Renewable and Sustainable Energy Reviews 241 (2026) 117243

Contents lists available at ScienceDirect

Renewable and Sustainable Energy Reviews

journal homepage: www.elsevier.com/locate/rser

Large language models for transportation energy integration: A functional framework and future directions

Yongjian Chen [a], Shiqi Shawn Ou [a,b,*], Zhilong Lv [a], Zhifeng Yang [a], Lin Gu [c], Wei Ma [d]

[a] School of Future Technology, South China University of Technology, Guangzhou, Guangdong, 511442, China
[b] Guangdong Artificial Intelligence and Digital Economy Laboratory (Guangzhou), Guangzhou, Guangdong, 510335, China
[c] Research Institute of Electrical Communication, Tohoku University, Sendai, Miyagi, 980-8577, Japan
[d] Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hong Kong SAR, China

ARTICLE INFO

Keywords:
Large language models
Transportation energy integration
Intelligent transportation systems
Decision intelligence
Multimodal fusion

ABSTRACT

The deep integration of Transportation Energy Integration (TEI) is a critical pathway for global decarbonization, as the transport and building sectors account for over half of global energy consumption and CO_2 emissions. However, achieving efficient TEI faces severe challenges, including processing vast, heterogeneous, and unstructured data (e.g., policy texts and human behavior); modeling complex nonlinear coupling among traffic, grid, and building systems; and navigating high decision uncertainty arising from driver actions and renewable energy fluctuations. Traditional models struggle with these issues. Large Language Models (LLMs) offer a transformative approach, leveraging powerful natural language understanding, cross-modal fusion, and complex reasoning. They show unique potential to parse unstructured text and simulate complex human behaviors, such as driving and charging choices. LLM–TEI studies emerged in 2023 following the 2022 ChatGPT breakthrough and have grown rapidly since. This review synthesizes 82 studies published up to 2025. It proposes a three-layer functional framework—Perception, Reasoning, and Decision—to characterize the roles of LLMs in TEI systematically. The synthesis demonstrates the expansive capabilities of LLMs across the TEI ecosystem, highlighting key applications spanning from foundational transport perception to cross-domain system optimization, including policy interpretation, behavioral simulation, and real-time scheduling. This review concludes by identifying four priority research directions: developing multimodal perception for understanding the physical world, enabling causal reasoning for high-stakes decisions, fostering continuous learning for dynamic adaptation, and designing lightweight architectures for edge intelligence. This review provides a comprehensive roadmap for advancing smarter, more resilient TEI systems.

List of abbreviations

AI	Artificial Intelligence
CoT	Chain-of-Thought
EVs	Electric Vehicles
ITS	Intelligent Transportation Systems
LLMs	Large Language Models
NLP	Natural Language Processing
RAG	Retrieval-Augmented Generation
TEI	Transportation Energy Integration
V2G	Vehicle-to-Grid
V2X	Vehicle-to-Everything

1. Introduction

The United Nations Environment Programme highlighted in its *2024 Emissions Gap Report* that global greenhouse gas emissions reached a record high in 2024. Without urgent and substantial action by countries to reduce emissions, the 1.5°C temperature control target set by the *Paris Agreement* is unlikely to be achieved [1]. The transportation and building energy sectors, which together account for approximately 55% of global energy consumption and nearly 50% of CO_2 emissions, have been recognized as among the most critical challenges for sustainable development [2]. In the transportation sector, energy demand remains heavily dependent on petroleum, making it a primary source of emissions, while its integration with other energy supply systems has

* Corresponding author. School of Future Technology, South China University of Technology, Guangzhou, Guangdong, 511442, China.
E-mail addresses: sou@scut.edu.cn, oushiqi@pazhou.cn (S.S. Ou).

https://doi.org/10.1016/j.rser.2026.117243
Received 8 December 2025; Received in revised form 3 March 2026; Accepted 28 June 2026
Available online 1 July 2026
1364-0321/

traditionally been weak [3]. However, recent advancements in electric mobility, vehicular connectivity, and the growing share of renewable energy generation are accelerating the convergence of transportation with building energy systems, power grids, and digital ecosystems [4]. This evolution has opened up new pathways for transportation decarbonization [5].

In this context, Transportation Energy Integration (TEI) is widely recognized as a core pathway to address climate challenges and promote the transformation of the energy structure [1]. This concept emphasizes the deep coupling of transportation infrastructure, such as electric vehicles (EVs), charging piles, and traffic networks, with energy systems, including buildings, power grids, and energy storage. The goal is to achieve coordinated optimal scheduling of multi-source resources, thereby enhancing the overall efficiency and sustainability of the system—for instance, Liu et al. [6] explored an innovative integration of energy storage and solar photovoltaic systems in a Beijing bus depot in China. Their findings showed that this integration could reduce the net grid load by 23% and the peak charging load by 8.6%, resulting in a net profit of 64% of the cost. This integration can be implemented across several dimensions: (1) Energy Coordination: Optimizing the charging of EVs with renewable energy sources like solar and wind to lower operational costs and alleviate grid pressure [7]; (2) Infrastructure Coupling: Enabling a two-way coupling of traffic and energy flows through Vehicle-to-Grid (V2G) technology and adaptive traffic signal control to enhance system resilience [8]; (3) Data-Driven Optimization: Utilizing Artificial Intelligence (AI) algorithms to merge transportation, energy, and environmental data for intelligent, real-time scheduling [9]. Collectively, these strategies herald a more efficient, resilient, and sustainable transportation-energy ecosystem.

Despite the promising prospects of TEI, achieving its deep integration faces persistent computational and analytical barriers. Traditional models are deficient in several areas: (1) Knowledge extraction barriers: They struggle to uniformly process heterogeneous information, such as structured grid loads and unstructured policy reports; (2) Dynamic modeling dilemmas: They have difficulty capturing the nonlinear interactions among physical systems, market forces, and uncertain human behavior; (3) Complex decision-making challenges: Optimizing conflicting objectives like cost, user convenience, and energy system stability under dynamic constraints is computationally formidable. To address these barriers, LLMs such as ChatGPT have emerged as transformative technologies. As foundational models trained on massive text and code corpora, they align closely with the primary challenges of TEI. The ability of LLMs to process multimodal information addresses knowledge extraction barriers; their sophisticated reasoning and simulation capabilities address dynamic modeling challenges; and their capacity to understand human intent and generate executable plans offers novel solutions for complex decision-making. For example, LLMs can not only interpret, predict, and respond to complex scenarios in transportation networks [10], but also provide powerful support for data analysis and prediction [11], smart charging management [12], energy dispatch optimization [13], and policy scenario simulation [14] in the TEI domain by integrating multi-source historical data and domain knowledge.

Although several reviews have explored the application of LLMs in related fields, as shown in Table 1, they have critical limitations that create a significant research gap. Existing works are often: (1) domain-specific and narrow in perspective, focusing on the technical details of isolated applications such as autonomous driving while lacking a systemic viewpoint; (2) overly prospective, concentrating on the future potential of LLMs in the power grid or EV sectors rather than systematically analyzing existing empirical studies; and (3) lacking an integrated framework that connects the foundational capabilities of LLMs, such as knowledge reasoning and dynamic modeling, with the interconnected challenges of TEI at the technical, economic, and policy levels. Therefore, to the best of our knowledge, this paper is the first to systematically review the applications of LLMs from an integrated perspective of transportation and energy.

To fill the gap mentioned above, this review aims to provide a comprehensive and systematic overview of the technological evolution, application scenarios, and core challenges of LLMs in TEI. The core contributions of this paper include: (1) Providing a holistic, TEI-centric literature review that integrates multiple domains, such as transportation, building energy, and power grids, into a unified analytical framework through extensive literature screening. (2) Establishing a novel three-layer functional framework, including perception and understanding, prediction and reasoning, and decision-making and optimization, to systematically elucidate the diverse roles LLMs play in TEI as shown in Fig. 1. (3) Identifying the key challenges of LLMs in TEI and proposing a clear future research roadmap, aiming to guide researchers and practitioners in this emerging area to promote its future development.

The structure of this paper is as follows. Section 2 first establishes the conceptual foundation by defining TEI, outlining its core challenges, and detailing the key characteristics of LLMs that make them an ideal enabling technology. Section 3 then presents our review methodology, detailing the literature selection process and establishing the functional

Table 1
Overview of existing survey on LLMs applied to transportation energy integration.

Survey	Year	TEI system	Contributions	No. of Included Studies
[15]	2024	EV energy consumption prediction	Summarizes traditional energy consumption estimation models and explores the application of machine learning, including the potential of neural networks, for energy consumption prediction.	Not defined
[16]	2024	Intelligent transportation systems (ITS)	Conducts a quantitative analysis of the need for large models for ITS and AV (Autonomous Vehicles), summarizes their applications in this field, and explores the deployment technologies for universities.	210
[10]	2025	ITS	Explores the applications of LLMs in ITS, including traffic flow prediction, vehicle detection and classification, autonomous driving, traffic sign recognition, and pedestrian detection.	Not defined
[17]	2025	ITS	Explores how LLMs address the EV challenges by optimizing routes, speeds, and behaviors to reduce energy consumption and emissions.	Not defined
[18]	2024	Electric energy sector	Investigate the capabilities and limitations of LLMs for enhancing power industry operations.	Not defined
[19]	2025	Smart gird	Explores the applications of LLMs in power system operations and the key technical challenges.	30
[20]	2025	Building energy applications	Explores the role of LLMs in the field of building energy systems and proposes a roadmap for enhancing the practicality of LLMs based on contextual learning, domain-specific fine-tuning, retrieval-augmented generation, and multimodal integration.	Not defined
[21]	2025	Building energy efficiency	Explores the applications, impacts, and potential of large language models in the building energy sector, with a particular focus on energy efficiency and decarbonization studies.	Not defined
[22]	2025	Energy systems	Divides into five areas of data generation, prediction, situational awareness, modeling, and optimal decision-making to explore the applications of deep generative models in energy systems.	228
This review	**Transportation Energy Integration**		**Explore the transformative potential of LLMs in advancing intelligent and integrated transportation-energy systems through a comprehensive analysis of their roles across key tasks and application domains.**	**82**

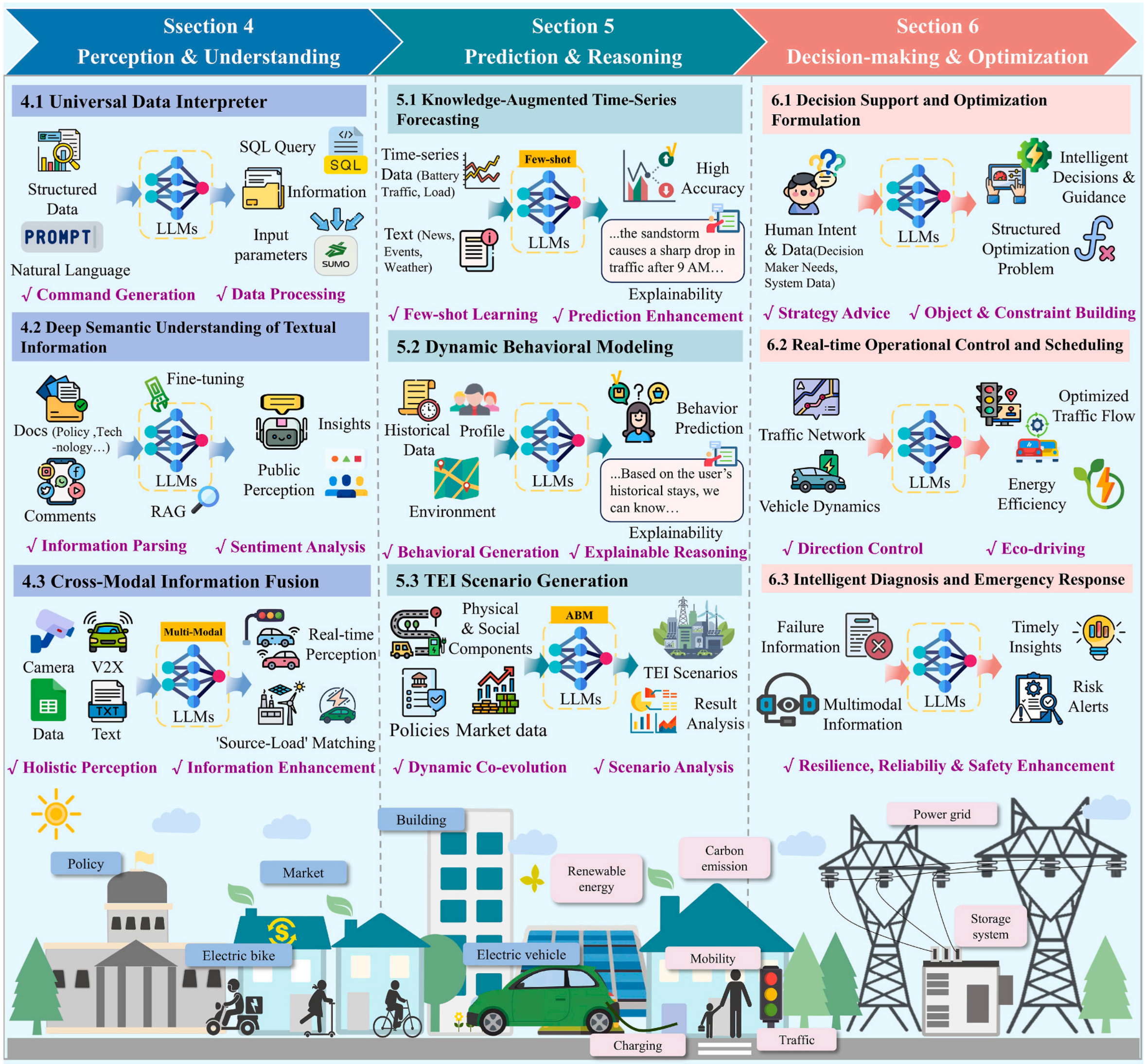

Fig. 1. The framework for the application of LLMs in TEI. This paper explores the key advances of LLMs in TEI across three tasks: perception and understanding, prediction and reasoning, and decision-making and optimization.

classification framework. The main body of the review, Sections 4-6, systematically explores the three primary functional roles of LLMs: Perception and Understanding (Section 4), Prediction and Reasoning (Section 5), and Decision-making and Optimization (Section 6). Following this analysis, Sections 7 and 8 delve into the critical challenges and future research directions for the field. Section 9 summarizes the key contributions and findings of the review.

2. Foundations for LLM-Driven Transportation Energy Integration

2.1. *Breaking barriers: Why transportation energy integration demands LLMs*

TEI is not merely an aggregation of transportation and energy systems, but a deeply coupled, complex socio-technical system. It aims to achieve deep decarbonization and sustainable development within the transportation sector by integrating transportation infrastructure, vehicles, and diversified energy networks [1]. However, the realization of an efficient and intelligent TEI faces a series of fundamental challenges, such as the barriers to knowledge extraction, the complexities of dynamic modeling, and the difficulties in robust decision-making, which in turn necessitate capabilities from next-generation intelligent technologies across three core dimensions: perception and understanding, prediction and reasoning, and decision-making and optimization.

The barriers to knowledge extraction. TEI systems are characterized by vast volumes of heterogeneous information, encompassing structured time-series data, such as grid load and traffic flow, as well as unstructured text, including policy documents, technical reports, and social media sentiment. Traditional models lack a unified framework for perceiving, parsing, and extracting knowledge from this multimodal data, creating information silos that hinder a comprehensive understanding of the system's overall state.

The complexities of dynamic modeling. The internal dynamics of TEI are replete with nonlinear interactions involving physical principles, market mechanisms, policy interventions, and highly uncertain human behaviors (e.g., charging choices, driving habits). Accurately modeling such complex system dynamics and making reliable inferences and predictions about future states is a task where traditional mathematical and statistical models often fall short.

The difficulties in robust decision-making. The system must continually strive for a dynamic balance among multiple conflicting objectives, including energy efficiency, operational cost, user convenience, and grid stability. Translating high-level, often ambiguous, human intent (e.g., “achieve the lowest cost while ensuring safety”) into executable, optimal scheduling and control strategies under high-dimensional, dynamic constraints is a significant challenge for conventional optimization methods.

These three challenges—the barriers to knowledge extraction, the complexities of dynamic modeling, and the difficulties in robust decision-making—constitute the core bottlenecks hindering the advancement of TEI. They directly correspond to the urgent need for intelligent systems capable of advanced perception, reasoning, and optimization. The advent of LLMs, therefore, offers a promising and systematic approach to overcoming these fundamental barriers.

2.2. From text to infrastructure: how LLMs enable intelligent transportation-energy systems

LLMs represented by pioneering models such as the GPT series [23] and the Llama series [24], are trained on vast amounts of data to acquire powerful capabilities in contextual understanding, reasoning, and generation [25]. To adapt these models for the complex demands of specific domains, a series of key technical paradigms has emerged. These paradigms form the foundational capabilities that enable LLMs to address the three core challenges of TEI, as illustrated in Fig. 2.

Pre-training and Fine-tuning. Pre-training endows LLMs with extensive world knowledge, while domain-specific fine-tuning adapts them to comprehend the specialized terminology and texts within TEI, forming the basis for effective knowledge extraction from multi-source information [26].

Prompt Engineering. Structured prompts, particularly advanced techniques like Chain-of-Thought (CoT), are designed to guide LLMs in performing step-by-step, logically rigorous reasoning on complex TEI scenarios, such as predicting policy impacts or analyzing accident causal chains [27].

Retrieval-Augmented Generation (RAG). This paradigm couples the generative power of LLMs with advanced information retrieval systems. By connecting to external knowledge bases, LLMs can dynamically retrieve and integrate up-to-date information (e.g., policies, news, technical documents) to generate more factually accurate and contextually relevant responses, thereby supporting a comprehensive understanding of the system [28].

Tool Use. This capability enables LLMs to invoke and interact with external specialized tools. By interfacing with software such as transportation simulators (e.g., SUMO [29]]) and energy analysis platforms (e.g., EnergyPlus [30]]), LLMs can bridge high-level language-based reasoning with rigorous physical-world modeling [31]].

Reinforcement Learning from Human Feedback (RLHF). RLHF integrates human evaluative feedback into the model's optimization process. This iterative alignment ensures that LLM-generated content better aligns with human preferences and values, a crucial step toward developing trustworthy decision-making systems.

LLM Agents. The LLM Agent paradigm elevates LLMs from passive language processors to autonomous entities capable of action. These agents can interact with a diverse set of tools, autonomously plan and execute multi-step workflows, thereby bridging the gap from language comprehension to tangible task resolution, and significantly expanding their application potential and value creation in TEI [32].

LLMs empowered by these technical capabilities have already demonstrated transformative potential in diverse fields such as finance [33], medicine [34], psychology [35], law [36,37], and education [38]. Analogously, the capabilities they offer for knowledge extraction, state modeling, and intelligent scheduling position them as an ideal technology for addressing TEI challenges. Although existing reviews have explored LLMs in discrete domains like power systems [18], smart grids [39], and building energy efficiency [21], a comprehensive TEI perspective remains underexplored. Therefore, the subsequent chapters of this paper will follow a logical framework of perception, reasoning, and decision-making to systematically review the specific applications and state-of-the-art advancements of LLMs in TEI.

3. Methodology

3.1. Search and selection strategy

To systematically delineate the research landscape of LLMs in TEI, a comprehensive literature review was conducted. The search scope covered the Web of Science, IEEE Xplore, and Google Scholar databases, spanning the period from ChatGPT's inception in 2022 to 2025. However, given that the application of LLMs in the TEI domain is still an emerging research direction, a preliminary investigation revealed a limited number of publications directly addressing LLM applications that span both transportation and energy. Therefore, the literature

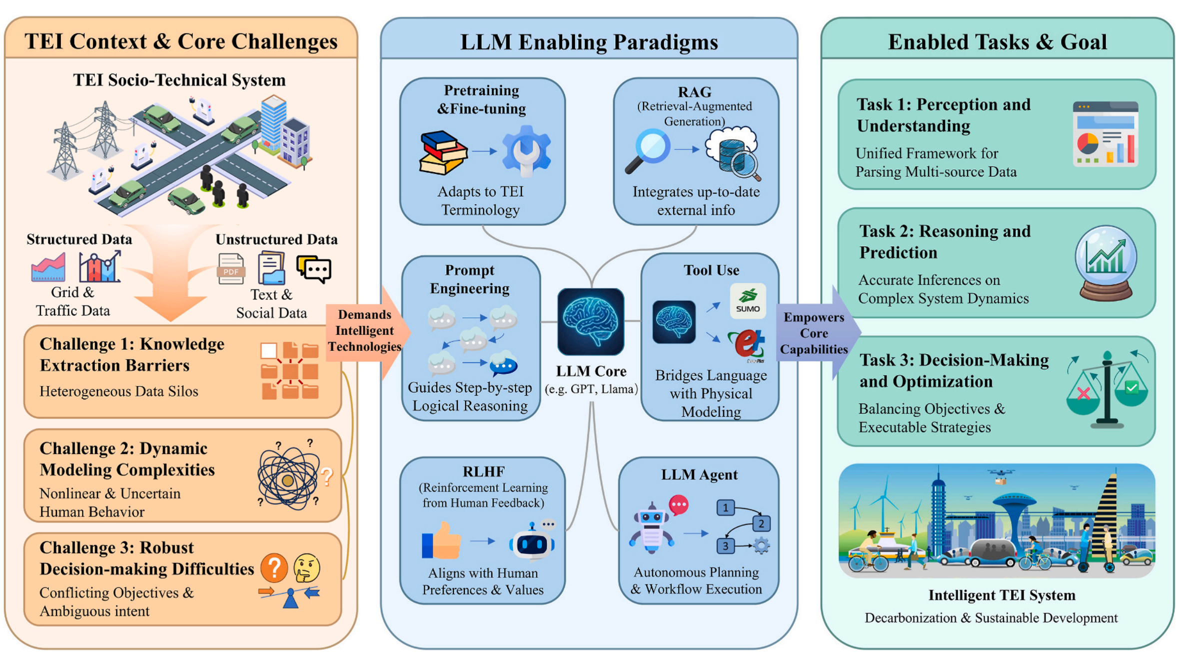

Fig. 2. Foundations for LLM-Driven TEI

selection criteria, while focusing on core integration applications, strategically included single-domain studies based on specific inclusion and exclusion criteria regarding system integration. To avoid including studies with weak relevance, this paper prioritized single-domain research only if it demonstrated functional connectivity to the coupled system: (1) Relevance of Coupling Variables, where the study optimizes a variable (e.g., travel flow) that serves as a critical boundary condition or input for the other domain (e.g., grid load modeling); or (2) Methodological Transferability, where the proposed LLM paradigm offers a direct technical solution to core TEI barriers—specifically knowledge extraction, dynamic behavioral modeling, or complex decision optimization—thereby demonstrating clear utility for the coupled system. Conversely, studies focusing on isolated component physics without system-level data interaction—such as internal combustion engine thermodynamics or photovoltaic material physics—were strictly excluded. Through this rigorous approach, this paper ultimately identified 82 relevant publications to serve as the basis for the in-depth analysis in Sections 4–6.

3.2. Proposed functional framework for LLM-driven TEI

To systematically elucidate the diverse roles LLMs play in TEI, this paper categorizes the 82 identified publications into a three-layer functional framework based on their core contributions. As will be detailed in the subsequent sections, the literature is organized into the following primary functional tasks:

- Perception and Understanding (Section 4): Studies focusing on extracting knowledge and processing multimodal information from diverse, heterogeneous data sources.
- Prediction and Reasoning (Section 5): Research dedicated to modeling complex states, forecasting time-series data, and reasoning about dynamic human behaviors.
- Decision-making and Optimization (Section 6): Applications centered on intelligent scheduling, operational control, and optimization problem formulation.

Furthermore, to capture the interdisciplinary nature of the field, the reviewed applications span multiple core domains, primarily including Transportation, Building Energy, Power Grids, and broader System-level integration. By mapping the literature through the "Perception-Reasoning-Decision" framework, the following sections provide an in-depth analysis of the current state-of-the-art and technological evolution of LLMs within the TEI ecosystem.

4. Perception and understanding: extracting transportation-energy knowledge from multi-source information

The foundational layer of an intelligent TEI system rests upon its ability to perceive and understand the operational environment from a multitude of complex information sources. This section reviews how LLMs are being leveraged to transform raw, heterogeneous data—spanning technical documents and social media—into structured, actionable knowledge, as illustrated by the framework in Fig. 3.

4.1. Universal data interpreter

LLMs are profoundly reshaping the paradigm of how researchers interact with the vast structured data in TEI systems. As a universal data interpreter, the core value of LLMs lies in their capacity to translate natural-language instructions into structured, executable data operations, thereby substantially lowering the technical barrier to data processing and analysis workflow.

The most direct application of this capability is the conversion of natural language queries into executable database commands, especially SQL, enabling "conversational" access to specialized databases. For example, the SQLMorpher framework demonstrates how to leverage an LLM to generate SQL queries for automatically completing complex building energy data transformation tasks, achieving an accuracy of up to 96% [40]. In the field of power dispatching, which demands extremely high reliability, Intelli-Dispatch-SQL constructs an LLM-based agent framework whose generated Text-to-SQL queries far exceed the accuracy of previous specialized models, achieving an execution accuracy of 86.2% [41].

Beyond querying existing databases, LLMs also show strong potential in data processing, particularly automated processing and information structuring. A key application is the extraction of structured information from unstructured text. For example, LLMs have been employed to parse building permit records and construct structured information frameworks covering six electrification technologies, including photovoltaics, EV charging infrastructure, energy storage systems, and electric water heaters [42], or to extract fuel consumption data from various reports to establish a localized carbon emission factor database [43]. Furthermore, this capability extends to broader data processing tasks, such as frameworks like TransCompressor that efficiently compress multimodal traffic sensor data from various transportation modes, including buses and taxis [44]. It also includes data integration and missing data recovery to assist with building energy consumption and power load forecasting [45,46], and the automatic calculation of carbon emission accounting based on structured data inputs from a building's entire life cycle [47].

Taking a step further, LLMs enable non-expert users to generate scripts using natural language in conjunction with software. For instance, LLMs can be used to automatically construct scene ontologies that support freight system optimization [48]. The ChatSUMO system integrates LLMs with the traffic simulation tool SUMO, allowing users to generate simulation scene scripts through natural language commands, reducing scenario creation time to 30 s with an accuracy of 96% [31]. For building energy consumption, studies have combined LLMs with specialized software like EnergyPlus, enabling users to complete the entire process from parameter setting to simulation script writing using natural language [49–53]. For specialized physical modeling languages like Modelica, LLM-based code generation workflows for power system

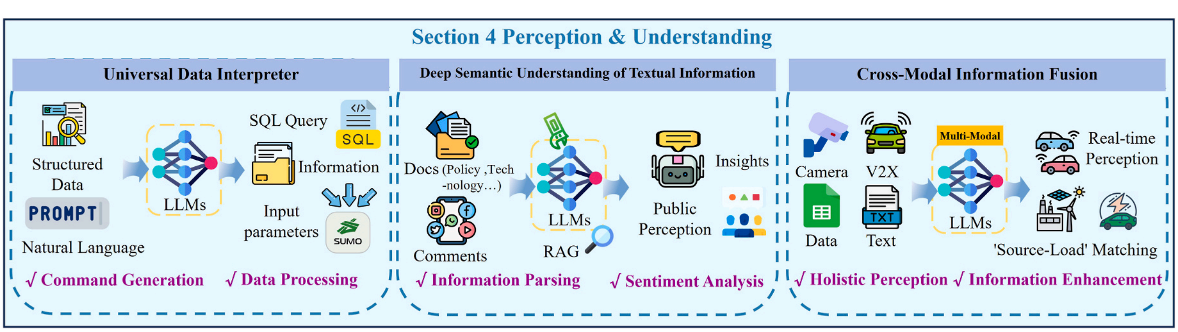

Fig. 3. Perception and understanding: Extracting knowledge from multi-source information.

tasks are also being developed [54].

4.2. Deep semantic understanding of textual information

TEI generates a massive volume of interdisciplinary text, encompassing diverse information from technical specifications and policy reports to public opinion. This presents a significant challenge for efficiently extracting key knowledge and understanding public sentiment. LLMs, with their superior contextual understanding and knowledge integration capabilities, offer a unified paradigm to address this, surpassing the limitations of traditional Natural Language Processing (NLP) methods. However, applying general-purpose LLMs to specific domains requires specialized adaptation techniques. The current mainstream approaches can be categorized into three types: Prompting, Fine-tuning, and RAG.

In the realm of extracting objective knowledge, these methods have been widely applied. For instance, prompt-based LLM frameworks have been developed to operationalize policy analysis, enabling quantitative evaluation of policy impacts on the EV industry [27]. Through fine-tuning, a team has trained the first open-source large model specialized for renewable energy using professional papers, demonstrating superior performance on zero-shot and few-shot tasks [26]. RAG is perhaps the most widely used paradigm currently, employed to accurately extract information from the latest energy storage technology literature to assist scientific research [28,55], to extract renewable energy siting regulations from legal documents with high precision [56], to build energy chatbots that provide comprehensive insights for small and medium-sized enterprises [57], and to enhance societal understanding and participation in achieving carbon neutrality [58].

At a higher level of social and behavioral cognition, LLMs are dedicated to discerning public attitudes and emotional patterns from text that influence the planning and development of TEI. A deep understanding of public attitudes is a core factor shaping policy and market development, and LLMs have demonstrated significant advantages over conventional methods in these tasks. For example, when analyzing user reviews of EV charging stations, the accuracy of GPT-4 has been proven to be higher than that of CNN models, and after incorporating human feedback, it can surpass previous benchmarks in EV market sentiment classification tasks with only a small amount of data [59]. Furthermore, when processing unstructured social media traffic complaints, the performance of LLMs in summary generation is also significantly superior to the traditional LexRank algorithm [60]. Existing work has combined LLMs with topic models like BERTopic or Aspect-Based Sentiment Analysis (ABSA) to extract features of public perception regarding charging infrastructure from social media texts [61,62]. Additionally, fine-tuning [63] and Aspect-level Sentiment Triplet Extraction (ASTE) approaches [64] have been used to synthesize expert knowledge from policy documents and social media discourse, producing structured recommendations for EV consumers. These behaviorally grounded insights are closely linked to EV adoption rates and charging patterns, thereby informing system-level planning and optimization within TEI frameworks.

4.3. Cross-modal information fusion

The complexity of TEI systems necessitates analysis that transcends single data sources; consequently, cross-modal analysis that fuses heterogeneous information from multiple sources has become a key frontier. LLMs with their powerful semantic understanding and unified representation capabilities, are emerging as the central hub for connecting different data modalities and building a comprehensive system cognition.

At the level of physical environment perception, cross-modal fusion aims to integrate various sensor data to construct rich, real-time world models, thereby directly enhancing the operational efficiency of TEI systems. For example, in autonomous driving, multimodal fusion of camera and LiDAR data has been employed to enhance 3D object detection accuracy. This capability underpins smoother and more energy-efficient vehicle control, thereby supporting energy optimization at both the individual vehicles and fleet levels [65]. AccidentGPT further fuses various perception data from Vehicle-to-Everything (V2X) sources to build a comprehensive scene-understanding model for traffic accident analysis and prevention [66]. By preventing accidents and optimizing post-accident responses, this technology can effectively reduce the significant energy waste caused by sudden congestion, thereby enhancing the energy resilience of the entire transportation network.

Beyond physical perception, cross-modal fusion also holds great potential for enhancing predictive and analytical capabilities. Its core contribution lies in the innovative information-fusion mechanism, which provides key support for "source-load" matching in TEI. To improve the accuracy of wind power generation forecasting, the CMLLM model innovatively "translates" numerical wind speed data into text. This allows it to be combined with the prior knowledge prompts of an LLM, providing a more reliable forecasting basis for matching volatile renewable energy with transportation charging demand [47]. Similarly, in power load forecasting, the MMGPT4LF model designs a multi-modal cross-attention mechanism to effectively fuse time-series data like historical load with non-sequential text such as news and holidays, which is crucial for precisely managing the increasingly complex grid loads caused by the large-scale integration of EVs [67].

5. Prediction and reasoning: Fusing Knowledge for State and Behavior Modeling

Building upon a foundation of structured knowledge, the next echelon of intelligence involves reasoning about system dynamics and predicting future states. This section examines how LLMs are surpassing traditional forecasting by integrating extracted knowledge with quantitative data to develop sophisticated models of system states, human behaviors, and potential future scenarios, as illustrated by the framework in Fig. 4.

5.1. Knowledge-augmented time-series forecasting

Time-series forecasting, a fundamental task in TEI, has long faced challenges such as high data dimensionality, strong stochasticity, and sample scarcity. The introduction of LLMs offers new solution pathways for these difficulties through two key advantages: their few-shot learning capability and their ability to deeply integrate external knowledge.

The exceptional few-shot and zero-shot learning capabilities of LLMs can effectively alleviate the model generalization problem caused by data scarcity. This advantage has been validated in both the transportation and energy domains. In the transportation sector, spatiotemporal-adapted architectures such as ST-LLM [66] and TPLLM [67] demonstrate strong few-shot performance by tailoring LLMs to capture structured traffic dynamics, highlighting their ability to generalize under sparse data conditions. In energy storage applications, prompt-based LLM frameworks have achieved robust State of Charge (SOC) estimation across multiple battery chemistries by encoding numerical battery measurements into text representations [68,69]. This paradigm also extends to renewable energy and carbon emission forecasting, where fine-tuned LLMs have demonstrated superior generalization compared with conventional models in tasks such as wind energy prediction [70] and zero-shot forecasting of provincial industrial carbon emissions [71].

Furthermore, LLMs can profoundly understand and quantify the impact of unstructured textual information on time-series evolution, enabling knowledge-augmented prediction and improving interpretability. For example, MMGPT4LF designed a multi-modal cross-attention mechanism to effectively fuse time-series data, such as historical load and weather, with textual information like news and holidays,

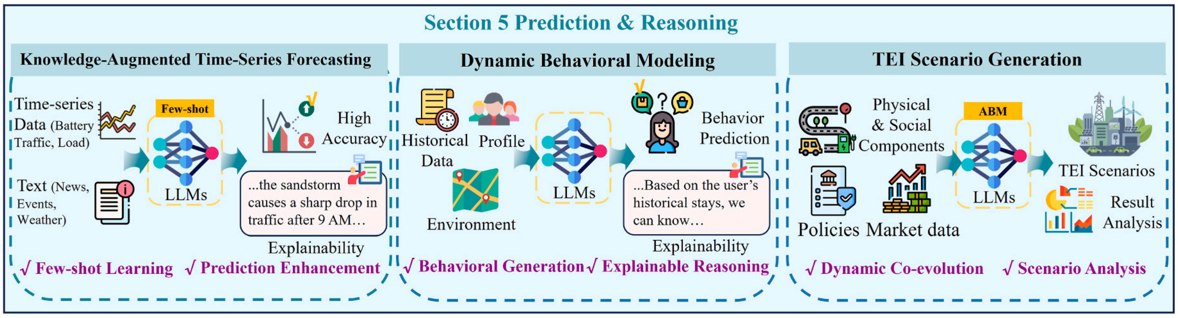

Fig. 4. Prediction and reasoning: Fusing knowledge for state and behavior modeling.

significantly improving the accuracy of load forecasting [67]. Similarly, forecasting performance has been enhanced by embedding prior knowledge—such as task instructions and statistical descriptors—into prompts [70], as well as by incorporating exogenous information, including extreme events [11,72], news sentiment [73], and weather conditions [74]. These strategies have improved the accuracy of energy supply–demand forecasting, carbon price prediction, traffic flow estimation, and EV charging demand projection. A parallel development is emerging in transportation systems, where efforts are increasingly directed toward bridging structured numerical data and unstructured textual representations. For example, Time-LLM aligns time-series signals with semantic text space through prototype learning and leverages instruction-based prompting to enhance temporal modeling performance [75]. A more efficient hybrid approach uses LLMs to process traffic-related textual context, such as weather forecasts, concerts, and accidents, to generate semantic embedding vectors as auxiliary features. These are then integrated into traditional spatiotemporal prediction models [76]. This method significantly improves prediction accuracy while also effectively circumventing the high computational costs associated with using LLMs for direct end-to-end numerical forecasting [77].

Although the application of LLMs in transportation and energy time-series forecasting is still in its early stages, the potential they have demonstrated is quite clear. In the future, deep integration of heterogeneous multi-source information, including traffic flow, grid load, macroeconomic data, and social events, is expected to enable precise forecasting of key indicators such as microgrid energy consumption and carbon emissions. This will provide more intelligent and dynamic decision support for urban planning and public policy formulation.

5.2. Dynamic Behavioral Modeling

Accurate reasoning about dynamic human behavior will be a critical link in achieving efficient, coordinated optimization of transportation and energy systems. The proliferation of new mobility services, such as EVs and ride-hailing, is triggering fundamental behavioral changes. The inherent sparsity and stochasticity of these behaviors pose significant challenges for traditional prediction models, while also exposing their interpretability deficiencies. LLMs, with their powerful contextual understanding and reasoning capabilities, provide a new paradigm for modeling and understanding this complex human mobility. At the level of travel choice, the predictive accuracy of LLMs has surpassed that of traditional models by leveraging contextual information, such as historical mobility data [78,79]. More importantly, they can provide interpretability for "black-box" predictions by clearly articulating their reasoning process, for example, inferring a future destination based on "historical visit records on weekday nights." This greatly enhances the transparency of the decision-making process [78] and allows for the effective integration of information about external, unforeseen events to improve prediction robustness [80]. At the level of driving behavior, the capability of LLMs extends from "predicting choices" to "generating actions". For instance, frameworks such as SurrealDriver have built natural language-driven driver agents whose decisions on steering and speed control exhibit features highly similar to those of humans [81]. By introducing mechanisms for driving style alignment, the agent's behavior can be better aligned with the habits of specific human drivers, significantly enhancing the credibility and realism of traffic simulations [82]. This also offers new approaches to addressing data privacy issues in transportation research [83,84].

The research mentioned above lays the foundation for connecting behavioral models with TEI systems. For example, individual mobility predictions generated by LLMs can be injected into macroscopic traffic flow models to quantify their overall impact on system energy consumption and carbon emissions. Integrating the travel and charging decisions of EV users can provide high-fidelity inputs for charging facility planning and V2G scheduling. It is even possible to couple changes in residential and work patterns with Building Energy Models (BEM) to reveal the energy consumption transfer effects between the transportation and building sectors. Applications of coupling LLMs with transportation and energy systems in this manner will be detailed in Section 5.3.

5.3. TEI scenario simulation

Building upon dynamic behavioral reasoning, TEI scenario simulation constructs complete virtual testing environments by integrating specific physical and social components. Traditional simulation methods often require the manual design of extensive rules and are primarily based on probability. They generally lack human-like cognitive abilities and natural language interaction mechanisms, making it difficult to accurately reflect complex social factors such as user preferences, policy feedback, and market evolution. In contrast, generative agents driven by LLMs demonstrate potential that surpasses traditional rule-based or learning models in such simulations, owing to their ability to dynamically capture heterogeneous agent behavior and market changes [85].

As research progresses, LLM-based agents have achieved deep integration with dynamic traffic simulators, providing decision support for system planning by accurately predicting commuting demands [86]. In the critical context of EV charging, generative modeling frameworks have been developed that integrate individual psychological attributes and environmental factors, enabling fine-grained representation and prediction of charging behavior [87]. Furthermore, to capture more complex system interactions, multi-agent game-theoretic structures have been introduced. For instance, the dual-agent model proposed by Niu et al. simulates the competition and cooperation between EV users and charging stations, achieving personalized recommendations and adaptive price adjustments [88].

To address the adaptive and nonlinear dynamics of transportation–energy systems, LLM-based surrogate modeling approaches have been introduced to analyze game-theoretic mechanisms linking regulatory policies and market behavior. Related studies demonstrate that such models can not only evaluate the impact of different policy orientations on EV market entry [14] but also accurately simulate the participation patterns of V2G services under real-time incentive mechanisms [32]. The advantage of this simulation strategy lies in its flexibility; by jointly inferring micro-level agent behaviors and macro-level economic variables (e.g., policy, pricing), it holds the potential to construct system co-evolution models that more closely align with real-world demands in the future [89].

6. Decision-making and optimization: language-based Intelligent Scheduling and Control

The ultimate purpose of an intelligent TEI system is to translate its understanding and predictions into optimal, real-world actions. This chapter examines the highest level of LLM application, focusing on how these models are employed to execute complex decision-making and optimization tasks, often by leveraging natural language as a powerful interface for defining objectives and controlling dynamic systems, as illustrated by the framework in Fig. 5.

6.1. Decision support and optimization formulation

In the field of TEI, LLMs are becoming a key bridge connecting human decision-makers with complex optimization tools. Their application is mainly reflected at two levels: first, as intuitive decision-support assistants, and second, as tools for formulating optimization problems.

A primary application direction is the development of domain-specific intelligent assistants that transform complex system operations into simple natural language dialogues. For system operators, systems like TP-GPT [90] for traffic monitoring, GAIA [91] for power dispatching, and intelligent assistants [92] for power simulation management, allow managers to query real-time data and execute complex commands via dialogue, significantly improving operational intuitiveness and efficiency. For high-level planning and end-users, LLMs can provide strategic advice and personalized recommendations. For example, they can generate in-depth strategic recommendations for transportation planning [93], explain and adjust building carbon emission strategies [94,95], and recommend the most suitable EV charging stations to users based on real-time data [96].

The ability of an LLM to understand human intent naturally extends to the highly structured task of formulating mathematical optimization problems. The core challenge in this area is to translate users' vague and personalized needs, such as "charge my car tonight when the electricity price is lowest", into precise constraints and objectives that an optimization solver can understand. The research path has likely evolved from early direct code generation to more reliable structured methods. For example, LLM4EV [12] utilizes function-calling templates to ensure the validity and accuracy of the output. The role of LLMs is not limited to "translating" user intent; research shows that in the field of carbon reduction, techniques like Retrieval-Augmented Generation (RAG) can be used to leverage external knowledge to construct more precise optimization objectives [97]. Furthermore, combining an LLM with traditional optimization algorithms, such as the whale optimization algorithm, has been proven to effectively enhance global solving performance in complex electricity-transportation networks, with performance improvements ranging from 1% to 20% compared to traditional methods [13]. These studies indicate that an emerging collaborative paradigm is taking shape: LLMs handle problem modeling, preference modeling, or heuristic guidance, while traditional optimization algorithms continue to undertake numerical solutions and constraint satisfaction tasks. While these hybrid approaches enhance numerical efficiency, collaborative optimization still faces fundamental challenges in semantic modeling, particularly within broader V2X scenarios such as vehicle-to-building coordination. Existing methods primarily rely on structured numerical inputs (e.g., battery state of charge, real-time electricity prices), often struggling to integrate qualitative contextual information like temporary facility scheduling or urgent user demands. This limitation reveals a key development direction for next-generation collaborative frameworks: large language models will serve as advanced semantic coordinators, extracting contextual knowledge from unstructured data and converting it into constraints or preference signals for processing by traditional optimization or learning algorithms. This layered architecture promises to significantly enhance coordination efficiency among heterogeneous subsystems while maintaining the reliability of existing numerical solvers.

From an economic perspective, the use of LLMs in TEI cost optimization is also emerging. Early studies mainly position LLMs as modeling assistants to help formulate cost-oriented optimization problems, where metrics such as total cost of ownership [98] or operational cost deviations from benchmark solvers [99] are used for evaluation. In parallel, research on V2G operation has established economic assessment models that account for battery degradation and bidirectional energy revenues [100]. These cost formulations provide useful references for future LLM-enabled approaches, highlighting the need to integrate long-term asset impacts into optimization objectives rather than focusing solely on short-term energy prices.

6.2. Real-time operational control and scheduling

In the real-time operation and scheduling of TEI systems, LLMs are transitioning from decision support tools to direct intelligent control agents. This capability is demonstrated through traffic flow optimization in macroscopic transportation networks. High-level applications like the TrafficGPT system enable managers to formulate traffic control strategies through natural language to alleviate congestion and optimize flow [101]. At the microscopic level, LLMs provide precise advice to drivers based on vehicle dynamics, excelling in conflict identification and decision-making [102]. Particularly in the field of Traffic Signal Control (TSC), LLMs have evolved from assisting in environmental understanding to directly acting as control agents. Research indicates that in multi-intersection scenarios, they can significantly reduce vehicle waiting times and the number of stops, thereby substantially improving

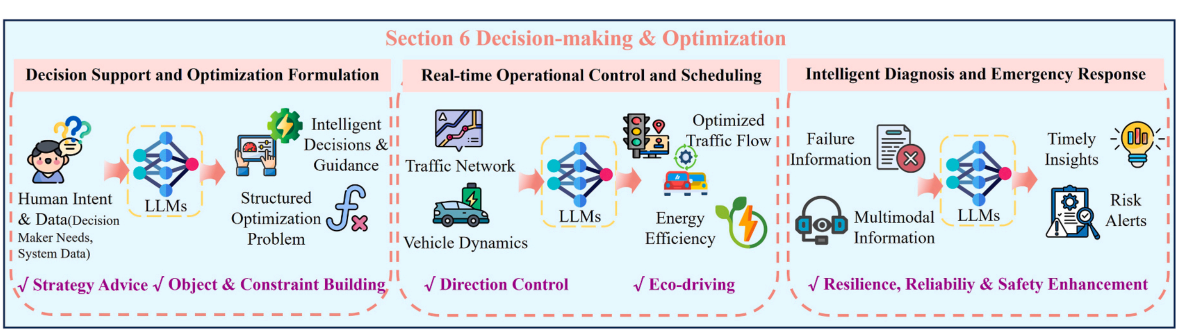

Fig. 5. Decision-making and optimization: Language-feedback-based intelligent scheduling and control.

network efficiency [103,104].

The strategic guidance capabilities of LLMs at the micro-operational level show potential for reducing energy consumption and carbon emissions. In the context of eco-driving assistance, for instance, LLM-based systems have provided real-time adaptive speed recommendations, achieving up to an 82% reduction in congestion and a 75% decrease in carbon emissions under experimental conditions [105]. In addition, LLMs have been explored as intelligent co-drivers that analyze environmental parameters to support dynamic control decisions, resulting in reported energy savings of up to 54% [106]. Achieving deeper efficiency improvements in heterogeneous powertrain systems, such as plug-in hybrid electric vehicles (PHEVs) and fuel cell electric vehicles (FCEVs), requires coordination under complex physical constraints [107–110]. Related research integrates vehicle operational states (e.g., state of charge and hydrogen reserves) into LLM-generated high-level strategies that guide the training of reinforcement learning agents, thereby improving overall system efficiency [111].Collectively, these studies indicate that LLMs can help translate system-level emission reduction objectives into context-aware operational strategies, offering a promising technical pathway to support deep decarbonization in the transportation sector.

However, it is necessary to acknowledge the inherent latency bottleneck of current LLMs in high-frequency control. Unlike traditional optimization algorithms that typically operate on millisecond timescales, the inference latency of LLMs generally ranges from hundreds of milliseconds to seconds [112]. This gap currently restricts their direct deployment in millisecond-level safety-critical closed loops (such as emergency braking or grid frequency regulation). Therefore, future research requires a further focus on model lightweighting and edge acceleration technologies.

6.3. Intelligent diagnosis and emergency response

LLMs are also poised to become a key tool for enhancing the emergency response capabilities and resilience of TEI systems. By virtue of their ability to deeply integrate and understand massive, unstructured fault information, LLMs can achieve in-depth diagnosis and precise identification of failures and anomalies. For example, LEMAD, a system based on multi-agent systems, utilizes an LLM to parse system logs for root cause analysis of power system anomalies, achieving an accuracy of 92.16% [113]. Meanwhile, DeCEN ensures the reliability of decentralized clean energy networks by analyzing sensor data to verify the authenticity of energy declarations and prevent fraud [114]. Building on this, the application of LLMs is evolving to “proactive warning”. By integrating and analyzing global news or driver multi-modal information, systems like SHIELD and PDA can respectively predict supply chain disruption risks and driving safety risks, thereby prospectively enhancing the system's ability to cope with disturbances [115,116]. A higher level of intelligence is demonstrated in the active generation and autonomous correction of safety strategies. The LLM within the Grid-Agent framework can autonomously generate action sequences to correct grid violations [117], while the LLM in the Rl2 framework, applied to active distribution networks, can even generate and iteratively optimize rules for safety-oriented reinforcement learning [118]. This progression of capabilities, from precise diagnosis and proactive warning to autonomous correction, could significantly enhance the resilience of TEI systems when facing failures, attacks, and uncertainties.

7. Challenges and limitations

Translating LLMs from theory into reliable TEI systems requires confronting four interconnected core challenges. This endeavor begins with the foundational problem of interoperability in integrating heterogeneous data and demands that user trust be earned when sensitive information is leveraged. Building on this, the models must resolve their inherent reliability dilemma for high-stakes decision-making while overcoming the sustainability paradox, in which their computational cost conflicts with TEI's low-carbon goals.

7.1. The interoperability challenge: integrating heterogeneous data across TEI silos

As a quintessential “System of Systems”, TEI is characterized by cross-domain and cross-scale properties that render the challenge of data interoperability significantly more acute than in any single domain. In the transportation and energy domains, achieving multidimensional data integration faces several thorny challenges. The data generated by transportation systems is highly complex, dynamic, and incomplete, and is influenced by factors such as weather, road infrastructure, driver behavior, and external events. Additionally, the data from energy systems also suffers from quality issues due to equipment failures and environmental changes. The data sources required for training LLM-based models are diverse, encompassing not only traffic sensors, cameras, and other monitoring devices within the energy system, but also social media data, academic papers, and research reports. These data from different domains exhibit significant differences in format, structure, and semantics, making it complex and resource-intensive for LLMs to integrate and preprocess them into a coherent and representative training corpus. Meanwhile, transportation and energy patterns change over time, requiring continuous data updates to maintain model accuracy and relevance. For example, in the context of smart transportation and smart grid collaboration, real-time changes in traffic flow affect EV charging demand, which, in turn, affects the grid load. However, there are differences between traffic flow data and grid load data in terms of collection frequency, accuracy, and data format. Traffic flow may be updated in minutes or even seconds, while grid load data is typically counted hourly, making data integration extremely challenging. If these data challenges are not properly addressed, LLMs are difficult to generalize effectively, potentially leading to bias in the formulation of TEI strategies. These challenges, however, can be mitigated through advanced alignment strategies [119]. For instance, Feature-based Alignment maps heterogeneous time series into a shared latent space to achieve synchronization [120], while Attention-based Alignment leverages cross-modal attention mechanisms to capture correlations across differing temporal granularities [121,122]. Furthermore, the advent of LLMs has introduced the ‘Text-Guided Analysis’ paradigm, offering a novel pathway to bridge the semantic gap between numerical data domains via textual modalities [123]. At the physical level, it is also imperative to advance the integration of multiple collection channels across the transportation and energy domains, building an integrated data collection network, and deploying intelligent collection devices in transportation and energy facilities. In addition, generative artificial intelligence techniques, such as LLMs and generative adversarial networks, are utilized to generate realistic data to compensate for missing values in real data [124].

In addition, the multifaceted and interdisciplinary nature of the transportation and energy domains significantly inflates the cost and complexity of constructing high-quality datasets. Currently, existing ontological resources remain siloed [125,126], leading to a lack of a comprehensive corpus capable of capturing the dynamic coupling inherent to TEI. This absence hinders the fine-tuning of LLMs for tasks such as energy consumption prediction and carbon emission assessment [21]. Therefore, there is an urgent imperative to recommend a standardized and reusable semantic annotation framework specifically for TEI. This paper proposes that such a system should be structured into three core layers: (1) a “Spatio-Temporal Alignment Layer” that employs unified anchors (e.g., GeoHash [127] and synchronized timestamps) to resolve granularity mismatches between traffic flows and grid loads; (2) an “Event-Driven Interaction Layer” that encapsulates heterogeneous data into semantic “coupling events” (e.g., ‘congestion-induced charging delays’) to explicate cross-domain causal logic; and (3) a “Goal-Oriented Reasoning Layer” that integrates policy constraints and multidimensional performance metrics (such as carbon intensity and grid stability metrics). Establishing this hierarchically structured corpus is essential for creating reproducible benchmarks and fostering open-source research [128].

7.2. The trust imperative: navigating privacy risks and security threats in TEI

The personalized services in TEI, such as intelligent charging, are deeply dependent on users' spatio-temporal trajectories and energy consumption data; the highly sensitive nature of this information creates an exceptionally sharp conflict between privacy and utility. Both the transportation and energy domains involve frequent user interactions, and data collection in these domains encompasses highly sensitive personal information, including travel and commuting patterns, energy usage data, and driving and charging behaviors. Inadequate system security poses a significant risk of data leakage. It has been demonstrated that attackers can gain access to system model code, raw data, and personal information through cyberattacks [129–131]. Such risks may result in financial loss or harm to users, which in turn may erode their trust in the system. Developers need to ensure strict fairness and secure operation in application development [132,133] and implement advanced data protection strategies such as differential privacy, synthetic data generation, and encrypted LLM systems [134].

7.3. The reliability dilemma: mitigating hallucinations and ensuring interpretability in critical TEI systems

Unlike in general applications, TEI is directly concerned with the safety and stability of critical infrastructure; consequently, the "black box" nature and "hallucination" issues of LLMs escalate from mere technical flaws to unacceptable security risks. As a critical infrastructure domain, transportation and energy systems require AI models to have decision transparency, operational reliability, and result credibility [135]. Interpretability refers to the ability of a model to present its behavioral logic in a human-understandable way [136], which is particularly important for understanding the functional boundaries and operational mechanisms of multimodal AI systems in transportation energy management [137]. For example, LLMs suffer from causal inference bias and over-rely on the assumption of temporal smoothness in driving behavior data, making it difficult to predict unexpected situations in dynamic traffic environments accurately [16,138]. Meanwhile, the implicit bias in the training data may be amplified by LLMs, which in turn affects the fairness of traffic scheduling decisions [10,139,140]. Methods for assessing interpretability include having LLMs perform problem-specific evaluations, evaluating automatic metrics such as their ability to simulate functions, and improving the accuracy of downstream tasks. This involves comparing arguments and post-hoc explanations of LLMs and utilizing sample explanations in the reasoning process [141]. Two main approaches are used to enhance model interpretability: one involves establishing a complete traceability mechanism for model architecture, training data, and version iteration, while the other consists of enhancing internal decision transparency through techniques such as thought chain simulation and knowledge graph visualization [142]. Xie et al. constructed a multimodal agent analysis framework that systematically elaborates on interpretability methods in visual-verbal interaction scenarios and provides a multimodal system with a new decision tracing scheme [143]. In the field of autonomous driving, the Cog-GA framework enhances the end-to-end interpretability of the navigation decision-making process through cognitive map construction and two-channel scene parsing [144].

The problem of model illusion seriously threatens the reliability of transportation energy systems, which is manifested in two types: intrinsic illusion, referring to the direct contradiction between the output content and the input context, and extrinsic illusion, which lacks a verifiable external basis [145]. In real-time traffic control scenarios, such deficiencies can lead to major decision-making errors; in energy extrapolation systems, they can result in significant deviations between simulation results and reality. Mitigation strategies for the illusion problem have been developed as multidimensional solutions: at the data level, knowledge gaps are filled by selecting training data sources and strict quality auditing combined with RAG techniques [24,146]; at the training stage, attention-sharpening regularization methods are used to enhance architectural sparsity and suppress the generation of illusions [147]; and at the inference stage, the model is strengthened through context-aware decoding techniques, response weights to immediate information [148]. These methods provide crucial technical support for constructing a reliable, large-model application system.

7.4. The sustainability paradox: balancing computational demands with the decarbonization goals of TEI

The core mission of TEI is to improve energy efficiency and support decarbonization. However, the growing adoption of LLMs introduces a sustainability paradox: the substantial computational resources required by LLMs may partially offset the energy savings achieved in physical systems.

First, LLMs consume significant computational resources during both training and inference. As illustrated in Fig. 6, which compares training resource requirements across models of different parameter scales, the computational power and memory demands of pre-training large models are several orders of magnitude higher than those of traditional deep learning models [149]. Moreover, in time-sensitive TEI applications such as traffic control and energy dispatch, the inference stage can also become resource-intensive [150,151]. Second, most existing studies evaluate the effectiveness of LLMs primarily in terms of operational energy savings achieved by TEI systems, such as congestion mitigation [105] or reductions in energy consumption [106], while rarely accounting for the computational energy cost of the AI models themselves. Consequently, the overall system-level energy balance has not yet been comprehensively quantified.

A key practical barrier lies in the widespread reliance on commercial, closed-source LLM services and the absence of standardized evaluation metrics. Limited transparency regarding model size, activation mechanisms, and data center efficiency makes it difficult to accurately estimate the energy consumption of each inference. Future evaluations of TEI systems should therefore consider not only control or optimization performance, but also the energy overhead associated with artificial intelligence components [155]. Establishing unified metrics for LLM energy consumption and carbon emissions is crucial for enabling such holistic assessments [152].

Several emerging technical trends may help alleviate this paradox and promote more sustainable LLM deployment. Model compression [156], parameter-efficient fine-tuning [157,158], and improved architectures are reducing computational demands [159,160]. At the system level, deploying models on specialized low-power hardware, transitioning toward smaller or edge-deployable language models, and increasing the use of renewable-powered data centers can further lower the carbon intensity of AI inference [161,162].

8. Future research directions

Based on the review presented in this study, this chapter outlines several future research directions for LLMs in TEI. These approaches can be beneficial in facilitating and enhancing the application of artificial intelligence in sympathetic energy fusion, as shown in Fig. 7.

8.1. From text to physical reality: multimodal fusion for TEI situational awareness

Current LLMs are predominantly text-centric, whereas TEI systems are inherently multimodal physical environments. In real-world

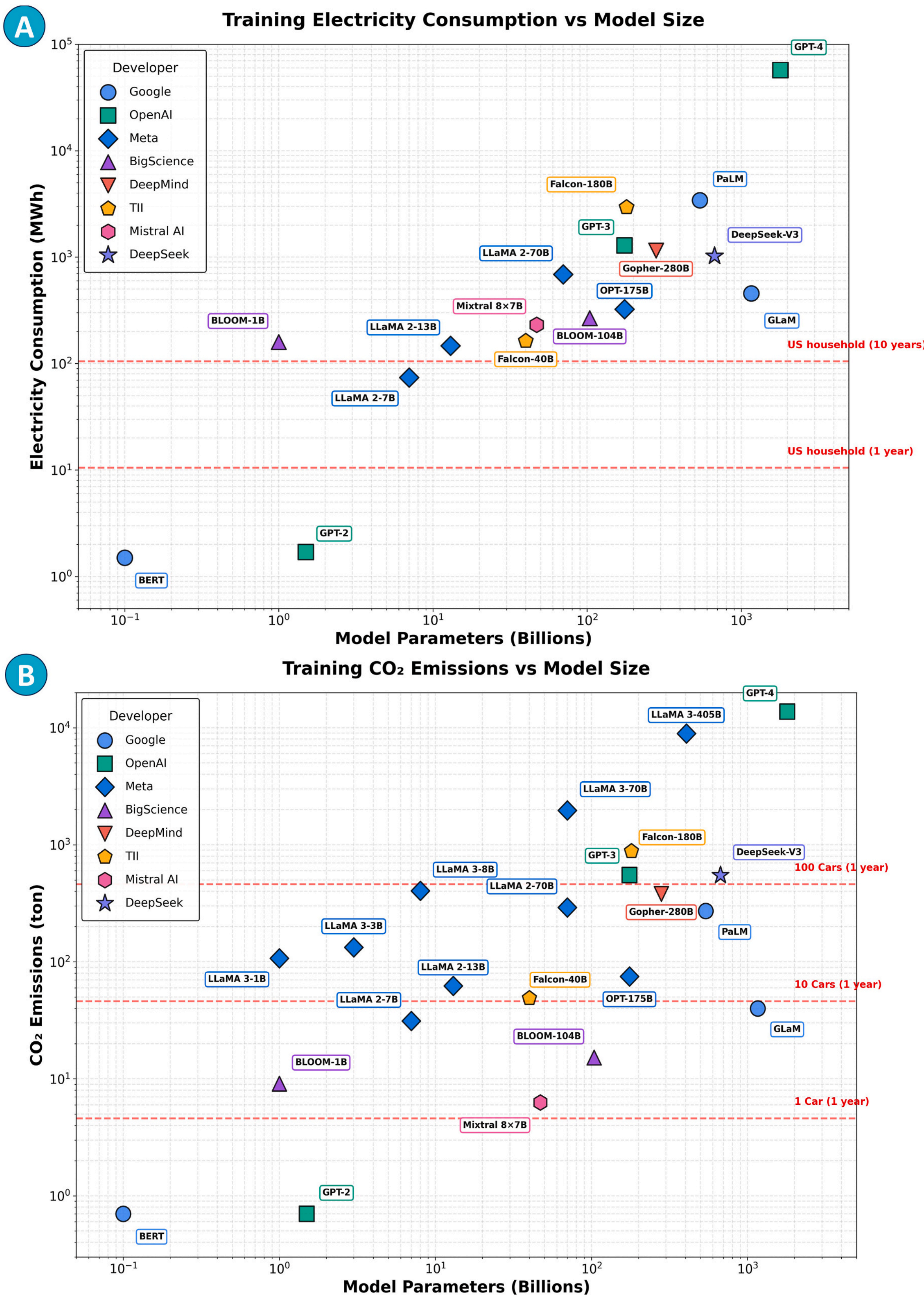

Fig. 6. Comparison of training resource requirements for LLMs, data sourced from Ref. [152]. (a) Comparison of Training Power Consumption and Model Parameters. A logarithmic scale plots the relationship between power consumption (megawatt-hours) and model scale (billions of parameters). The red dashed lines provide reference benchmarks: one indicates the average annual power consumption of a U.S. household (10.5 MWh/year [153]), and the other represents the cumulative power consumption of that household over ten years. (b) Training CO_2 emissions versus model parameters. Plot model size against total CO_2 emissions (metric tons) on a logarithmic axis. The red dashed lines represent the equivalent annual emissions of 1, 10, and 100 typical passenger vehicles (each emitting approximately 4.6 tons of CO_2 annually [154]).

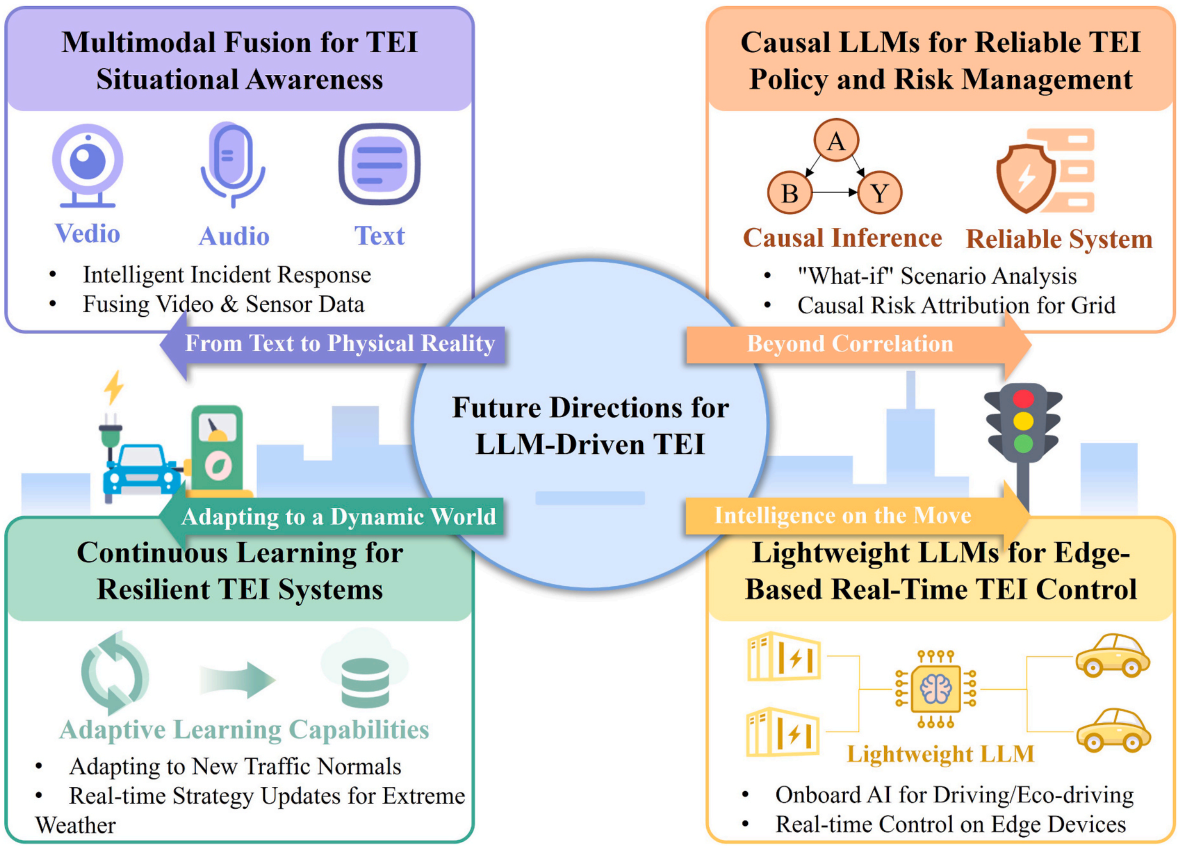

Fig. 7. Future directions for LLM-Driven transportation energy integration.

operation, dispatch and control decisions rely on heterogeneous data streams such as video surveillance, sensor measurements, equipment alarms, and spoken commands, rather than textual reports alone. Therefore, the further evolution of LLMs in TEI depends on their ability to integrate deeply with vision, time-series sensing, and speech models to form a unified understanding of physical operating conditions. The goal of multimodal fusion is not merely to aggregate diverse inputs, but to establish cross-modal semantic alignment and causal reasoning [163]. For example, in a transportation–energy coordination scenario, a system may need to jointly interpret a detected road incident from video feeds, a measurable drop in traffic throughput, and contextual information such as an upcoming large public event, then translate this combined understanding into actionable control strategies. Such capabilities would allow LLMs to move beyond the role of "text interpreters" and toward becoming central hubs for multimodal situational awareness, providing a more complete environmental context for complex TEI decision-making. In summary, the development of multimodal LLMs will be crucial in determining whether these models can meaningfully participate in high-level TEI operations. Only when language models can reliably interface with real-world data streams can they support fast, context-aware perception and coordinated control in practical TEI environments.

8.2. Beyond correlation: causal LLMs for reliable TEI policy and risk management

While current LLMs excel at identifying statistical regularities—for instance, that high temperatures are typically concomitant with peak EV charging demand—they often struggle to distinguish whether these patterns stem from genuine causal links or merely spurious associations within the data. In complex TEI systems, distinct variables such as traffic flow and grid load, though lacking direct physical connection, may exhibit highly synchronous fluctuations driven by shared environmental confounders like weather or holidays. Mistaking these correlations for direct causation leads to prediction failure once environmental conditions shift. Moreover, relying solely on observational data prone to contextual mismatch: the model may capture only superficial features like "localized load spikes" or "traffic stoppages" while missing the critical causal context triggering these phenomena (e.g., "temporary price subsidies" or "defensive control due to accidents"). Neglecting such causal incentives behind behaviors leads to fundamental biases in risk assessment [164]. To address these limitations, future research should introduce Causal Intervention Encoders [165] into LLMs to disentangle the interference of environmental noise. Furthermore, utilizing prompting techniques such as Chain-of-Actions can decompose complex policy interventions—where a single instruction triggers cascading effects—into verifiable causal steps [165]. This causal modeling capability enables operators to accurately evaluate: "Did the new time-of-use pricing policy genuinely alter user charging habits, or does it merely coincide with weather-driven fluctuations?"

8.3. Adapting to a dynamic world: continuous learning for resilient TEI systems

TEI systems operate in environments that evolve. Changes in urban structures reshape mobility demand, the expansion of charging infrastructure alters load distribution, and emerging storage and propulsion technologies introduce new operational characteristics. Extreme weather and unexpected disruptions can further invalidate established scheduling patterns. Under such conditions, LLMs trained once in an offline setting may gradually lose relevance, showing reduced capability in interpreting new operational contexts, such as traffic redistribution caused by newly opened highways, shifts in regional charging peaks, or altered EV energy consumption patterns under abnormal weather.

Research attention is therefore moving toward model frameworks with continuous learning capabilities, enabling LLMs to update their knowledge structures and decision bases after deployment using newly observed data. Technical directions include self-supervised learning, incremental learning, and human-in-the-loop feedback, which allow progressive adaptation without full retraining [166]. These mechanisms support the incorporation of new traffic patterns, load dynamics, and equipment behaviors into existing representations, followed by corresponding adjustments in scheduling logic and resource allocation strategies. Continuous learning capability is closely related to the long-term effectiveness of TEI intelligence. Models equipped with ongoing adaptation mechanisms are more likely to maintain stable decision performance in complex and uncertain environments, supporting sustained coordination between transportation and energy systems.

8.4. Intelligence on the move: lightweight LLMs for edge-based real-time TEI control

Real-time TEI control tasks, such as autonomous trajectory planning and microgrid frequency regulation, require low-latency, locally executable intelligence [167]. Cloud-only deployment introduces communication delays and bandwidth dependence, whereas standalone edge devices constrain model scale. Consequently, research is shifting toward hierarchical device–edge–cloud architectures to balance latency and computational capacity [168]. Lightweight LLMs enabled by quantization and knowledge distillation reduce storage and computation overhead, allowing deployment on resource-constrained hardware. Quantized 7 B models can operate within approximately 4 GB automotive-grade memory [169], and mobile multimodal inference can reached basic real-time performance [170]. These advances support onboard perception and preliminary decision-making, reducing reliance on continuous cloud connectivity [171]. To further overcome device-level constraints, collaborative inference distributes workloads across nearby edge nodes, such as charging stations or building-side servers [114], expanding computational capacity while minimizing cross-domain data transmission. The integration of lightweight edge models and distributed collaboration provides a scalable pathway for low-latency LLM-driven TEI control systems.

9. Conclusion

Leveraging their remarkable emergent abilities, superior generalization performance, and multi-task learning advantages, LLMs offer an unprecedented opportunity to build resilient and efficient TEI systems. Although applications of LLMs have proliferated across subdomains of TEI, a holistic review assessing how they address the core interconnections between these systems has been absent. To fill this gap, this study systematically analyzes 82 seminal papers published since the field's inflection point in 2023. The core significance of this work lies in providing the first structured cognitive map and development roadmap for this nascent and disparate field, achieved through a holistic TEI framework and a novel three-tiered functional hierarchy of Perception, Reasoning, and Decision-making.

The three-layer analytical framework proposed in this review redefines how LLMs perceive, model, and manage complex dynamics within TEI systems. (1) At the perception layer, LLMs dismantle the information barriers of multi-source heterogeneous data; their capability to automatically parse data and extract key information promises to significantly reduce manual costs and technical barriers. (2) At the reasoning layer, LLMs make it possible to simulate complex system dynamics and human behavior. Existing research generally indicates their superiority over traditional machine learning methods in time-series forecasting and behavioral reasoning tasks, particularly regarding few-shot learning and interpretability. They have also demonstrated the potential to move beyond conventional modeling frameworks to simulate the evolution of TEI scenarios and achieve realistic results. (3) At the decision layer, LLMs act as a bridge, translating high-level human intent into optimal control strategies. They can leverage their knowledge-driven and tool-calling capabilities to provide feasible solutions and language feedback mechanisms for complex scheduling problems.

The systematic classification in this review also reveals the expansive scope of LLM applications in TEI: the research landscape comprehensively spans from foundational, transportation-centric perception tasks to complex, system-level decision-making and optimization problems tightly integrated with domains such as power grids and building energy. Meanwhile, the path to fully unleashing the LLM potential still faces challenges, including the complexity of data integration, the reliability and safety of the models, and high computational costs.

To address these challenges and guide future development, this study crystallizes four key strategic directions. These are not merely incremental technical improvements but are core capabilities essential for building the next generation of intelligent TEI systems: (1) Building models for multimodal perception to achieve a panoramic understanding of complex physical transportation-energy systems; (2) Establishing causal inference models to ensure the reliability of decision-making in high-risk infrastructure control; (3) Developing continual learning frameworks to guarantee the long-term effectiveness of models in dynamically evolving markets and environments; and (4) Researching and developing lightweight intelligence to empower distributed, low-latency edge devices with autonomous control capabilities. As technology continues to evolve and methodologies mature, LLMs are poised to become a major driving force in the intelligent, green transformation of transportation-energy systems, providing solid support for achieving strategic carbon reduction targets and developing future smart cities.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This research is supported by the Introduced Innovative R&D Team of Guangdong (2023ZT10L145). All opinions expressed in this paper are the authors' and do not necessarily reflect the policies and views of sponsors.

Data availability

Data will be made available on request.

References

[1] Jia L, Shi R, Ji L, et al. Road transportation and energy integration strategy in China. Chin J Eng Sci 2022;24:163. https://doi.org/10.15302/J-SSCAE-2022.03.017.
[2] IEA – international energy agency. IEA n.d. https://www.iea.org (accessed May 7, 2025).
[3] Zhou Y, Cao S, Hensen JLM, et al. Energy integration and interaction between buildings and vehicles: a state-of-the-art review. Renew Sustain Energy Rev 2019;114:109337. https://doi.org/10.1016/j.rser.2019.109337.
[4] Muratori M, Jadun P, Bush B, et al. Future integrated mobility-energy systems: a modeling perspective. Renew Sustain Energy Rev 2020;119:109541. https://doi.org/10.1016/j.rser.2019.109541.
[5] Bogdanov D, Ram M, Khalili S, et al. Effects of direct and indirect electrification on transport energy demand during the energy transition. Energy Policy 2024;192:114205. https://doi.org/10.1016/j.enpol.2024.114205.
[6] Liu X, Plötz P, Yeh S, et al. Transforming public transport depots into profitable energy hubs. Nat Energy 2024;9:1206–19. https://doi.org/10.1038/s41560-024-01580-0.
[7] Li D, Zouma A, Liao J-T, et al. An energy management strategy with renewable energy and energy storage system for a large electric vehicle charging station. eTransportation 2020;6:100076. https://doi.org/10.1016/j.etran.2020.100076.
[8] Tan Z, Yang P, Nehorai A. An optimal and distributed demand response strategy with electric vehicles in the smart grid. IEEE Trans Smart Grid 2014;5:861–9. https://doi.org/10.1109/TSG.2013.2291330.

[9] Soares J, Almeida J, Gomes L, et al. Electric vehicles local flexibility strategies for congestion relief on distribution networks. Energy Rep 2022;8:62–9. https://doi.org/10.1016/j.egyr.2022.01.036.
[10] Mahmud D, Hajmohamed H, Almentheri S, et al. Integrating LLMs with ITS: recent advances, potentials, challenges, and future directions. IEEE Trans Intell Transport Syst 2025:1–36. https://doi.org/10.1109/TITS.2025.3528116.
[11] Guo X, Zhang Q, Jiang J, et al. Towards explainable traffic flow prediction with large language models. Commun Transp Res 2024;4:100150. https://doi.org/10.1016/j.commtr.2024.100150.
[12] Huang C, Li S, Liu R, et al. Large foundation models for power systems. 2024 IEEE Power Energy Soc Gen Meet PESGM 2024:1–5. https://doi.org/10.1109/PESGM51994.2024.10688670.
[13] Chen X, Lu X, Li Q, et al. Integration of LLM and human–AI coordination for power dispatching with connected electric vehicles under SAGVNs. IEEE Trans Veh Technol 2025;74:1992–2002. https://doi.org/10.1109/TVT.2024.3434969.
[14] Ansaldo G. AgentSpeak: a framework for agent-based modeling with integrated large language models; case study: analyzing policy interventions in electric vehicle adoption. Northeastern University; 2023. https://doi.org/10.17760/D20621578.
[15] Zhang X, Zhang Z, Liu Y, et al. A review of machine learning approaches for electric vehicle energy consumption modelling in urban transportation. Renew Energy 2024;234:121243. https://doi.org/10.1016/j.renene.2024.121243.
[16] Gan L, Chu W, Li G, et al. Large models for intelligent transportation systems and autonomous vehicles: a survey. Adv Eng Inform 2024;62:102786. https://doi.org/10.1016/j.aei.2024.102786.
[17] Abraham A, Aldhanhani T, Hamidouche W, et al. Harnessing the power of large language models for sustainable and intelligent transportation systems in the electric vehicle era. In: Abraham A, Prasad S, Alhammadi A, et al., editors. Internet veh. Comput. Vis. Solut. Smart City transform. Cham: Springer Nature Switzerland; 2025. p. 85–113. https://doi.org/10.1007/978-3-031-72959-1_5.
[18] Majumder S, Dong L, Doudi F, et al. Exploring the capabilities and limitations of large language models in the electric energy sector. Joule 2024;8:1544–9. https://doi.org/10.1016/j.joule.2024.05.009.
[19] Madani S, Tavasoli A, Astaneh ZK, et al. Large language models integration in smart grids. https://doi.org/10.48550/arXiv.2504.09059; 2025.
[20] Liu M, Zhang L, Chen J, et al. Large language models for building energy applications: opportunities and challenges. Build Simulat 2025;18:225–34. https://doi.org/10.1007/s12273-025-1235-9.
[21] Zhang L, Chen Z. Opportunities of applying large language models in building energy sector. Renew Sustain Energy Rev 2025;214:115558. https://doi.org/10.1016/j.rser.2025.115558.
[22] Zhang X, Glaws A, Cortiella A, et al. Deep generative models in energy system applications: review, challenges, and future directions. Appl Energy 2025;380:125059. https://doi.org/10.1016/j.apenergy.2024.125059.
[23] Brown T, Mann B, Ryder N, et al. Language models are few-shot learners. Adv. Neural inf. Process. Syst., vol. 33. Curran Associates, Inc.; 2020. p. 1877–901.
[24] Touvron H, Martin L, Stone K, et al. Llama 2: open foundation and fine-tuned chat models. https://doi.org/10.48550/arXiv.2307.09288; 2023.
[25] Chang Y, Wang X, Wang J, et al. A survey on evaluation of large language models. ACM Trans Intell Syst Technol 2024;15(39):1–39. https://doi.org/10.1145/3641289. 45.
[26] Gabber HA, Hemied OS. Domain-specific large language model for renewable energy and hydrogen deployment strategies. Energies 2024;17:6063. https://doi.org/10.3390/en17236063.
[27] Gu J, Wu H, Liu M, et al. A framework for automatic interpretation of electric vehicle industry policies based on LLMs. 2024 6th Int Conf Mach Learn Big Data Bus Intell MLBDBI 2024:214–8. https://doi.org/10.1109/MLBDBI63974.2024.10823863.
[28] Leng Y, Zhong Y, Gu Z, et al. Intelligent, personalized scientific assistant via large language models for solid-state battery research. ACS Mater Lett 2025;7:1807–16. https://doi.org/10.1021/acsmaterialslett.4c02674.
[29] Eclipse SUMO - simulation of urban MObility. Eclipse SUMO - Simul urban Mobil n.d. https://eclipse.dev/sumo/. [Accessed 28 November 2025].
[30] EnergyPlus n.d. https://energyplus.net/. [Accessed 28 November 2025].
[31] Li S, Azfar T, Ke R. ChatSUMO: large language model for automating traffic scenario generation in simulation of urban MObility. IEEE Trans Intell Veh 2024:1–12. https://doi.org/10.1109/TIV.2024.3508471.
[32] Sun Y, Cui C, Zhang C, et al. Dynamic incentive strategies for smart EV charging stations: an LLM-Driven user digital twin approach. https://doi.org/10.48550/arXiv.2504.01423; 2025.
[33] Wu S, Irsoy O, Lu S, et al. BloombergGPT: a large language model for finance. https://doi.org/10.48550/arXiv.2303.17564; 2023.
[34] Li J, Wang S, Zhang M, et al. Agent hospital: a simulacrum of hospital with evolvable medical agents. https://doi.org/10.48550/arXiv.2405.02957; 2024.
[35] Kovač G, Portelas R, Dominey PF, et al. The SocialAI school: insights from developmental psychology towards artificial socio-cultural agents. https://doi.org/10.48550/arXiv.2307.07871; 2023.
[36] Izzidien A, Sargeant H, Steffek F. LLM vs. lawyers: identifying a subset of summary judgments in a large UK case law dataset. https://doi.org/10.48550/arXiv.2403.04791; 2024.
[37] Cheong I, Xia K, Feng KJK, et al. (a)I am not a lawyer, but…: engaging legal experts towards responsible LLM policies for legal advice. Proc. 2024 ACM Conf. Fairness Account. Transpar. New York, NY, USA: Association for Computing Machinery; 2024. p. 2454–69. https://doi.org/10.1145/3630106.3659048.
[38] Lan Y-J, Chen N-S. Teachers' agency in the era of LLM and generative AI: designing pedagogical AI agents. Educ Technol Soc 2024;27:I–XVIII.
[39] Shi H, Fang L, Chen X, et al. Review of the opportunities and challenges to accelerate mass-scale application of smart grids with large-language models. IET Smart Grid 2024;7:737–59. https://doi.org/10.1049/stg2.12191.
[40] Sharma A, Li X, Guan H, et al. Automatic data transformation using large language model - an experimental study on building energy data. 2023 IEEE Int Conf Big Data BigData 2023:1824–34. https://doi.org/10.1109/BigData59044.2023.10386931.
[41] Ni B, Cai X, Shen Z, et al. Intelli-dispatch-SQL: an LLM-based agent for reliable text-to-SQL in power dispatching. Energy AI 2025;22:100591. https://doi.org/10.1016/j.egyai.2025.100591.
[42] Liu T, Zanocco C, Wang Z, et al. Large language model enabled knowledge discovery of building-level electrification using permit data. Energy Build 2025;343:115890. https://doi.org/10.1016/j.enbuild.2025.115890.
[43] Li L, Wang C, Gan J, et al. A fuel consumption-based method for developing local-specific CO2 emission rate database using open-source big data. J Big Data 2024;11:74. https://doi.org/10.1186/s40537-024-00932-7.
[44] Yang H, Wu R, Xu W. TransCompressor: LLM-powered multimodal data compression for smart transportation. In: Proc. 30th annu. Int. Conf. Mob. Comput. Netw. Washington D.C. DC USA: ACM; 2024. p. 2335–40. https://doi.org/10.1145/3636534.3698120.
[45] Choi S, Yoon S. GPT-based data-driven urban building energy modeling (GPT-UBEM): concept, methodology, and case studies. Energy Build 2024;325:115042. https://doi.org/10.1016/j.enbuild.2024.115042.
[46] Hu Y, Kim H, Ye K, et al. Applying fine-tuned LLMs for reducing data needs in load profile analysis. Appl Energy 2025;377:124666. https://doi.org/10.1016/j.apenergy.2024.124666.
[47] Gu X, Chen C, Fang Y, et al. CECA: an intelligent large-language-model-enabled method for accounting embodied carbon in buildings. Build Environ 2025;272:112694. https://doi.org/10.1016/j.buildenv.2025.112694.
[48] Tupayachi J, Xu H, Omitaomu OA, et al. Towards next-generation urban decision support systems through AI-powered construction of scientific ontology using large language Models—A case in optimizing intermodal freight transportation. Smart Cities 2024;7:2392–421. https://doi.org/10.3390/smartcities7050094.
[49] Zhang L, Chen Z, Ford V. Advancing building energy modeling with large language models: exploration and case studies. Energy Build 2024;323:114788. https://doi.org/10.1016/j.enbuild.2024.114788.
[50] Zhang Y, Wang D, Wang G, et al. Data-driven building load prediction and large language models: comprehensive overview. Energy Build 2025;326:115001. https://doi.org/10.1016/j.enbuild.2024.115001.
[51] Jiang G, Ma Z, Zhang L, et al. Prompt engineering to inform large language model in automated building energy modeling. Energy 2025;316:134548. https://doi.org/10.1016/j.energy.2025.134548.
[52] Zhang L, Ford V, Chen Z, et al. Automatic building energy model development and debugging using large language models agentic workflow. Energy Build 2025;327:115116. https://doi.org/10.1016/j.enbuild.2024.115116.
[53] Jiang G, Ma Z, Zhang L, et al. EPlus-LLM: a large language model-based computing platform for automated building energy modeling. Appl Energy 2024;367:123431. https://doi.org/10.1016/j.apenergy.2024.123431.
[54] Xiang J, Ye T, Liu P, et al. ModiGen: a large language model-based workflow for multi-task modelica code generation. https://doi.org/10.48550/arXiv.2503.18460; 2025.
[55] Zhao S, Chen S, Zhou J, et al. Potential to transform words to watts with large language models in battery research. Cell Rep Phys Sci 2024;5. https://doi.org/10.1016/j.xcrp.2024.101844.
[56] Buster G, Pinchuk P, Barrons J, et al. Supporting energy policy research with large language models: a case study in wind energy siting ordinances. Energy AI 2024;18:100431. https://doi.org/10.1016/j.egyai.2024.100431.
[57] Arslan M, Mahdjoubi L, Munawar S. Driving sustainable energy transitions with a multi-source RAG-LLM system. Energy Build 2024;324:114827. https://doi.org/10.1016/j.enbuild.2024.114827.
[58] Han T, Cong R-G, Yu B, et al. Integrating local knowledge with ChatGPT-like large-scale language models for enhanced societal comprehension of carbon neutrality. Energy AI 2024;18:100440. https://doi.org/10.1016/j.egyai.2024.100440.
[59] Chaturvedi S, Chen EW, Sharma IP, et al. A generative AI approach to pricing mechanisms and consumer behavior in the electric vehicle charging market. Proc AAAI Symp Ser 2023;2:54–61. https://doi.org/10.1609/aaaiss.v2i1.27649.
[60] Pullanikkat R, Basu M, Ghosh S. Hear the commute: a generative AI-based framework to summarize transport grievances from social media. Int J Intell Transp Syst Res 2025. https://doi.org/10.1007/s13177-025-00502-y.
[61] Ashby C, Weir D, Fussey P. Understanding public views on electric vehicle charging: a thematic analysis. Transp Res Interdiscip Perspect 2025;29:101325. https://doi.org/10.1016/j.trip.2025.101325.
[62] Wang Z, Abolarin E, Wu K, et al. Beyond charging anxiety: an explainable approach to understanding user preferences of EV charging stations using review data. https://doi.org/10.48550/arXiv.2507.03243; 2025.
[63] Ziemba P, Majewski F. Using the large language model ChatGPT to support decisions in sustainable transport. Sustainability 2025;17:7520. https://doi.org/10.3390/su17167520.
[64] Carrasco-Aguilar Á, Carmona-Martínez MM, Parra-Meroño MC, et al. Integrating technical analysis into sentiment analysis: an ASTE framework for electric car purchase decision support based on LLMs and semantic BNF. Electronics 2025;14:1020. https://doi.org/10.3390/electronics14051020.
[65] Jiang J, Li Y, Nie J, et al. Integrating large language models with cross-modal data fusion for advanced intelligent transportation systems in sustainable cities

development. Appl Soft Comput 2025;177:113278. https://doi.org/10.1016/j.asoc.2025.113278.
[66] Wang L, Ren Y, Jiang H, et al. AccidentGPT: a V2X environmental perception multi-modal large model for accident analysis and prevention. 2024 IEEE Intell Veh Symp IV 2024:472–7. https://doi.org/10.1109/IV55156.2024.10588374.
[67] Gao M, Zhou S, Gu W, et al. MMGPT4LF: leveraging an optimized pre-trained GPT-2 model with multi-modal cross-attention for load forecasting. Appl Energy 2025;392:125965. https://doi.org/10.1016/j.apenergy.2025.125965.
[68] Bian C, Duan Z, Hao Y, et al. Exploring large language model for generic and robust state-of-charge estimation of li-ion batteries: a mixed prompt learning method. Energy 2024;302:131856. https://doi.org/10.1016/j.energy.2024.131856.
[69] Bian C, Han X, Duan Z, et al. Hybrid prompt-driven large language model for robust state-of-charge estimation of multitype li-ion batteries. IEEE Trans Transp Electrif 2025;11:426–37. https://doi.org/10.1109/TTE.2024.3391938.
[70] Zhu G, Jia W, Xing Z, et al. CMLLM: a novel cross-modal large language model for wind power forecasting. Energy Convers Manag 2025;330:119673. https://doi.org/10.1016/j.enconman.2025.119673.
[71] Feng Z, Sun Y, Ning J, et al. Implementing a provincial-level universal daily industrial carbon emissions prediction by fine-tuning the large language model. Appl Energy 2025;383:125372. https://doi.org/10.1016/j.apenergy.2025.125372.
[72] Wang X, Feng M, Qiu J, et al. From news to forecast: integrating event analysis in LLM-Based time series forecasting with reflection. 2024.
[73] Chen R, Jiang H, Guo T, et al. Can large language models forecast carbon price movements? Evidence from Chinese carbon markets. https://doi.org/10.2139/ssrn.4984589; 2024.
[74] Qu H, Li H, You L, et al. ChatEV: predicting electric vehicle charging demand as natural language processing. Transp Res Part Transp Environ 2024;136:104470. https://doi.org/10.1016/j.trd.2024.104470.
[75] Jin M, Wang S, Ma L, et al. Time-LLM: time series forecasting by reprogramming large language models. https://doi.org/10.48550/arXiv.2310.01728; 2024.
[76] Huang X. Enhancing traffic prediction with textual data using large language models. https://doi.org/10.48550/arXiv.2405.06719; 2024.
[77] Karim MM, Shi Y, Zhang S, et al. Large language models and their applications in roadway safety and mobility enhancement: a comprehensive review. Artif Intell Transp 2025;1:100004. https://doi.org/10.1016/j.ait.2025.100004.
[78] Wang X, Fang M, Zeng Z, et al. Where would I go next? Large language models as human mobility predictors. https://doi.org/10.48550/arXiv.2308.15197; 2024.
[79] Gong L, Lin Y, Zhang X, et al. Mobility-LLM: learning visiting intentions and travel preferences from human mobility data with large language models. 2024.
[80] Liang Y, Liu Y, Wang X, et al. Exploring large language models for human mobility prediction under public events. Comput Environ Urban Syst 2024;112:102153. https://doi.org/10.1016/j.compenvurbsys.2024.102153.
[81] Jin Y, Yang R, Yi Z, et al. SurrealDriver: designing LLM-powered generative driver agent framework based on human drivers' driving-thinking data. 2024 IEEERSJ Int Conf Intell Robots Syst IROS 2024:966–71. https://doi.org/10.1109/IROS58592.2024.10802229.
[82] Yang R, Zhang X, Fernandez-Laaksonen A, et al. Driving style alignment for LLM-Powered driver agent. 2024 IEEERSJ Int. Conf. Intell. Robots Syst. IROS. 2024. p. 11318–24. https://doi.org/10.1109/IROS58592.2024.10802629.
[83] Wang J, Jiang R, Yang C, et al. Large language models as urban residents: an LLM agent framework for personal mobility generation. 2024.
[84] Ju C, Liu J, Sinha S, et al. TrajLLM: a modular LLM-Enhanced agent-based framework for realistic human trajectory simulation. Companion Proc. ACM Web Conf. 2025. Sydney NSW Australia: ACM; 2025. p. 2847–50. https://doi.org/10.1145/3701716.3715201.
[85] Wang X, Qiu J, Cheng Y, et al. Context-aware carbon market analysis using generative agents and retrieval-augmented generation. https://doi.org/10.2139/ssrn.5158291; 2025.
[86] Chen Y, Chi B, Li C, et al. Toward interactive next location prediction driven by large language models. IEEE Trans Comput Social Syst 2025:1–17. https://doi.org/10.1109/TCSS.2024.3522965.
[87] Feng J, Cui C, Zhang C, et al. Large language model based agent framework for electric vehicle charging behavior simulation. 2025 10th Asia Conf Power Electr Eng ACPEE 2025:2683–7. https://doi.org/10.1109/ACPEE64358.2025.11040586.
[88] Niu Z, Li J, Ai Q, et al. EV charging system considering power dispatching based on multi-agent LLMs and CGAN. IEEE Trans Intell Transport Syst 2025:1–13. https://doi.org/10.1109/TITS.2025.3564389.
[89] Li N, Gao C, Li M, et al. EconAgent: large language model-empowered agents for simulating macroeconomic activities. In: Ku L-W, Martins A, Srikumar V, editors. Proc. 62nd annu. Meet. Assoc. Comput. Linguist. Vol. 1 long pap. Bangkok, Thailand: Association for Computational Linguistics; 2024. p. 15523–36. https://doi.org/10.18653/v1/2024.acl-long.829.
[90] Wang B, Cai Z, Karim MM, et al. Traffic performance GPT (TP-GPT): Real-Time data informed intelligent ChatBot for transportation surveillance and management. 2024 IEEE 27th Int. Conf. Intell. Transp. Syst. ITSC; 2024. p. 460–7. https://doi.org/10.1109/ITSC58415.2024.10919940.
[91] Cheng Y, Zhao H, Zhou X, et al. A large language model for advanced power dispatch. Sci Rep 2025;15:8925. https://doi.org/10.1038/s41598-025-91940-x.
[92] Jia M, Cui Z, Hug G. Enhancing LLMs for power system simulations: a feedback-driven multi-agent framework. IEEE Trans Smart Grid 2025:1. https://doi.org/10.1109/TSG.2025.3589114. 1.
[93] Ying S, Li Z, Yu M. Beyond words: evaluating large language models in transportation planning. Geo Spatial Inf Sci 2025;0:1–23. https://doi.org/10.1080/10095020.2025.2493073.
[94] Zhang L, Chen Z. Large language model-based interpretable machine learning control in building energy systems. Energy Build 2024;313:114278. https://doi.org/10.1016/j.enbuild.2024.114278.
[95] Zhang L, Huang J, Chen C. Multi-condition building decarbonization using deep reinforcement learning and large language model. Energy Build 2025:115810. https://doi.org/10.1016/j.enbuild.2025.115810.
[96] Teimoori Z. LLM-enabled EV charging stations recommendation. https://doi.org/10.48550/arXiv.2505.01447; 2025.
[97] Wen J, Zhang R, Niyato D, et al. Generative AI for low-carbon artificial intelligence of things with large language models. IEEE Internet Things Mag 2025;8:82–91. https://doi.org/10.1109/IOTM.001.2400074.
[98] Zheng X, Jiang C, Wang H. Large language model-assisted planning of electric vehicle charging infrastructure with real-world case study. Sustain Energy Technol Assessments 2025;84:104723. https://doi.org/10.1016/j.seta.2025.104723.
[99] Mohammadi S, Hassan A, Haghighi R, et al. Large Language models for solving economic dispatch problem. 2025 IEEE energy convers. Conf. Congr. Expo. ECCE; 2025. p. 1–5. https://doi.org/10.1109/ECCE58356.2025.11259643.
[100] Zhang H, Yang G, Lei N, et al. Scenario-aware electric vehicle energy control with enhanced vehicle-to-grid capability: a multi-task reinforcement learning approach. Energy 2025;335:138189. https://doi.org/10.1016/j.energy.2025.138189.
[101] Zhang S, Fu D, Liang W, et al. TrafficGPT: viewing, processing and interacting with traffic foundation models. Transp Policy 2024;150:95–105. https://doi.org/10.1016/j.tranpol.2024.03.006.
[102] Masri S, Ashqar HI, Elhenawy M. Large language models (LLMs) as traffic control systems at urban intersections: a new paradigm. Vehicles 2025;7:11. https://doi.org/10.3390/vehicles7010011.
[103] Movahedi M, Choi J. The crossroads of LLM and traffic control: a study on large language models in adaptive traffic signal control. IEEE Trans Intell Transport Syst 2025;26:1701–16. https://doi.org/10.1109/TITS.2024.3498735.
[104] Da L, Gao M, Mei H, et al. Prompt to transfer: Sim-to-real transfer for traffic signal control with prompt learning. Proc AAAI Conf Artif Intell 2024;38:82–90. https://doi.org/10.1609/aaai.v38i1.27758.
[105] Moraga Á, de Curtò J, de Zarzà I, et al. AI-driven UAV and IoT traffic optimization: large language models for congestion and emission reduction in smart cities. Drones 2025;9:248. https://doi.org/10.3390/drones9040248.
[106] Aliazam M, Javadi A, Hosseini Monazzah AM, et al. MAPS: Energy-Reliability tradeoff management in autonomous vehicles through LLMs penetrated science. 2024 5th CPSSI Int. Symp. Cyber-Phys. Syst. Appl. Theory CPSAT; 2024. p. 1–8. https://doi.org/10.1109/CPSAT64082.2024.10745445.
[107] Lei N, Zhang H, Wang H, et al. An improved co-optimization of component sizing and energy management for hybrid powertrains interacting with high-fidelity model. IEEE Trans Veh Technol 2023;72:15585–96. https://doi.org/10.1109/TVT.2023.3296114.
[108] Wan H, Ruan J, Xia J, et al. The continuous training of machine learning-based energy management strategy for plug-in hybrid electric vehicle, part I: electric driving mode. Energy 2025;333:137467. https://doi.org/10.1016/j.energy.2025.137467.
[109] Lei N, Zhang H, Hu J, et al. Sim-to-real design and development of reinforcement learning-based energy management strategies for fuel cell electric vehicles. Appl Energy 2025;393:126030. https://doi.org/10.1016/j.apenergy.2025.126030.
[110] Lei N, Zhang H, Wang H, et al. Theory-constrained neural network with modular interpretability for fuel cell vehicle modeling. IEEE Trans Veh Technol 2025;74:8907–20. https://doi.org/10.1109/TVT.2025.3533486.
[111] Zhang H, Dong J, Lei N, et al. Optimal vehicle dynamics and powertrain control of carbon-free autonomous vehicles: large language model assisted heterogeneous-agent learning. Energy 2025;338:138786. https://doi.org/10.1016/j.energy.2025.138786.
[112] Cui C, Ma Y, Yang Z, et al. Large language models for autonomous driving (LLM4AD): concept, benchmark, experiments, and challenges. https://doi.org/10.48550/arXiv.2410.15281; 2025.
[113] Ji X, Zhang L, Zhang W, et al. LEMAD: LLM-empowered multi-agent system for anomaly detection in power grid services. Electronics 2025;14:3008. https://doi.org/10.3390/electronics14153008.
[114] Jiang S, Chai W, Zhang M, et al. Verifying energy generation via edge LLM for web3-based decentralized clean energy networks. Inf Fusion 2026;127:103752. https://doi.org/10.1016/j.inffus.2025.103752.
[115] Cheng Z-Q, Dong Y, Shi A, et al. SHIELD: LLM-driven schema induction for predictive analytics in EV battery supply chain disruptions. In: Dernoncourt F, Preoţiuc-Pietro D, Shimorina A, editors. Proc. 2024 conf. Empir. Methods nat. Lang. Process. Ind. Track. Miami, Florida: US: Association for Computational Linguistics; 2024. p. 303–33. https://doi.org/10.18653/v1/2024.emnlp-industry.24.
[116] Xu Z, Chen T, Huang Z, et al. Personalizing driver agent using large language models for driving safety and smarter human–machine interactions. IEEE Intell Transp Syst Mag 2025:2–18. https://doi.org/10.1109/MITS.2025.3551736.
[117] Zhang Y, Saber AM, Youssef A, et al. Grid-agent: an LLM-powered multi-agent system for power grid control. https://doi.org/10.48550/arXiv.2508.05702; 2025.
[118] Yang X, Lin C, Liu H, et al. RL2: reinforce large language model to assist safe reinforcement learning for energy management of active distribution networks.

IEEE Trans Smart Grid 2025;16:3419–31. https://doi.org/10.1109/TSG.2025.3568226.
[119] Jiang Y, Ning K, Pan Z, et al. Multi-modal time series analysis: a tutorial and survey. Proc. 31st ACM SIGKDD Conf. Knowl. Discov. Data Min. V2. New York, NY, USA: Association for Computing Machinery; 2025. p. 6043–53. https://doi.org/10.1145/3711896.3736567.
[120] Zheng L, Chen Z, He J, et al. MULAN: Multi-Modal causal structure learning and root cause analysis for microservice systems. Proc. ACM Web Conf. 2024. New York, NY, USA: Association for Computing Machinery; 2024. p. 4107–16. https://doi.org/10.1145/3589334.3645442.
[121] Cheng M, Chen Y, Liu Q, et al. InstrucTime: advancing time series classification with multimodal language modeling. In: Proc. Eighteenth ACM int. Conf. Web search data min. New York, NY, USA: Association for Computing Machinery; 2025. p. 792–800. https://doi.org/10.1145/3701551.3703499.
[122] Liu C, Xu Q, Miao H, et al. TimeCMA: towards LLM-empowered multivariate time series forecasting via cross-modality alignment. Proc. Thirty-Ninth AAAI Conf. Artif. Intell. Thirty-Seventh Conf. Innov. Appl. Artif. Intell. Fifteenth symp. Educ. Adv. Artif. Intell., vol. 39. AAAI Press; 2025. p. 18780–8. https://doi.org/10.1609/aaai.v39i18.34067.
[123] Shen C, Chen Z, Luo D, et al. Exploring multi-modal data with tool-augmented LLM agents for precise causal discovery. https://doi.org/10.48550/arXiv.2412.13667; 2025.
[124] Zhang K, Zhou F, Wu L, et al. Semantic understanding and prompt engineering for large-scale traffic data imputation. Inf Fusion 2024;102:102038. https://doi.org/10.1016/j.inffus.2023.102038.
[125] Li H, Hong T. A semantic ontology for representing and quantifying energy flexibility of buildings. Adv Appl Energy 2022;8:100113. https://doi.org/10.1016/j.adapen.2022.100113.
[126] Qi Y, Mai G, Zhu R, et al. EVKG: an interlinked and interoperable electric vehicle knowledge graph for smart transportation system. Trans GIS 2023;27:949–74. https://doi.org/10.1111/tgis.13064.
[127] Geohash. Wikipedia; 2025.
[128] Nie T, Sun J, Ma W. Exploring the roles of large language models in reshaping transportation systems: a survey, framework, and roadmap. https://doi.org/10.48550/arXiv.2503.21411; 2025.
[129] Ishihara S. Training data extraction from pre-trained language models: a survey. In: Ovalle A, Chang K-W, Mehrabi N, et al., editors. Proc. 3rd workshop trust. Nat. Lang. Process. Trust. 2023. Toronto, Canada: Association for Computational Linguistics; 2023. p. 260–75. https://doi.org/10.18653/v1/2023.trustnlp-1.23.
[130] Truong J-B, Maini P, Walls RJ, et al. Data-free model extraction. 2021 IEEECVF Conf. Comput. Vis. Pattern Recognit. CVPR. Nashville, TN: USA: IEEE; 2021. p. 4769–78. https://doi.org/10.1109/CVPR46437.2021.00474.
[131] Carlini N, Tramer F, Wallace E, et al. Extracting training data from large language models. https://doi.org/10.48550/arXiv.2012.07805; 2021.
[132] Liyanage UP, Ranaweera ND. Ethical considerations and potential risks in the deployment of large language models in diverse societal contexts. J Comput Soc Dyn 2023;8:15–25.
[133] Kumar A, Murthy SV, Singh S, et al. The ethics of interaction: mitigating security threats in LLMs. https://doi.org/10.48550/arXiv.2401.12273; 2024.
[134] Rho D, Kim T, Park M, et al. Encryption-friendly LLM architecture. https://doi.org/10.48550/arXiv.2410.02486; 2025.
[135] Dang Y, Huang K, Huo J, et al. Explainable and interpretable multimodal large language models: a comprehensive survey. https://doi.org/10.48550/arXiv.2412.02104; 2024.
[136] Du M, Liu N, Hu X. Techniques for interpretable machine learning. Commun ACM 2019;63:68–77. https://doi.org/10.1145/3359786.
[137] Danesh T, Ouaret R, Floquet P, et al. Hybridization of model-specific and model-agnostic methods for interpretability of neural network predictions: application to a power plant. Comput Chem Eng 2023;176:108306. https://doi.org/10.1016/j.compchemeng.2023.108306.
[138] Wen L, Yang X, Fu D, et al. On the road with GPT-4V(ision): early explorations of visual-language model on autonomous driving. https://doi.org/10.48550/arXiv.2311.05332; 2023.
[139] Chen GH, Chen S, Liu Z, et al. Humans or LLMs as the judge? A study on judgement biases. https://doi.org/10.48550/arXiv.2402.10669; 2024.
[140] Sreeram S, Wang T-H, Maalouf A, et al. Probing multimodal LLMs as world models for driving. https://doi.org/10.48550/arXiv.2405.05956; 2024.
[141] Singh C, Inala JP, Galley M, et al. Rethinking interpretability in the era of large language models. https://doi.org/10.48550/arXiv.2402.01761; 2024.
[142] Huang Y, Sun L, Wang H, et al. TrustLLM: trustworthiness in large language models. https://doi.org/10.48550/arXiv.2401.05561; 2024.
[143] Xie J, Chen Z, Zhang R, et al. Large multimodal agents: a survey. https://doi.org/10.48550/arXiv.2402.15116; 2024.
[144] Li Z, Lu Y, Mu Y, et al. Cog-GA: a large language models-based generative agent for vision-language navigation in continuous environments. https://doi.org/10.48550/arXiv.2409.02522; 2024.
[145] Huang L, Yu W, Ma W, et al. A survey on hallucination in large language models: principles, taxonomy, challenges, and open questions. ACM Trans Inf Syst 2025;43:1–55. https://doi.org/10.1145/3703155.
[146] Slobodkin A, Goldman O, Caciularu A, et al. The curious case of hallucinatory (un)answerability: finding truths in the hidden states of over-confident large language models. https://doi.org/10.48550/arXiv.2310.11877; 2023.
[147] Liu B, Ash JT, Goel S, et al. Exposing attention glitches with flip-flop language modeling. https://doi.org/10.48550/arXiv.2306.00946; 2023.
[148] Shi W, Han X, Lewis M, et al. Trusting your evidence: hallucinate less with context-aware decoding. https://doi.org/10.48550/arXiv.2305.14739; 2023.
[149] Koonce B. ResNet 50. In: Koonce B, editor. Convolutional neural netw. Swift tensorflow image recognit. Dataset categ. Berkeley: CA: Apress; 2021. p. 63–72. https://doi.org/10.1007/978-1-4842-6168-2_6.
[150] JiZiwei LeeNayeon, Frieske Rita, et al. Survey of hallucination in natural language generation. ACM Comput Surv 2023. https://doi.org/10.1145/3571730.
[151] Luo Z, Xu C, Zhao P, et al. Augmented large language models with parametric knowledge guiding. https://doi.org/10.48550/arXiv.2305.04757; 2023.
[152] Ji Z, Jiang M. A systematic review of electricity demand for large language models: evaluations, challenges, and solutions. Renew Sustain Energy Rev 2026;225:116159. https://doi.org/10.1016/j.rser.2025.116159.
[153] Electricity use in homes - U.S. energy information administration (EIA). n.d. https://www.eia.gov/energyexplained/use-of-energy/electricity-use-in-homes.php. [Accessed 2 February 2026].
[154] Us EPA O. Greenhouse gas emissions from a typical passenger vehicle n.d. https://www.epa.gov/greenvehicles/greenhouse-gas-emissions-typical-passenger-vehicle. [Accessed 2 February 2026].
[155] Mao Y, Yu X, Huang K, et al. Green edge AI: a contemporary survey. https://doi.org/10.48550/arXiv.2312.00333; 2024.
[156] Yang H, Wu R, Xu W. TransCompressor: LLM-powered multimodal data compression for smart transportation. In: Proc. 30th annu. Int. Conf. Mob. Comput. Netw. New York, NY, USA: Association for Computing Machinery; 2024. p. 2335–40. https://doi.org/10.1145/3636534.3698120.
[157] Dettmers T, Pagnoni A, Holtzman A, et al. QLORA: efficient finetuning of quantized LLMs n.d.
[158] Li J, Xu J, Huang S, et al. Large language model inference acceleration: a comprehensive hardware perspective. https://doi.org/10.48550/arXiv.2410.04466; 2025.
[159] DeepSeek-Ai, Guo D, Yang D, et al. DeepSeek-R1: incentivizing reasoning capability in LLMs via reinforcement learning. https://doi.org/10.48550/arXiv.2501.12948; 2025.
[160] Xu J, Guo Z, He J, et al. Qwen2.5-omni technical report. https://doi.org/10.48550/arXiv.2503.20215; 2025.
[161] Jiang P, Sonne C, Li W, et al. Preventing the immense increase in the life-cycle energy and carbon footprints of LLM-powered intelligent chatbots. Engineering 2024;40:202–10. https://doi.org/10.1016/j.eng.2024.04.002.
[162] Zhu X, Chen S, Liang X, et al. Next-generation generalist energy artificial intelligence for navigating smart energy. Cell Rep Phys Sci 2024;5. https://doi.org/10.1016/j.xcrp.2024.102192.
[163] Yin S, Fu C, Zhao S, et al. A survey on multimodal large language models. Natl Sci Rev 2024;11:nwae403. https://doi.org/10.1093/nsr/nwae403.
[164] Das D, Pargal S, Chakraborty S, et al. Why slammed the brakes on? Auto-annotating driving behaviors from adaptive causal modeling. 2022 IEEE Int. Conf. Pervasive Comput. Commun. Workshop Affil. Events PerCom Workshop; 2022. p. 587–92. https://doi.org/10.1109/PerComWorkshops53856.2022.9767397.
[165] Li S, Li H, Li X, et al. Causal intervention is what large language models need for spatio-temporal forecasting. IEEE Trans Cybern 2025;55:3825–37. https://doi.org/10.1109/TCYB.2025.3569333.
[166] Qi B, Chen X, Gao J, et al. Interactive continual learning: fast and slow thinking. 2024. p. 12882–92.
[167] A review on edge large language models: design, execution, and applications. ACM Comput Surv n.d.
[168] Qu G, Chen Q, Wei W, et al. Mobile edge intelligence for large language models: a contemporary survey. IEEE Commun Surv Tutor 2025;27:3820–60. https://doi.org/10.1109/COMST.2025.3527641.
[169] Uprety I, Zhao X. Edge-deployable LLMs for autonomous vehicle intelligence. In: Proc. Tenth ACMIEEE symp. Edge Comput., the Hilton Arlington national landing Arlington. VA USA: ACM; 2025. p. 1–7. https://doi.org/10.1145/3769102.3774639.
[170] Yao Y, Yu T, Zhang A, et al. Efficient GPT-4V level multimodal large language model for deployment on edge devices. Nat Commun 2025;16:5509. https://doi.org/10.1038/s41467-025-61040-5.
[171] Accelerating LLM and VLM inference for automotive and robotics with NVIDIA TensorRT edge-LLM. NVIDIA Tech Blog; 2026. https://developer.nvidia.com/blog/accelerating-llm-and-vlm-inference-for-automotive-and-robotics-with-nvidia-tensorrt-edge-llm/. [Accessed 3 February 2026].