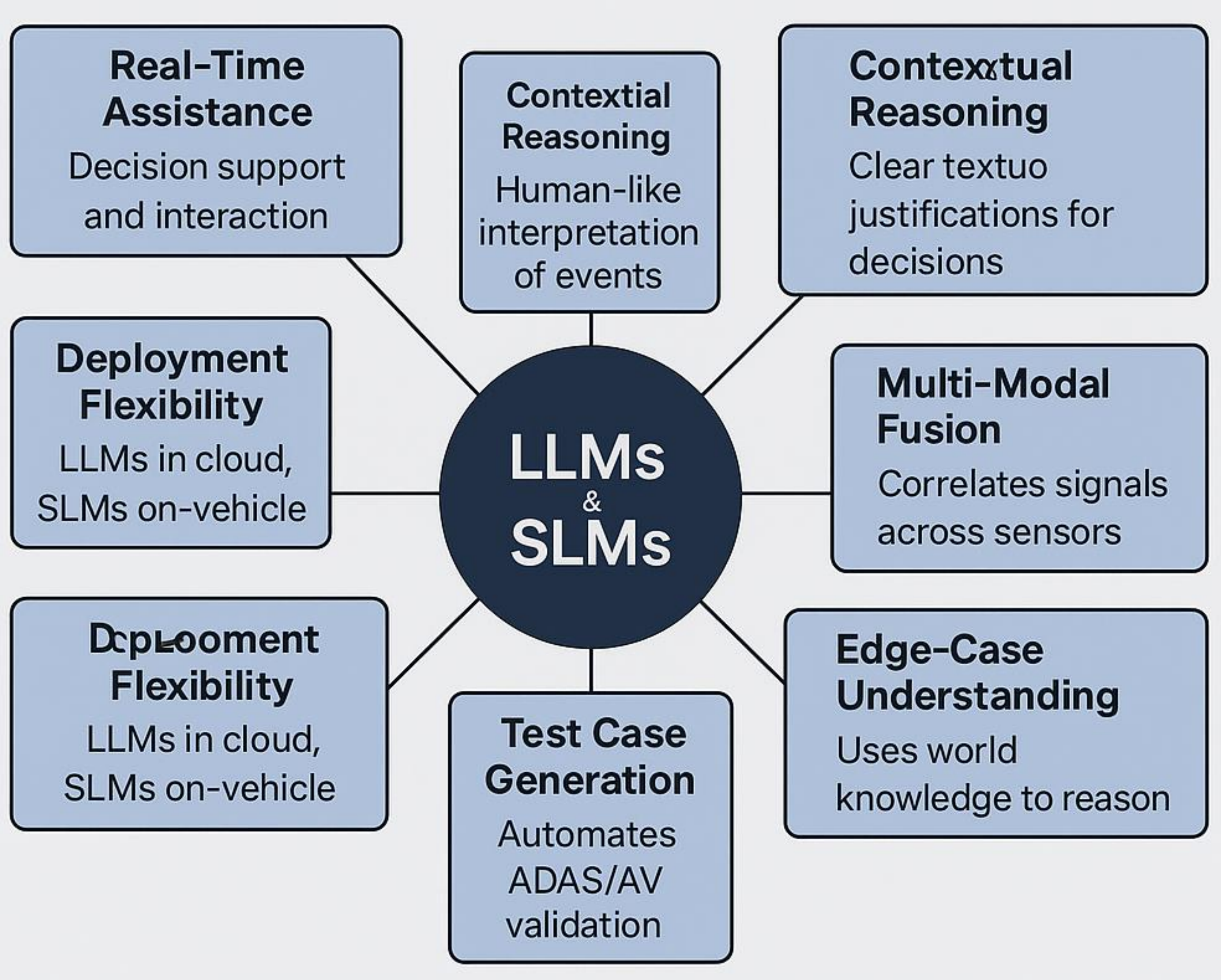


Edge

- Small scale
- Low latency
- Specific task

3. Why are LLM and SLM so good for Mobility?

LLMs and SLMs bring human-like reasoning, explanation, and semantic understanding to complex automotive data - closing the gap between perception and high-level decision-making, and enabling safer, more transparent mobility.



4. ADAS → IDAS → SDAS → FAS

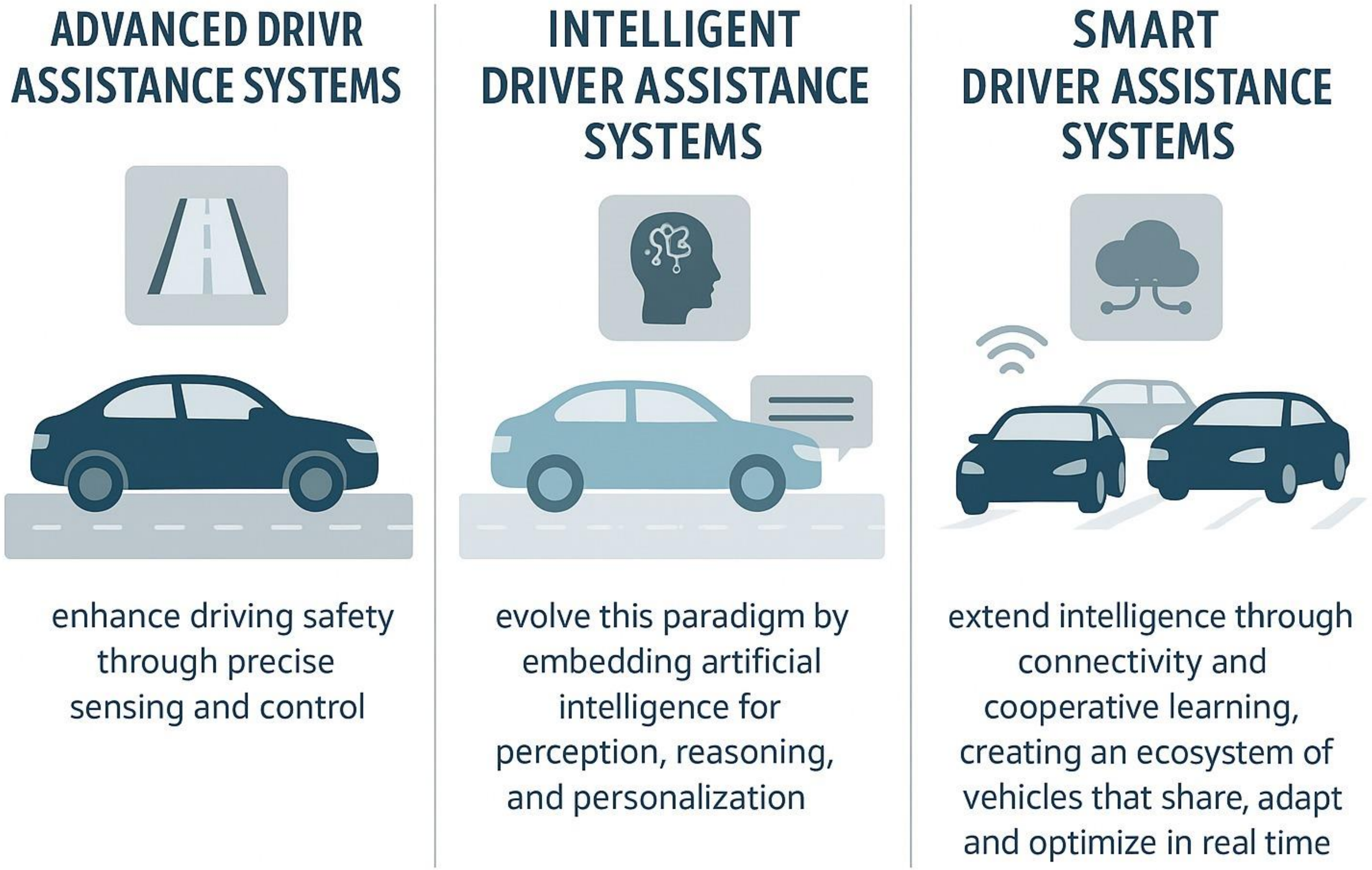
Advanced Driver Assistance Systems (ADAS) enhance driving safety through precise sensing and control.

Intelligent Driver Assistance Systems (IDAS) evolve this paradigm by embedding artificial intelligence for perception, reasoning, and personalization.

Smart Driver Assistance Systems (SDAS) extend intelligence through connectivity and cooperative learning, creating an ecosystem of vehicles that share, adapt, and optimize in real time.

4. ADAS → IDAS → SDAS → FAS

xDAS
Everything Driver Assistance Systems

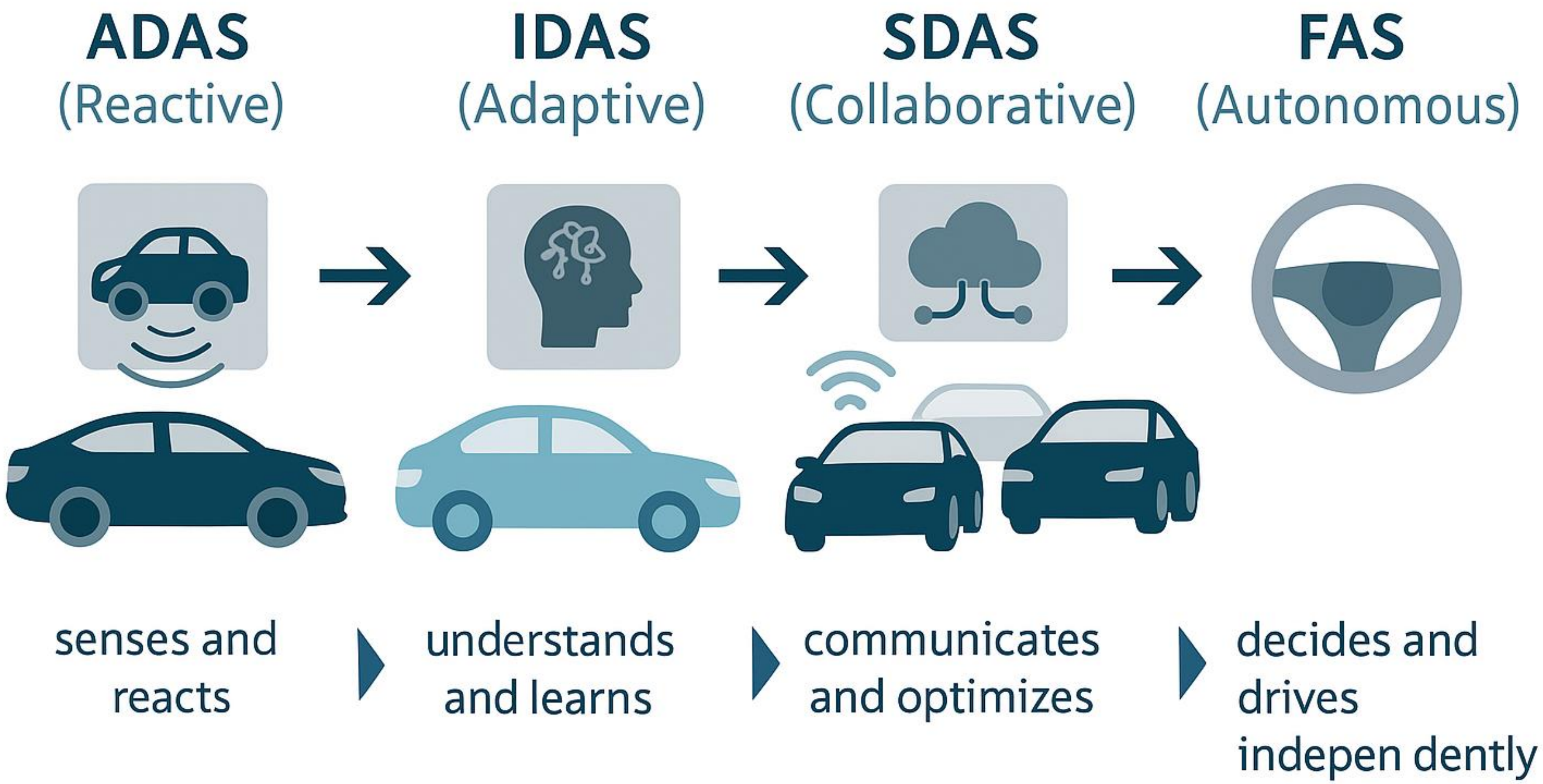


Toronto Metropolitan University

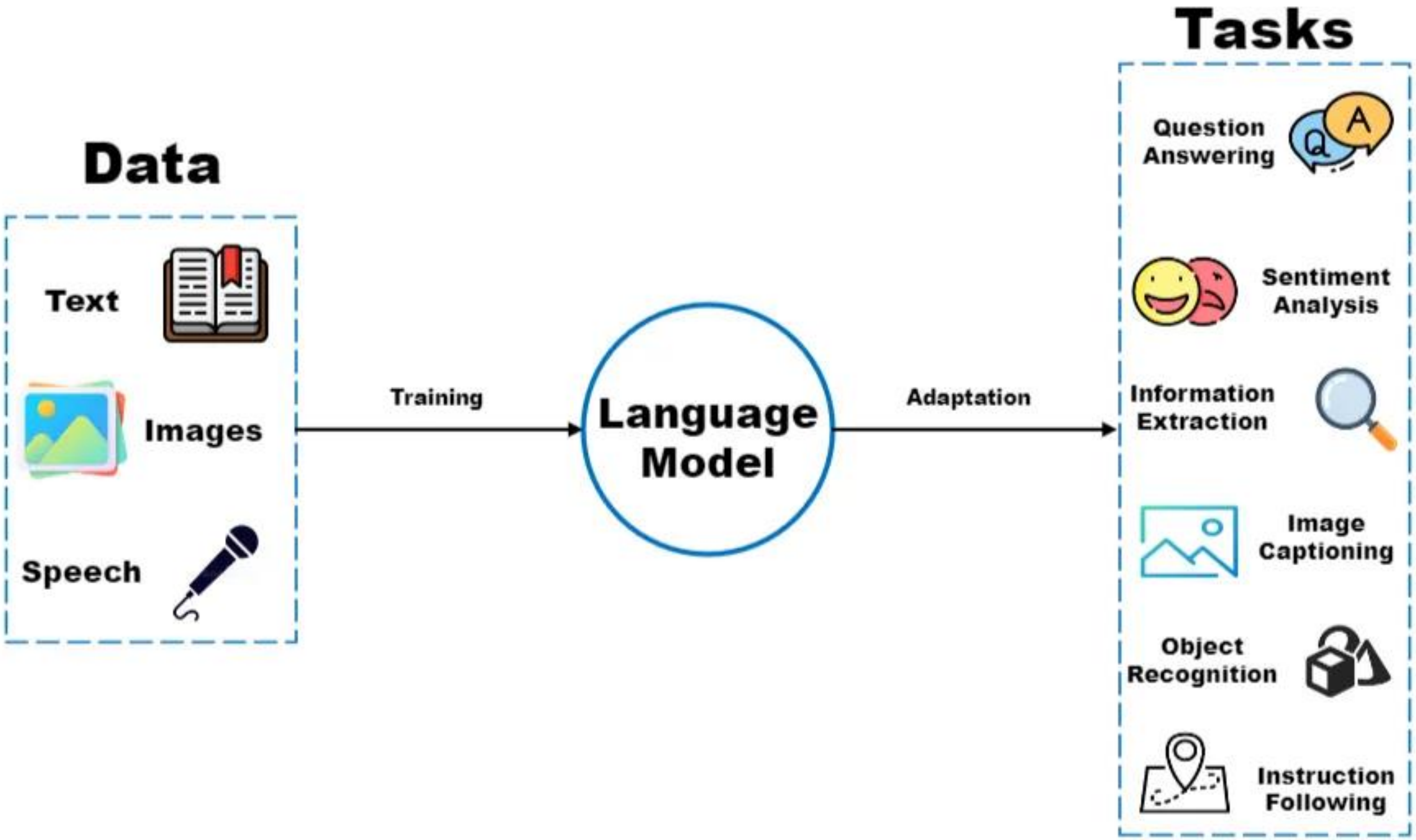
4. ADAS → IDAS → SDAS → FAS

Feature	ADAS	IDAS	SDAS
Core Concept	Advanced sensing & control	AI reasoning & learning	Connected, cooperative, and predictive AI
Decision Logic	Deterministic / rule-based	Cognitive / adaptive	Collective / distributed
Data Source	Local sensors	Sensors + driver + context	Sensors + V2X + cloud
AI Integration	Minimal (mostly perception)	Deep integration (perception, planning, reasoning)	Integrated across vehicles and infrastructure
Learning Ability	None (fixed logic)	Online/offline learning	Shared / federated learning
Communication	Isolated vehicle	Context-aware vehicle	Connected ecosystem
Objective	Driver assistance & safety	Intelligent support & personalization	System-level optimization & coordination
Example Features	ACC, LKA, AEB	Predictive ACC, AI co-pilot	Cooperative ACC, cloud-based hazard prediction

4. ADAS → IDAS → SDAS → FAS

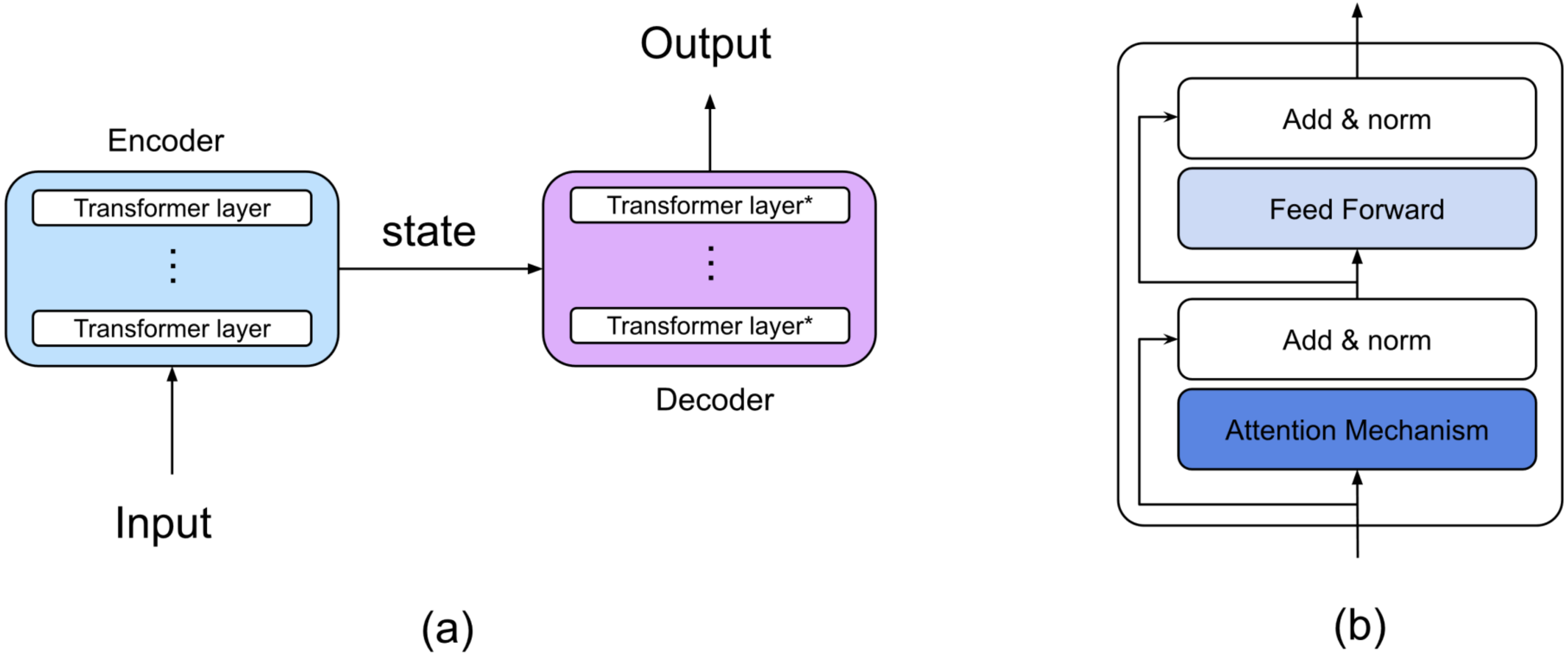


5. Foundations of Large Language Model

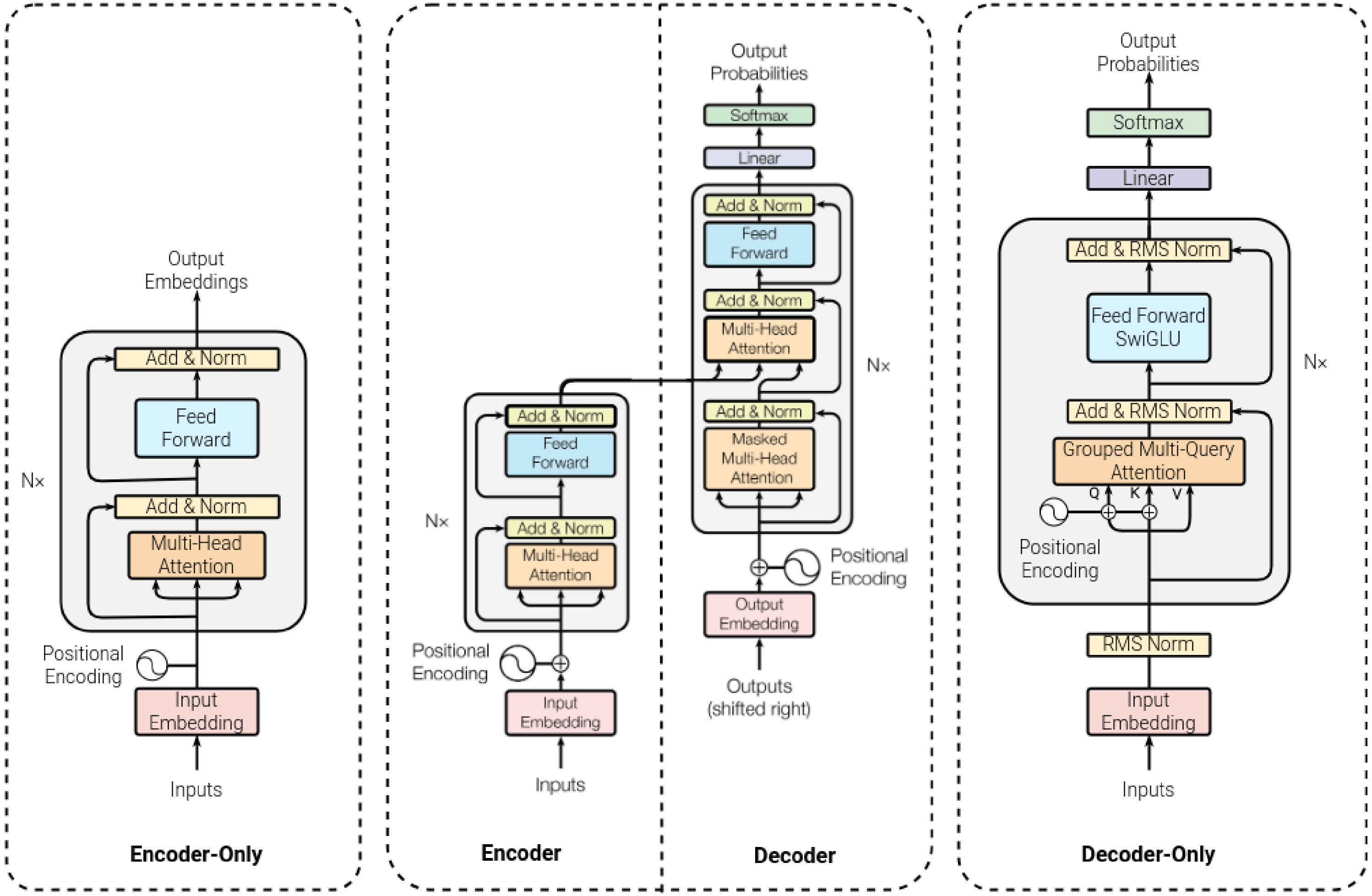


Toronto Metropolitan University

5. Foundations of Large Language Model



5. Foundations of Large Language Model



BERT (Devlin et al., 2018)

Original Transformer (Vaswani et al., 2017)

LLaMA (Touvron et al., 2023)

https://www.luminis.eu/blog/llm-series-part-1-a-comprehensive-introduction-to-large-language-models/

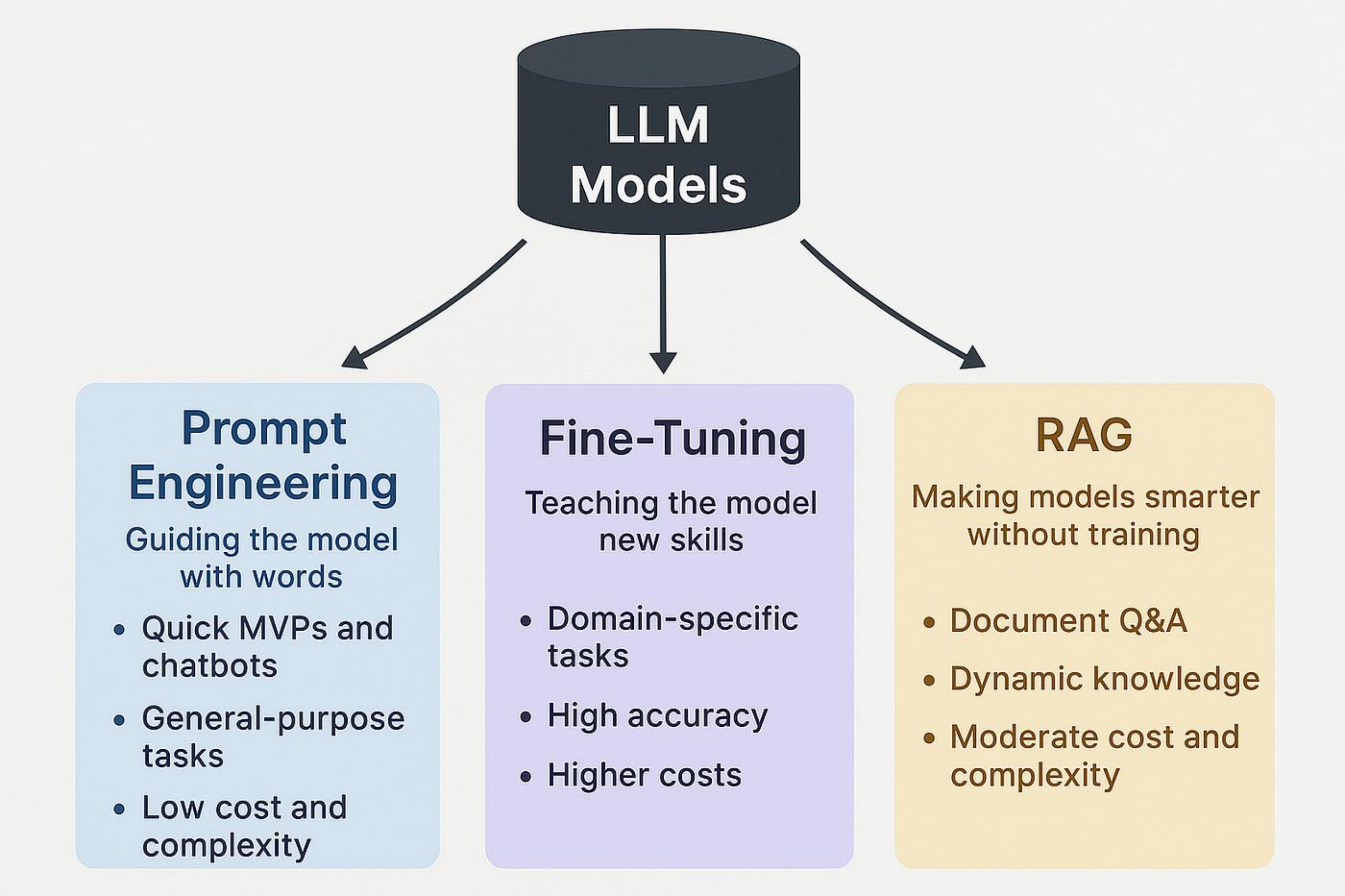
5. Foundations of Large Language Model

Large Language Models (LLMs) in three levels:

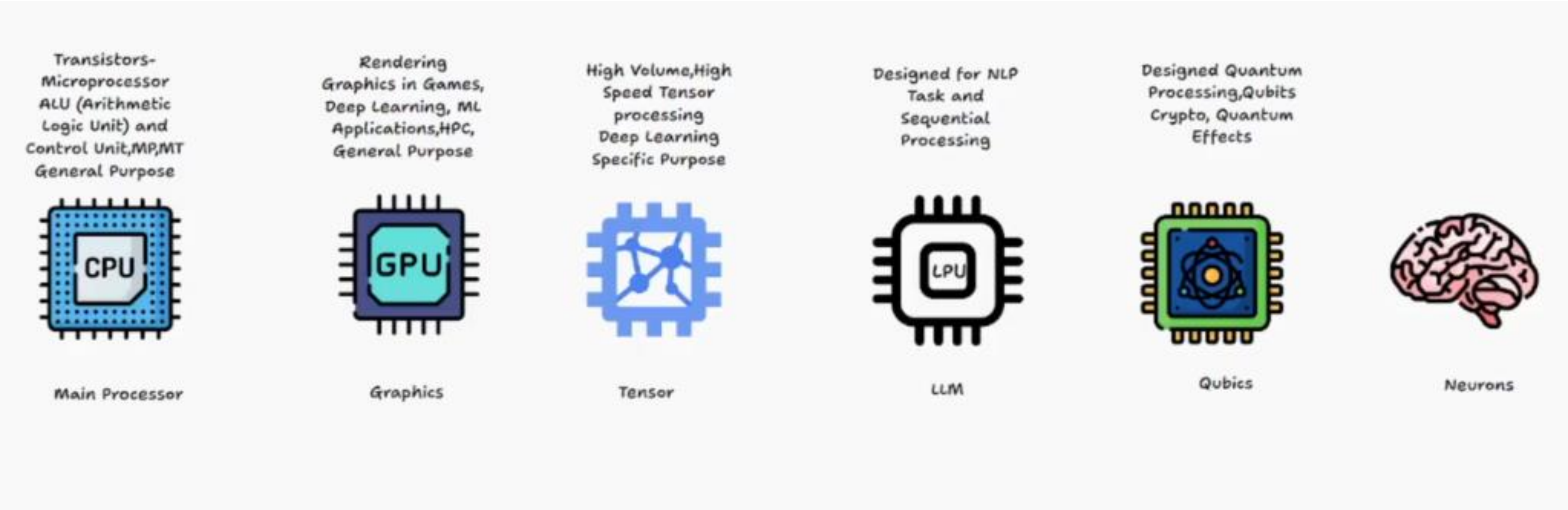
- **Prompt Engineering** – using pre-trained LLMs with structured prompts.
- **Model Fine-Tuning** – adapting LLMs to domain-specific datasets.
- **Build Your Own** – creating a customized LLM for autonomous vehicle data.

Toronto Metropolitan University

5. Foundations of Large Language Model



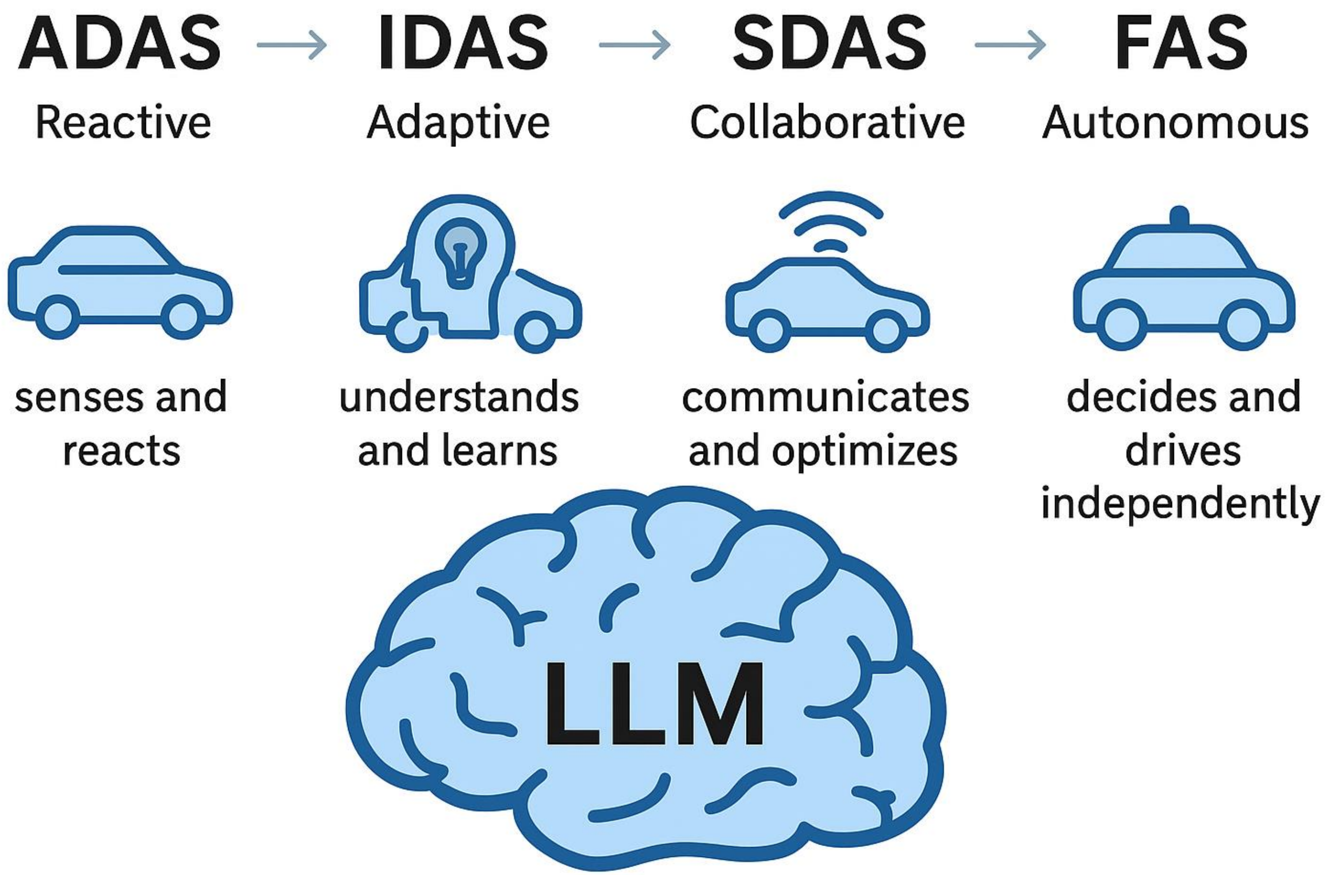
5. Foundations of Large Language Model



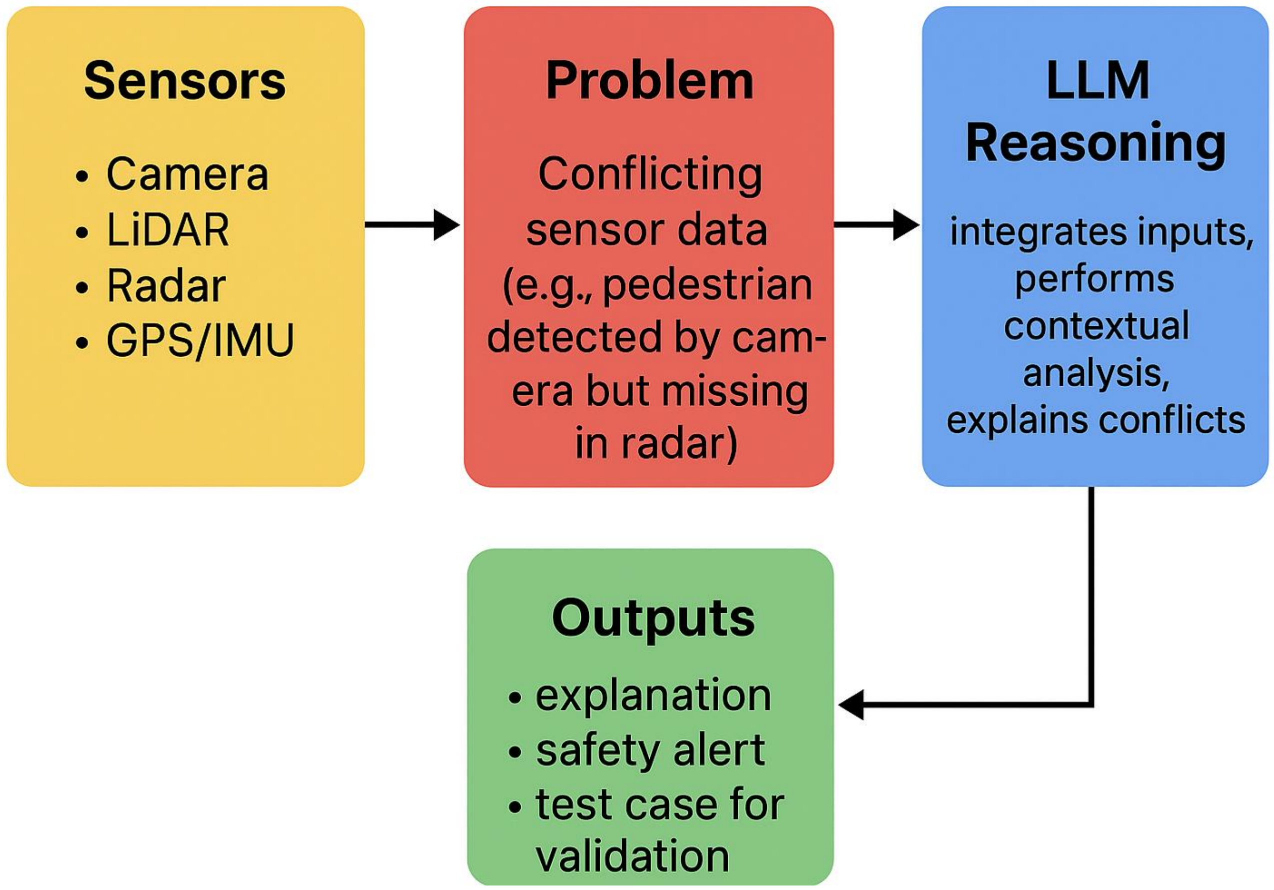
Toronto Metropolitan University

https://blog.jcharistech.com/2024/02/28/types-of-processing-units-for-llm-and-machine-learning-apps/

6. LLMs for the Automotive Domain



6. LLMs for the Automotive Domain

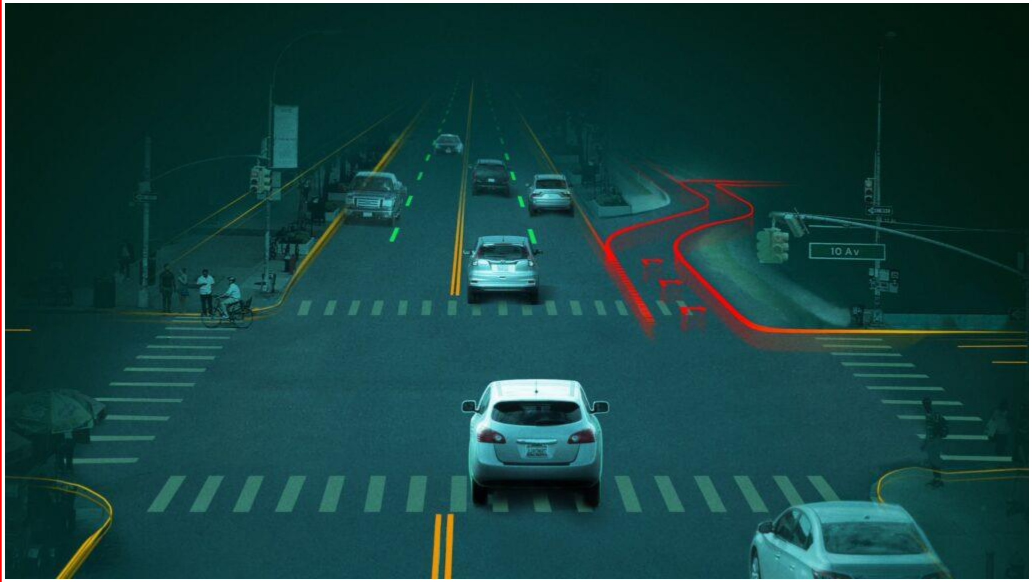


6. LLMs for the Automotive Domain

1. Data Sources (Vehicle Ecosystem)
2. Preprocessing and Representation
3. Large Language Model (LLM) Core
4. Applications in Automotive Systems
5. Human-in-the-Loop Outputs

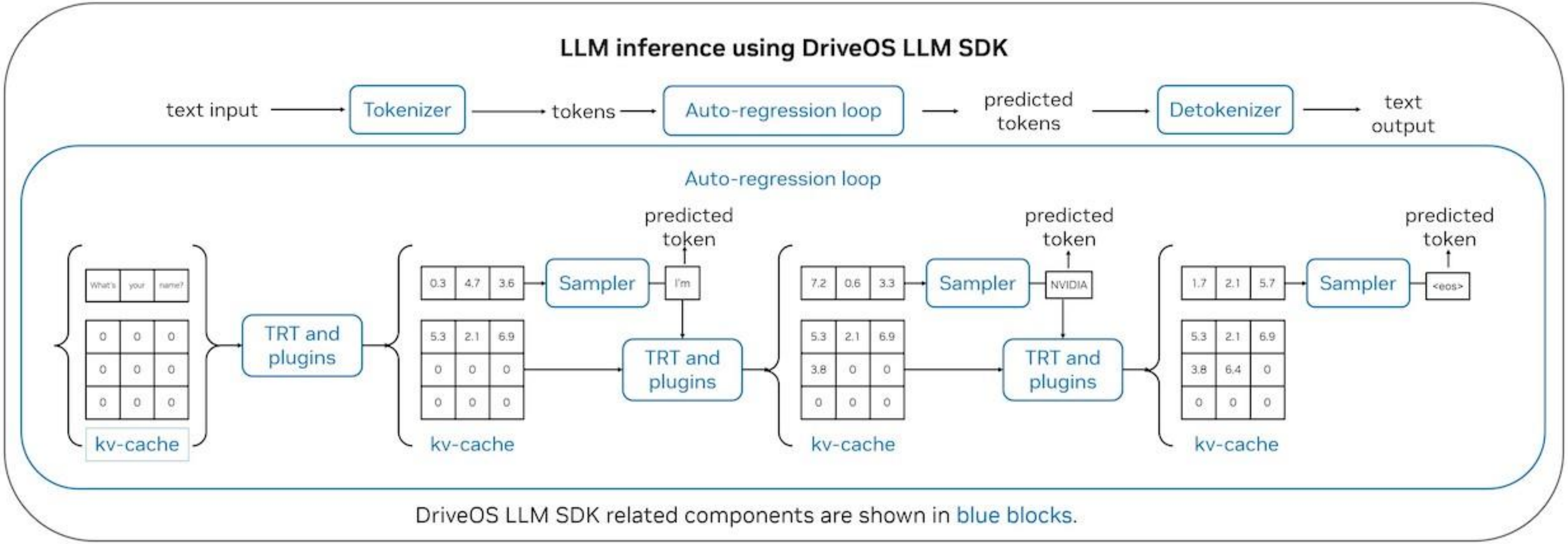
6. LLMs for the Automotive Domain

Large Language Models (LLMs) demonstrate strong generalization in natural language processing, powering applications such as translation, digital assistants, recommendation systems, code generation, and cybersecurity. In automotive domains, demand is rising for LLM-based solutions in both autonomous driving and in-cabin features. However, deploying LLMs and Vision-Language Models (VLMs) on resource-constrained automotive platforms remains a critical challenge.



Toronto Metropolitan University

6. LLMs for the Automotive Domain



7. SLMs for the Automotive Domain

- A **Small Language Model (SLM)** is a **lightweight neural language model** designed to perform natural-language understanding, reasoning, or generation with **low computational cost**, making it suitable for **embedded, edge, or resource-constrained devices**.
- Compared to LLMs, an SLM typically has:
  - **fewer parameters** (50M to ~3B)
  - **lower memory footprint**
  - **faster inference**
  - **lower power consumption**
  - ability to run **on-device** or **in real time**
- SLMs maintain many capabilities of LLMs - such as question answering, summarization, and basic reasoning—but with optimized architectures so they can run **without requiring cloud computing**.

Toronto Metropolitan University

7. SLMs for the Automotive Domain

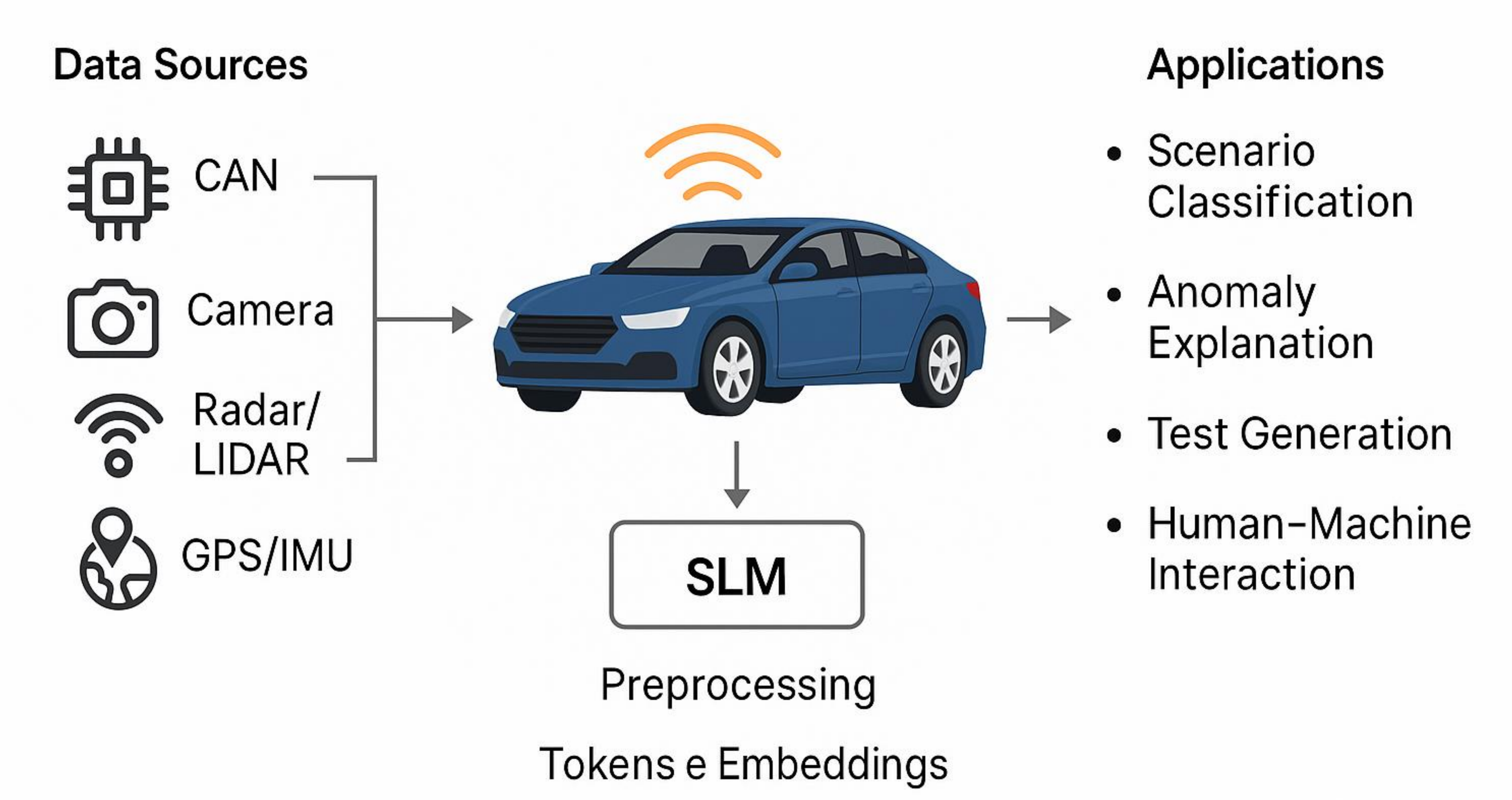


Figure from ChatGPT

7. SLMs for the Automotive Domain

SLMs for Human-Machine Interaction

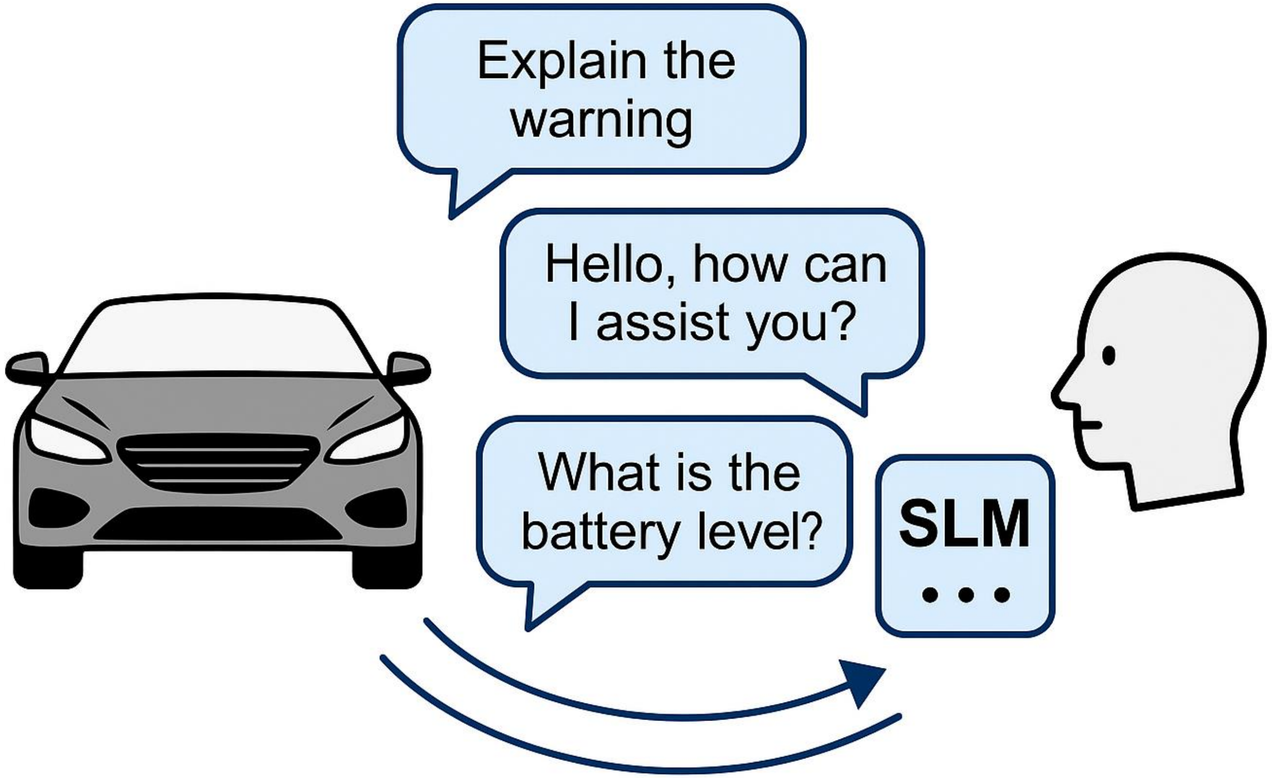
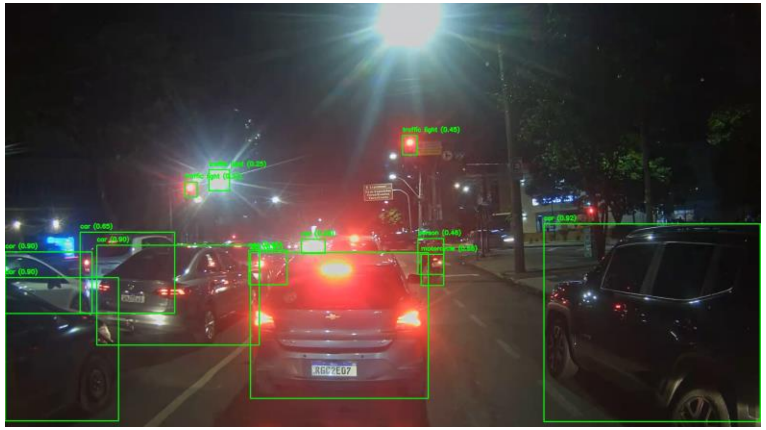


Figure from ChatGPT

8. Our Contributions in LLMs for the Automotive

- **Conference**: IEEE ISIE 2025 – Toronto - CA
- **Title**: LLM-Powered Framework for Interpretable Traffic Rule Processing in Autonomous Driving (UTFPR-PG and TMU)

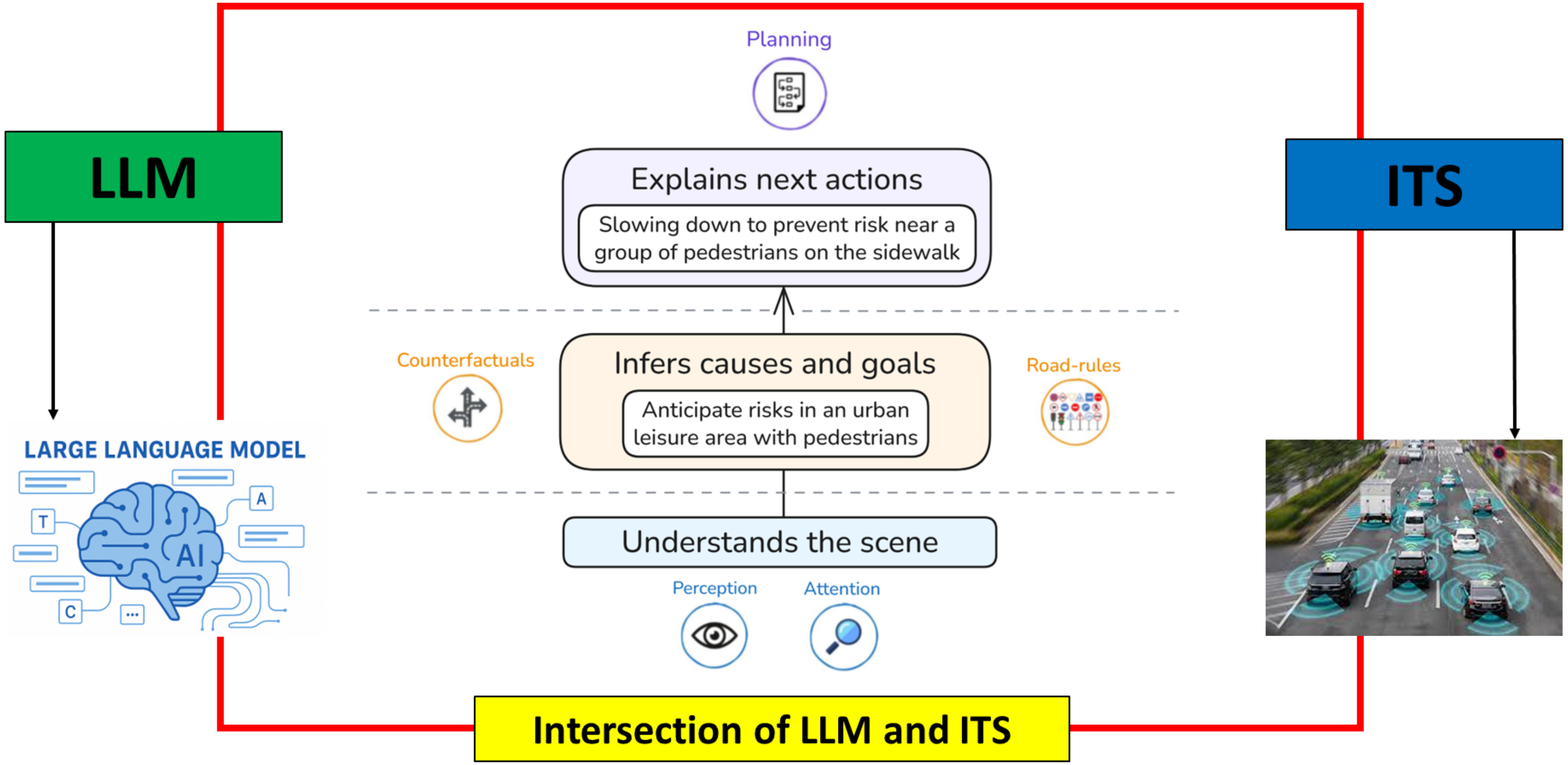
LLMs for driving assistance.



Toronto Metropolitan University

8. Our Contributions in LLMs for the Automotive

- **Conference**: IEEE ISGT 2025 – Dubai - UAE
- **Title**: Multimodal Large Language Model Framework for Safe and Interpretable Grid-Integrated EVs (UTFPR-PG, TMU, and UofT)



8. Our Contributions in LLMs for the Automotive

- **CarCara**
  - GSA has vehicles for DAQ

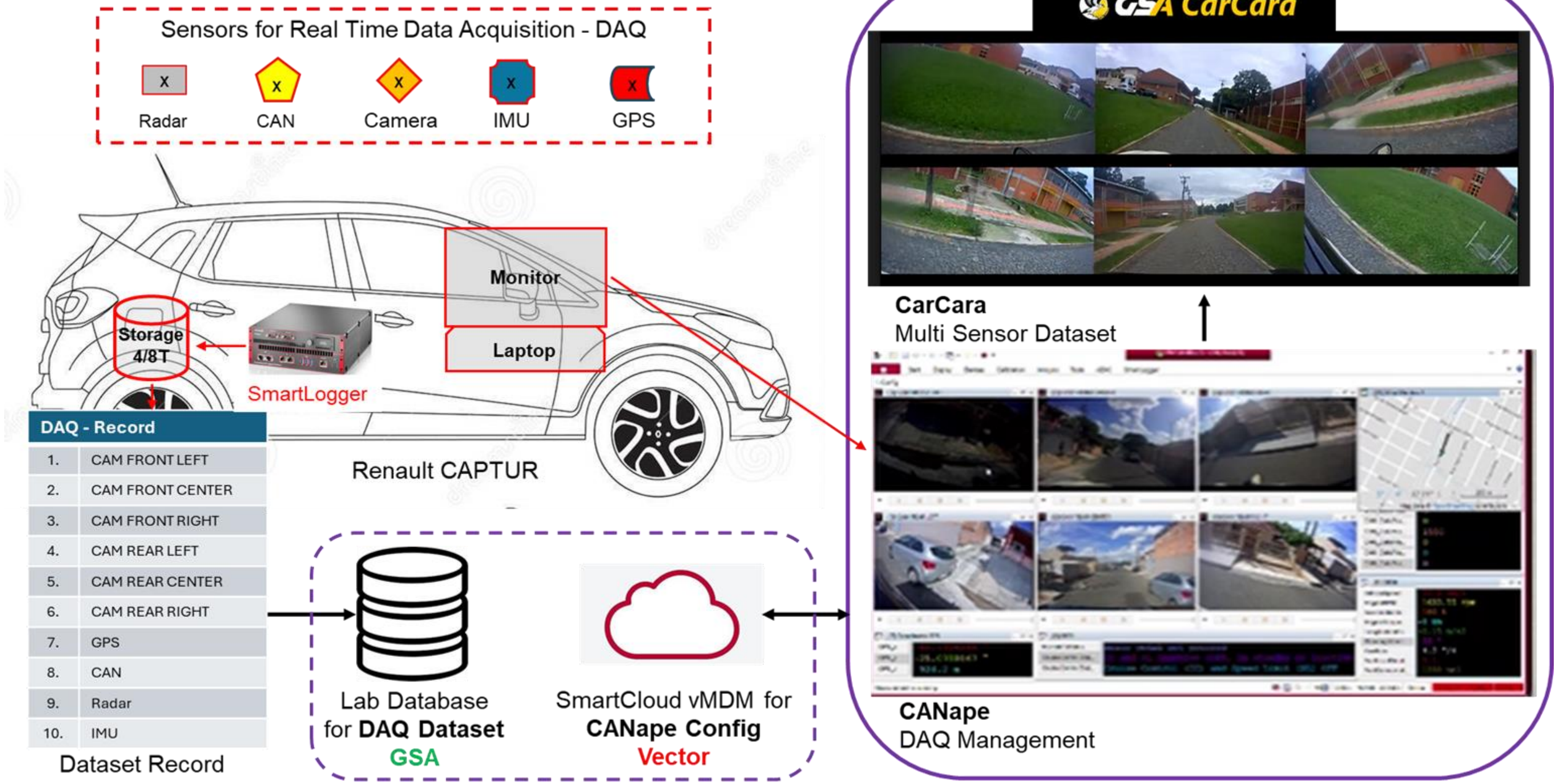
DAF

STELLANTIS

RENAULT

8. Our Contributions in LLMs for the Automotive

- **CarCara**
  - DAQ for Dataset



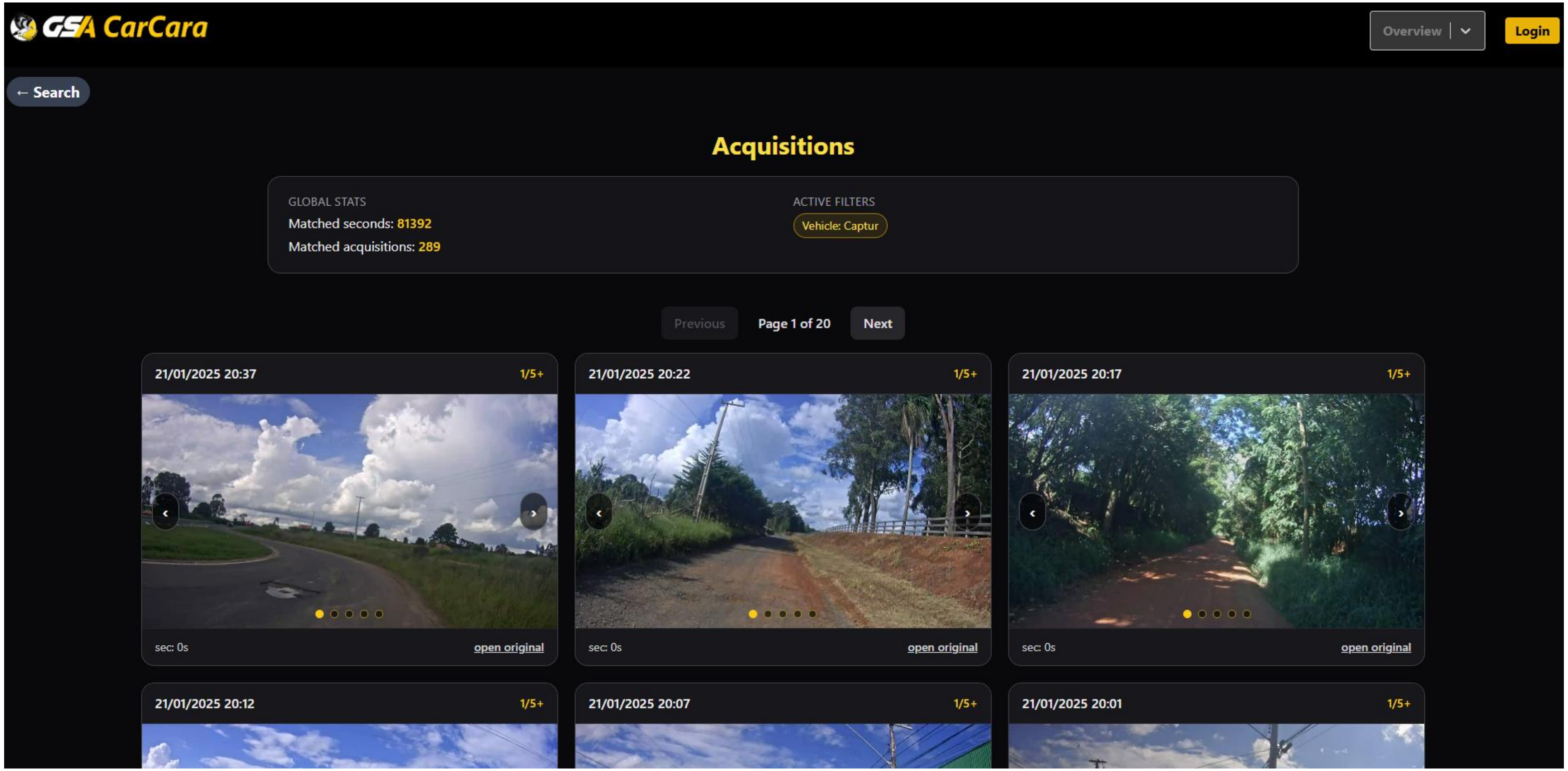
https://carcara.onrender.com/#/search?b.vehicle=Captur

Toronto Metropolitan University

UTFPR-PG

8. Our Contributions in LLMs for the Automotive

- **CarCara**
  - Powered by AI

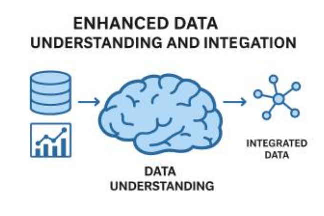


https://carcara.onrender.com/#/search?b.vehicle=Captur

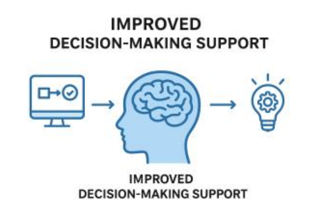
9. Potential Use Cases

1. Sensor Data Fusion and Conflict Resolution
2. Autonomous Driving Decision Support
3. Natural Language Human-Vehicle Interaction
4. Predictive Maintenance and Diagnostics
5. Cybersecurity Threat Detection and Response
6. Test Case Generation for ADAS/AV Validation
7. Driver Behavior Monitoring and Alerts
8. Automated Documentation and Compliance Reports
9. Fleet Management and Operational Optimization
10. Context-Aware In-Vehicle Assistants

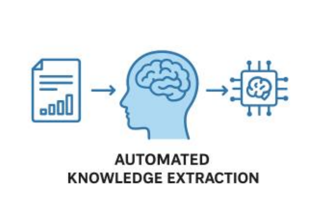
10. Expected Benefits



Enhanced Data Understanding and Integration

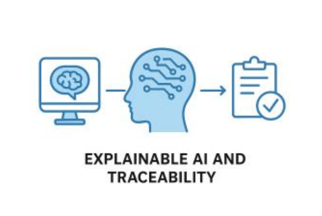


Improved Decision-Making Support



Automated Knowledge Extraction

Accelerated Validation and Verification



Explainable AI and Traceability

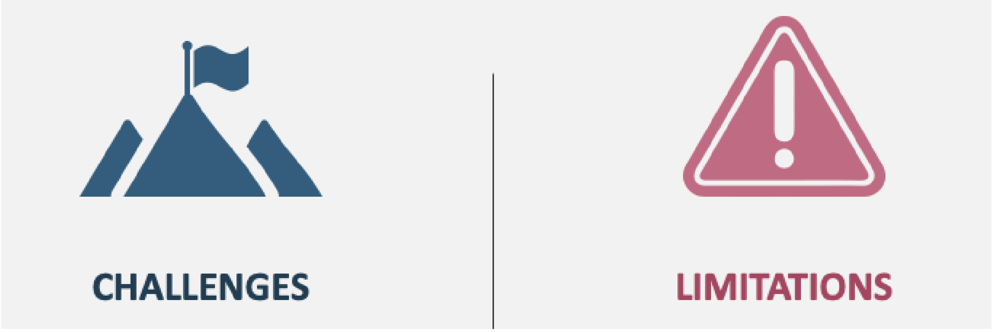
Enhanced Simulation and Training

Scalability and Adaptability

Toronto Metropolitan University

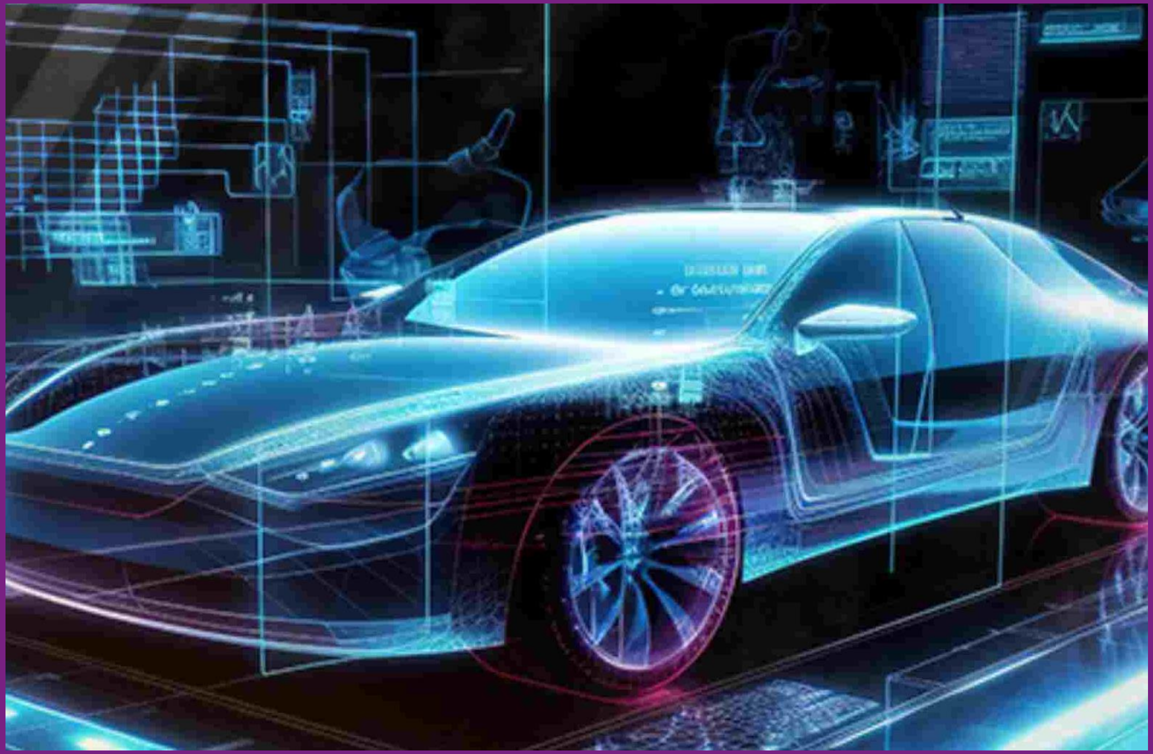
11. Challenges and Limitations

Challenge Area	Key Limitation
Data Complexity	Heterogeneous, high-dimensional, temporal multimodal data
Semantic Understanding	Ambiguity, lack of domain knowledge, limited context reasoning
Computational Requirements	Large model size, inference latency, edge deployment constraints
Data Annotation	High labeling cost, rare events, noisy sensor data
Interpretability	Black-box outputs, limited explainability in safety-critical applications
Generalization	Domain shift, rare events, multi-agent interactions
System Integration	Requires hybrid architectures, complex validation pipelines
Regulatory and Security	Privacy, cybersecurity risks, compliance with safety standards



12. Future Perspectives

The future of LLMs in autonomous and semi-autonomous vehicle datasets lies in their ability to provide context-aware, explainable, and real-time reasoning across complex multi-modal data, enhance testing and safety validation, enable fleet intelligence, and support regulatory compliance. With continued research in efficient deployment and multi-modal integration, LLMs could transform how autonomous systems interpret, reason, and act on their data.



Toronto Metropolitan University

13. References

[1] S. Smith, J. Doe, and A. Johnson, "Enhancing Autonomous Vehicle Decision-Making with Large Language Models," IEEE Transactions on Intelligent Vehicles, vol. 8, no. 3, pp. 123-135, Mar. 2025.DOI: 10.1109/TIV.2025.1234567

[2] M. Zhang, L. Wang, and H. Li, "Interpretable Deep Learning Models for Autonomous Driving Systems," IEEE Transactions on Vehicular Technology, vol. 74, no. 5, pp. 2345-2357, May 2025.DOI: 10.1109/TVT.2025.9876543

[3] R. Patel, S. Kumar, and P. Gupta, "Natural Language Processing Techniques for Autonomous Vehicle Data Analysis," IEEE Access, vol. 13, pp. 4567-4579, Apr. 2025.DOI: 10.1109/ACCESS.2025.1234567

[4] A. Lee, B. Kim, and C. Park, "Benchmarking Large Language Models for Autonomous Driving Datasets," IEEE Transactions on Neural Networks and Learning Systems, vol. 36, no. 7, pp. 1234-1245, Jul. 2025.DOI: 10.1109/TNNLS.2025.1234567

[5] J. Chen, Y. Zhao, and X. Liu, "Data Fusion Strategies for Interpreting Sensor Data in Autonomous Vehicles," IEEE Transactions on Intelligent Transportation Systems, vol. 26, no. 2, pp. 678-690, Feb. 2025.DOI: 10.1109/TITS.2025.1234567

13. References

[6] K. Nguyen, D. Tran, and E. Hoang, "Real-Time Natural Language Interfaces for Semi-Autonomous Vehicles," IEEE Transactions on Human-Machine Systems, vol. 55, no. 4, pp. 123-135, Apr. 2025.DOI: 10.1109/THMS.2025.1234567

[7] F. Wang, G. Zhang, and H. Xu, "Explainable AI Models for Autonomous Driving Decision Systems," IEEE Transactions on Artificial Intelligence, vol. 6, no. 1, pp. 45-57, Jan. 2025.DOI: 10.1109/TAI.2025.1234567

[8] L. Yang, M. Li, and J. Zhang, "Leveraging Transformer Models for Interpreting Autonomous Vehicle Datasets," IEEE Transactions on Robotics, vol. 41, no. 3, pp. 123-135, Mar. 2025.DOI: 10.1109/TRO.2025.1234567

[9] P. Singh, Q. Zhang, and R. Lee, "Ethical Implications of Large Language Models in Autonomous Vehicle Systems," IEEE Transactions on Computational Social Systems, vol. 12, no. 2, pp. 234-245, Feb. 2025.DOI: 10.1109/TCSS.2025.1234567

[10] S. Gupta, T. Sharma, and V. Mehta, "Human-AI Collaboration in Semi-Autonomous Vehicles: A Natural Language Approach," IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 55, no. 6, pp. 1234-1246, Jun. 2025.DOI: 10.1109/TSMC.2025.1234567


Q
&
A