Journal of Contemporary Educational Research, 2026, Volume 10, Issue 6
http://ojs.bbwpublisher.com/index.php/JCER
Online ISSN: 2208-8474
Print ISSN: 2208-8466

Teaching Large Model Methods in Transportation Graduate Education: A Scenario-Driven Curriculum Reform

Qingying Lai[1]*, Xinyue Xu[1], Chunyan Li[1], Shasha Liu[1], Liujiang Kang[2]

[1]School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China
[2]School of Systems Science, Beijing Jiaotong University, Beijing 100044, China

***Corresponding author:** Qingying Lai, qylai@bjtu.edu.cn

Abstract: Generative artificial intelligence is reshaping transportation data analysis, knowledge service, and decision support, but graduate courses in transportation programs still need a systematic way to connect large model methods with research training. This paper presents a curriculum reform design for Application and Practice of Large Models in Transportation, a 32-hour elective course for first-year master's students. The reform responds to four problems: fragmented method learning, weak task modeling for multi-source transportation data, insufficient evidence for project-based outputs, and inadequate training in trustworthy use. Based on outcome-based education, the course reconstructs learning objectives, teaching modules, scenario-based projects, and assessment evidence. Prompt design, retrieval-augmented generation, agent-based tool use, experimental evaluation, and academic norms are organized into an integrated pathway. The design emphasizes reproducible project records, data cards, model evaluation cards, and system risk cards. It provides a practical framework for cultivating transportation problem formulation, intelligent application development, research reporting, and trustworthy artificial intelligence awareness.

Keywords: Research training; Knowledge retrieval; Intelligent agents; Outcome-based assessment; Reproducible evidence; Trustworthy artificial intelligence

Online publication: July 15, 2026

1. Introduction

Generative artificial intelligence and large model technologies are changing how transportation data are interpreted, how operational knowledge is retrieved, and how decision-support tools are prototyped. China's *New Generation Artificial Intelligence Development Plan* and the national Artificial Intelligence Plus policy both emphasize deep integration between artificial intelligence and economic and social sectors [1,2]. The implementation opinions on Artificial Intelligence Plus Transportation further propose building integrated transportation large models, promoting intelligent agents, and applying artificial intelligence in typical transportation scenarios [3]. For graduate education, these policy directions mean that transportation programs need courses that connect computational methods with transportation problems, not only general tool demonstrations.

Higher education policy also supports this curricular transformation. The Ministry of Education's artificial intelligence innovation plan for universities calls for improving artificial intelligence talent cultivation and interdisciplinary capacity in higher education [4], while the Artificial Intelligence Plus Education action plan stresses intelligent technology integration across educational processes [5]. The policy on classified development of academic and professional degree graduate education requires stronger differentiation and better support for academic innovation and practical innovation [6]. These requirements are relevant to first-year master's students in transportation-related programs because this stage is usually when students move from course learning to research topic formation, laboratory work, and participation in engineering-oriented projects.

At the methodological level, retrieval-augmented generation (RAG) uses external knowledge retrieval to support generated responses [7]. ReAct, or Reasoning and Acting, provides a way to combine reasoning steps with tool use in agent-based applications [8]. Outcome-based education (OBE) emphasizes the alignment of learning outcomes, teaching activities, and assessment evidence [9], and constructive alignment further supports the coherence between intended learning outcomes, learning tasks, and assessment methods [10]. For large model education, recent studies show that structured prompt training can improve students' data analysis and programming performance [11], but large language models in education still require attention to fact-checking, bias, human oversight, and responsible use [12]. Therefore, a graduate course on transportation large-model applications should teach students to define transportation problems, organize data, build and evaluate application prototypes, and report trustworthy results.

This paper develops a scenario-driven curriculum reform for Application and Practice of Large Models in Transportation. The course is designed as a 32-hour professional elective for first-year master's students. It aims to form a research-oriented pathway that integrates transportation scenarios, large model methods, reproducible project evidence, and trustworthy artificial intelligence governance.

2. Course context and reform problems

2.1. Course orientation

The course is positioned for first-year master's students in transportation planning and management, transportation engineering, and related programs. Students usually have prior exposure to programming, transportation theory, operations research, and transportation data analysis, but their experience with RAG, agent-based tool calling, model evaluation, and academic reporting is uneven. The course is therefore not a general artificial intelligence literacy course. It is also not a software development course that only asks students to build a working interface. Its central task is to help students transform transportation research or practice needs into verifiable large model application tasks.

Four course objectives are set according to this orientation. Objective 1 is to understand the principles, applicability, and limitations of large language models, RAG, agents, tool calling, and workflow orchestration in transportation settings. Objective 2 is to complete problem definition, data organization, knowledge-base construction, prompt design, tool connection, and application development around transportation tasks. Objective 3 is to build a transportation large-model application prototype with reproducible experimental records, testing results, and a research-style report. Objective 4 is to analyze reliability, safety, privacy protection, academic norms, and engineering ethics in model applications and to propose corresponding governance measures.

2.2. Reform problems

The course reform responds to four connected problems. First, method learning is fragmented. Students may read papers, call model interfaces, and discuss transportation cases separately, but they often lack a pathway for turning a research question into an application task. Second, task modeling for multi-source transportation data is weak. Textual records, structured data, time-series data, multimodal data, and operational rules require different data descriptions, constraints, and evaluation criteria. Third, project outputs may remain at the demonstration level. A graduate course needs evidence on baselines, parameter choices, failure cases, reproducibility, and result limitations. Fourth, trustworthy use and academic norms are often not trained as continuous requirements. Transportation applications may involve location data, passenger comments, equipment status, operational safety, and management decisions, so students need explicit training in data desensitization, model hallucination checks, citation norms, human review, and accountability boundaries. **Table 1** maps these problems to corresponding strategies.

Table 1. Course reform problems and corresponding strategies

Reform problem	Corresponding strategy	Expected improvement
Fragmented method learning	Organize the course through transportation research questions and scenario-based tasks.	Improve students' ability to translate professional needs into verifiable model tasks.
Weak task modeling for multi-source transportation data	Add data governance, task definition, and evaluation criteria as core modules.	Align data sources, model use, and evaluation indicators.
Project outputs staying at the demonstration level	Require experiment records, baseline comparison, reproducibility checks, and research-style reporting.	Move course projects from working prototypes to evidence-supported arguments.
Insufficient trustworthy use and academic norms	Embed risk labeling, human handover, data desensitization, and citation requirements throughout the project.	Strengthen reliable, safe, and responsible model application awareness.

3. Materials and methods

3.1. Design principles and evidence sources

The curriculum design uses three principles. The first principle is scenario-driven learning. Transportation scenarios are selected as the starting point because graduate students need to connect large model methods with research problems and engineering contexts. The second principle is research-oriented evidence. Each project must provide process records, data descriptions, evaluation results, and risk analysis instead of only a runnable demonstration. The third principle is trustworthy application. Trustworthy artificial intelligence research emphasizes reliability, robustness, privacy, fairness, security, transparency, and accountability across the artificial intelligence lifecycle [13]. These dimensions are translated into classroom requirements such as unified testing, error analysis, risk labeling, and human handover.

The design evidence comes from the course syllabus, transportation graduate training needs, national and sectoral artificial intelligence policy documents, and published research on RAG, agents, OBE, generative artificial intelligence in education, trustworthy artificial intelligence, and classroom assessment. Because this manuscript reports a reform design rather than completed outcome data, claims about effectiveness are stated as intended or expected instructional effects. Formal learning-outcome evidence should be collected after course implementation.

3.2. Course modules and class-hour allocation

The course has 32 class hours and 2 credits. Eight modules are arranged from conceptual understanding to research reporting. Prompt design is treated as the practice of writing prompts or prompt templates that define roles, tasks, context, constraints, and output formats. Schema is used as a structured output specification, especially when model responses need to be checked, compared, or connected to downstream tools. **Table 2** presents the class-hour allocation and main outputs.

Table 2. Course modules and class-hour allocation

Module	Main content	Hours	Teaching form	Main output
Foundations and transportation applications	Model principles, capability boundaries, typical transportation cases, and hallucination risks	4	Lecture and case discussion	Scenario analysis record
Transportation data governance and task modeling	Text, structured data, time-series data, multimodal data, task definition, and evaluation indicators	4	Lecture and seminar	Task and data description
Prompt design and task interaction	Role setting, task decomposition, context organization, output constraints, error analysis, and schema-based outputs	4	Computer-based practice	Prompt templates and comparison notes
Transportation knowledge base and RAG	Document segmentation, vectorization, retrieval strategy, reranking, citation traceability, and answer verification	6	Lecture and practice	Knowledge base and RAG testing record
Agents and tool calling	Task planning, tool access, transportation data query, computation, visualization, and log recording	4	Lecture and case demonstration	Tool-calling workflow and log
Transportation scenario application design	Travel knowledge service, review-text analysis, equipment maintenance support, prediction-support analysis, logistics classification, and dispatching support	4	Group seminar	Project plan and technical route
Model application evaluation and optimization	Accuracy, completeness, stability, interpretability, latency, cost, and user experience	3	Practice and seminar	Evaluation report and revision record
Comprehensive project and research report	Solution design, testing, risk analysis, course paper or project report, and defense	3	Project work and defense	Project report and defense materials

3.3. Scenario-based project organization

Scenario-based projects are designed to match graduate students' research readiness without turning the course into a large engineering development task. Students may select a public, desensitized, or teaching-simulated dataset, define a transportation task, implement a small prototype, and complete an evaluation report. The scenarios emphasize problem formulation, data explanation, method choice, and evidence quality. **Table 3** lists representative project scenarios.

Table 3. Scenario-based project design

Scenario	Task focus	Recommended method	Expected output
Travel knowledge question answering	Answer service questions on policies, routes, rules, or travel guidance with traceable sources.	RAG and citation checking	Knowledge base, test set, and answer evaluation report
Passenger review-text analysis	Summarize service complaints, identify themes, and support service-improvement suggestions.	Prompt templates and structured output schema	Theme labels, prompt comparison, and error analysis

Equipment maintenance support	Retrieve maintenance rules and assist fault-description interpretation.	RAG and human review workflow	Maintenance knowledge card and risk card
Traffic prediction support analysis	Explain data trends, understand indicators, and assist analysis of prediction results produced by existing models or datasets.	Data cards, model explanation prompts, and evaluation notes	Data trend explanation and auxiliary analysis report
Logistics order classification	Classify service requests or order descriptions and explain classification reasons.	Few-shot prompting and rule-constrained output	Classification prompt, confusion cases, and revision record
Vehicle dispatching decision support	Generate alternative dispatching suggestions under explicit constraints and compare their risks.	Agent workflow and tool calling	Workflow log, constraint checklist, and defense presentation

3.4. Ethics statement

Ethics approval was not required for this manuscript because it reports a curriculum reform design and does not involve the collection or analysis of identifiable student data, human-subject experimental data, or intervention outcome data. If the course is formally evaluated in later teaching cycles, institutional requirements for educational research ethics and informed consent should be followed.

4. Results

4.1. Scenario-driven teaching framework

The expected result of the reform design is a scenario-driven research-oriented teaching framework. The framework links course input, research training pathway, evidence-chain outputs, and trustworthy governance. **Figure 1** illustrates the framework. The figure is intentionally presented in grayscale to meet the academic manuscript style and to avoid relying on color for meaning.

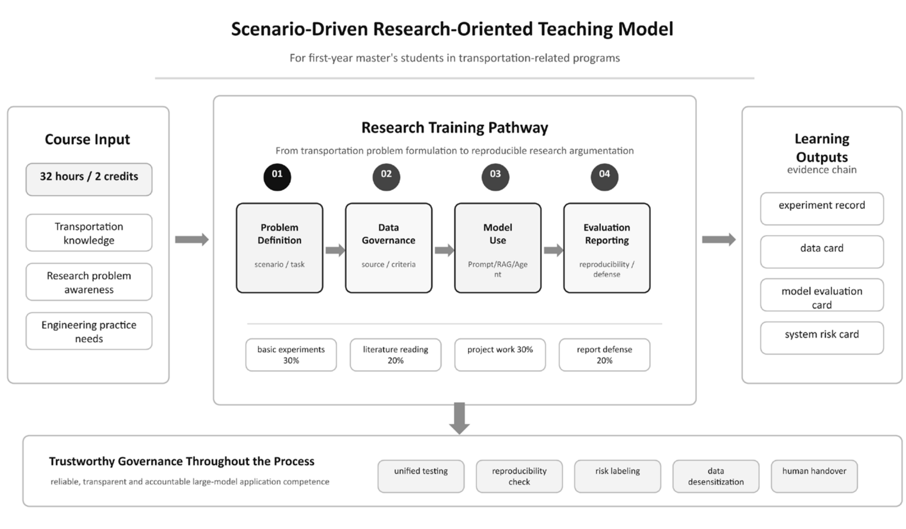

Figure 1. Scenario-driven research-oriented teaching model for large model methods in transportation graduate education

4.2. Case-based project workflow

A typical project follows four steps. First, students define a transportation problem and specify users, data sources, constraints, and expected outputs. Second, they prepare a data card that records data source, preprocessing, desensitization, limitations, and permitted use. Third, they implement the method through prompt templates, RAG, agents, or tool calling, and they keep experiment records. Fourth, they conduct evaluations and write a report that includes baselines, failure cases, reproducibility information, risks, and human handover rules. This workflow prevents students from treating model calls as final answers and requires them to defend the evidence behind their outputs.

The evidence chain consists of four documents. The experiment record card captures prompts, parameter settings, model versions, tool calls, and test cases. The data card records data source, preprocessing, missing values, sensitive information, and usage constraints. The model evaluation card records test design, evaluation indicators, comparison results, failure cases, and revision actions. The system risk card records privacy risk, hallucination risk, safety risk, academic integrity risk, and the planned human intervention mechanism. This evidence-chain design is consistent with formative assessment, which emphasizes feedback and learning process evidence rather than only final scores [14].

4.3. Assessment and objective attainment

Assessment is organized around both course scores and objective attainment. The course score consists of basic experiments, literature reading, and seminars, comprehensive project implementation, and project report with defense. **Table 4** shows how each component supports the four course objectives.

Table 4. Assessment components and supported course objectives

Assessment component	Score weight	Main evidence	Supported objectives
Basic experiments	30%	Prompt comparison, RAG testing record, agent workflow log, and short experiment notes	Objectives 1 and 2
Literature reading and seminar	20%	Reading notes, method comparison, and seminar participation	Objectives 1 and 4
Comprehensive project implementation	30%	Prototype, data card, evaluation card, reproducibility record, and risk card	Objectives 2, 3, and 4
Project report and defense	20%	Research-style project report, defense slides, question responses, and revision record	Objectives 3 and 4

Table 5 further describes internal support weights for objective attainment. These weights are used only inside each course objective to judge which assessment evidence supports the objective most strongly. They are not additional course-score percentages and should not be added to the score weights in **Table 4**.

Table 5. Internal support weights for course objective attainment

Course objective	Supporting evidence	Internal support weight
Objective 1: understand principles and boundaries	Basic experiments; literature reading and seminar	60%; 40%
Objective 2: complete task modeling and application development	Basic experiments; comprehensive project implementation	40%; 60%
Objective 3: build prototype and research-style evidence	Comprehensive project implementation; project report and defense	55%; 45%

Objective 4: analyze trustworthy use and academic norms	Literature reading and seminar; comprehensive project implementation; project report and defense	30%; 40%; 30%

5. Discussion

5.1. Pedagogical significance

The reform has three pedagogical implications. First, it adapts large model education to first-year graduate students in transportation-related programs. These students already have professional foundations, but they still need structured support for research task formulation and responsible model use. Second, it changes project-based learning from interface demonstration to evidence-based research training. The required records and cards make the process auditable and help students report methods and results in an academic form. Third, it embeds trustworthy artificial intelligence into ordinary project work. Students do not learn ethics as a separate lecture only; they must connect privacy, reliability, citation, risk labeling, and human handover to their project design.

5.2. Implementation safeguards

Several safeguards are necessary for implementation. Teachers should provide public or desensitized teaching datasets, baseline test cases, prompt templates, and reporting templates at the beginning of the course. The difficulty of projects should be controlled so that students can complete a verifiable prototype within one semester. Group work should be paired with individual evidence submission, because model outputs and project reports need traceable individual contributions. The defense should include questions on data source, baseline design, failure cases, risk control, and reproducibility rather than only asking whether the prototype can run.

5.3. Limitations

This paper reports a curriculum reform design rather than empirical evidence from a completed teaching cycle. Therefore, it does not claim measured improvement in learning outcomes. After formal implementation, the course should collect objective-attainment data, project-quality rubrics, student feedback, defense records, and follow-up evidence related to research topic development. These data can be used to test whether the reform improves transportation problem formulation, model application ability, research reporting quality, and trustworthy artificial intelligence awareness.

6. Conclusion

Application and Practice of Large Models in Transportation is designed as a 32-hour graduate elective course that connects large model methods with transportation research and engineering practice. The proposed reform reconstructs course objectives, modules, scenario-based projects, assessment evidence, and trustworthy governance around the needs of first-year master's students. The central innovation is a scenario-driven research-oriented teaching model in which transportation problems, data governance, prompt and RAG design, agent-based tool use, reproducible evaluation, and risk control are organized as one learning pathway. The design is expected to support graduate students in forming transportation problem-modeling ability, intelligent application development ability, research-style reporting ability, and responsible artificial intelligence awareness. Future work should evaluate the course through formal teaching data and refine the assessment rubrics according to implementation evidence.

Disclosure statement

The authors declare no conflict of interest.

References

[1] The State Council of the People's Republic of China, 2017, Notice on Issuing the New Generation Artificial Intelligence Development Plan, viewed June 4, 2026, https://www.gov.cn/zhengce/zhengceku/2017-07/20/content_5211996.htm.

[2] The State Council of the People's Republic of China, 2025, Opinions on Deeply Implementing the Artificial Intelligence Plus Action, viewed June 4, 2026, https://www.gov.cn/zhengce/content/202508/content_7037861.htm.

[3] Ministry of Transport of the People's Republic of China, National Development and Reform Commission, Ministry of Industry and Information Technology, et al., 2025, Implementation Opinions on Artificial Intelligence Plus Transportation, viewed June 4, 2026, https://xxgk.mot.gov.cn/2020/jigou/kjs/202509/t20250925_4177256.html.

[4] Ministry of Education of the People's Republic of China, 2018, Notice on Issuing the Artificial Intelligence Innovation Action Plan for Institutions of Higher Education, viewed June 4, 2026, https://www.moe.gov.cn/srcsite/A16/s7062/201804/t20180410_332722.html.

[5] Ministry of Education of the People's Republic of China, National Development and Reform Commission, Ministry of Science and Technology, et al., 2026, Notice on Issuing the Artificial Intelligence Plus Education Action Plan, viewed June 4, 2026, https://www.moe.gov.cn/srcsite/A16/s3342/202604/t20260410_1433240.html.

[6] Ministry of Education of the People's Republic of China, 2023, Opinions on Further Promoting Classified Development of Academic Degree and Professional Degree Graduate Education, viewed June 4, 2026, https://hudong.moe.gov.cn/srcsite/A22/moe_826/202312/t20231218_1095043.html.

[7] Lewis P, Perez E, Piktus A, et al., 2020, Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. Advances in Neural Information Processing Systems, 33: 9459–9474.

[8] Yao S, Zhao J, Yu D, et al., 2023, ReAct: Synergizing Reasoning and Acting in Language Models. The Eleventh International Conference on Learning Representations.

[9] Spady WG, 1994, Outcome-Based Education: Critical Issues and Answers, American Association of School Administrators, Arlington, VA.

[10] Biggs J, Tang C, 2011, Teaching for Quality Learning at University, 4th ed., Open University Press, Maidenhead.

[11] Garg A, Soodhani KN, Rajendran R, 2025, Enhancing Data Analysis and Programming Skills through Structured Prompt Training: The Impact of Generative AI in Engineering Education. Computers and Education: Artificial Intelligence, 8: 100380. https://doi.org/10.1016/j.caeai.2025.100380.

[12] Kasneci E, Sessler K, Kuechemann S, et al., 2023, ChatGPT for Good? On Opportunities and Challenges of Large

Language Models for Education. Learning and Individual Differences, 103: 102274. https://doi.org/10.1016/j.lindif.2023.102274.

[13] Li B, Qi P, Liu B, et al., 2023, Trustworthy AI: From Principles to Practices. ACM Computing Surveys, 55(9): 1–46. https://doi.org/10.1145/3555803.

[14] Black P, Wiliam D, 1998, Assessment and Classroom Learning. Assessment in Education: Principles, Policy and Practice, 5(1): 7–74. https://doi.org/10.1080/0969595980050102.