Artificial Intelligence and Computing Innovations

Intelligence Science Publishing

https://www.ispublishing.com

ISSN: 3070-4944

Vo.6, No.4, 2026

Data-Driven Intelligent Systems: A Survey of Representation Learning, Predictive Modeling, Causal Reasoning, and AI-Driven Decision Optimization

Liora Veyne[1*]

[1]Department of Computer Science, Lakehead University, Thunder Bay, Canada

**Corresponding author: lv878@lakeheadu.ca*

Abstract: Data-driven intelligent systems have become an important research direction in modern artificial intelligence because they connect large-scale data processing, representation learning, predictive modeling, causal reasoning, and adaptive decision optimization into a unified computational workflow. Different from traditional data analytics, which often emphasizes descriptive statistics or isolated prediction, recent data-driven systems are designed to learn from heterogeneous, dynamic, noisy, and high-dimensional data while supporting decisions under uncertainty. This survey reviews recent studies on data-driven intelligent systems and synthesizes the literature from a methodological perspective rather than from a domain-by-domain perspective. The reviewed works are organized around representation learning, sequential and spatiotemporal prediction, counterfactual and causal reasoning, anomaly and risk modeling, large language model agents, and system-level challenges. The survey shows that Transformer models, graph neural networks, contrastive learning, continual learning, retrieval-augmented modeling, and agentic large language model frameworks are becoming key foundations for robust data-driven intelligence. Meanwhile, several limitations remain unresolved, including data heterogeneity, distribution shift, interpretability, causal validity, evaluation beyond accuracy, and governance of automated decision pipelines. The paper concludes by identifying future directions toward unified representation-reasoning frameworks, drift-aware decision systems, auditable AI pipelines, human-in-the-loop decision coordination, and more comprehensive benchmarks for evaluating data-driven intelligent systems.

Keywords: data-driven intelligent systems, representation learning, predictive modeling, causal reasoning, anomaly detection, large language model agents, decision optimization, artificial intelligence.

1 INTRODUCTION

Data-driven intelligent systems are becoming a central paradigm for the design of modern artificial intelligence applications. In earlier computational systems, data was often treated as an auxiliary resource used to estimate parameters, validate assumptions, or generate reports after decisions had already been defined by human analysts. Recent intelligent systems adopt a different logic. They use data as the primary

basis for discovering latent structures, learning predictive relationships, estimating uncertainty, identifying abnormal states, reasoning about interventions, and optimizing downstream actions. This transformation has made data-driven intelligence a broad but coherent research direction that connects machine learning, statistical inference, representation learning, causal modeling, optimization, and increasingly large language model based reasoning.

The growth of data-driven intelligence is closely related to three technical shifts. First, the data available to intelligent systems has become more heterogeneous. Modern datasets may combine structured tables, multivariate time series, graphs, text, logs, trajectories, sensor measurements, user interactions, and domain-specific metadata. Such data cannot be handled effectively by simple feature engineering alone. Second, model architectures have become more expressive. Transformer models, graph neural networks, self-supervised learning, contrastive objectives, and retrieval-based mechanisms allow systems to represent complex dependencies that are difficult to capture with shallow models. Third, the objective of data-driven systems has moved beyond prediction accuracy. In many decision settings, systems must also remain robust under distribution shift, provide interpretable evidence, support counterfactual evaluation, and adapt continuously to changing data environments.

Because of these shifts, data-driven intelligent systems should not be understood as a collection of independent applications. Rather, they can be viewed as a methodological framework in which raw observations are transformed into learned representations, representations are used to support prediction and reasoning, and reasoning outputs are converted into optimized decisions. This perspective makes it possible to study diverse recent works under a unified structure. For example, sequential modeling studies emphasize temporal dependency and dynamic state evolution; causal studies emphasize counterfactual utility and robust intervention; anomaly detection studies emphasize rare-event identification and risk reasoning; and large language model agent studies emphasize semantic abstraction, memory, planning, and auditable decision traces [1]-[4], [10], [14], [16], [18].

This paper does not aim to provide a detailed domain-by-domain catalogue of data-driven applications. Instead, it synthesizes the literature from a methodological perspective and examines how representation learning, predictive modeling, causal reasoning, anomaly detection, and LLM-based agents jointly reshape the design of modern data-driven intelligent systems. This organization is intentionally broad enough to cover multiple research threads while avoiding a fragmented discussion of separate business or engineering domains. The main contribution of this survey is therefore a structured conceptual review that identifies the core technical layers, representative methodological patterns, major limitations, and future research directions of data-driven intelligent systems.

The remainder of this paper is organized as follows. Section II introduces the conceptual foundation and general architecture of data-driven intelligent systems. Section III reviews representation learning for heterogeneous and dynamic data. Section IV discusses predictive modeling and sequential intelligence. Section V analyzes causal reasoning and counterfactual decision optimization. Section VI summarizes anomaly detection, risk modeling, and robust intelligence. Section VII discusses large language models and agentic decision intelligence. Section VIII identifies system-level challenges. Section IX presents future research directions, and Section X concludes the paper.

TABLE I. METHODOLOGICAL ORGANIZATION OF REVIEWED STUDIES

Technical Stream	Representative Focus	Related References
Representation learning	Self-supervised learning, graph dependency modeling, spatiotemporal representation, shared latent spaces	[12], [28], [30], [34], [35], [40], [41]
Predictive and sequential modeling	Time-series forecasting, non-stationary modeling, dynamic distribution monitoring, multi-task forecasting	[4], [7], [25], [26], [27], [34], [41]
Causal and counterfactual reasoning	Counterfactual guidance, causal ranking, exposure bias correction, interpretable decision support	[1], [2], [6], [8], [11], [22], [37]
Anomaly and risk intelligence	Fraud detection, retrieval-augmented anomaly detection, semantic risk reasoning, few-shot detection	[9], [10], [20], [24], [32], [38], [40]
LLM and agentic decision intelligence	Multi-agent reasoning, memory-driven planning, semantic mapping, auditable attribution	[14], [16], [18], [32], [38]

2 CONCEPTUAL FOUNDATION OF DATA-DRIVEN INTELLIGENT SYSTEMS

A data-driven intelligent system can be defined as a computational system that learns from data to support adaptive reasoning and decision-making. This definition is broader than conventional supervised learning because the goal is not limited to mapping inputs to labels. A complete data-driven system usually includes data acquisition, preprocessing, representation construction, predictive inference, decision optimization, feedback integration, and continuous adaptation. Each layer introduces a different technical challenge. Data acquisition must handle noise and incompleteness; representation learning must preserve useful information while reducing irrelevant variation; prediction must estimate future or hidden states; reasoning must evaluate alternative actions; and optimization must select decisions that remain useful under uncertainty.

From a system perspective, data-driven intelligence can be represented as a pipeline that converts heterogeneous observations into increasingly structured and actionable forms. Raw observations are first cleaned, aligned, normalized, and transformed into features or embeddings. Representation models then map these features into latent spaces that encode temporal, semantic, structural, or causal relationships. Predictive and causal modules infer outcomes, risks, or utility changes. Decision optimization modules translate inference results into actions, rankings, alerts, or policy recommendations. Finally, feedback from the environment is used to update the system, detect drift, and improve future decisions. This pipeline is shown in Fig. 1.

This abstract architecture is useful because it highlights the difference between data-driven systems and isolated models. A model may be trained once on a static dataset and evaluated only by accuracy, but an intelligent system must coordinate multiple components across time. It must decide what data

should be trusted, what patterns should be extracted, which predictions are reliable, how uncertainty should be communicated, and how decisions should be revised when the environment changes. Recent literature increasingly reflects this broader system-level view, especially in work involving spatiotemporal modeling, causal ranking, retrieval-augmented anomaly detection, multi-agent reasoning, and memory-driven planning [2], [10], [16], [30], [34], [41].

The conceptual foundation also suggests that data-driven intelligence is not simply a matter of using larger datasets or deeper models. Larger datasets may improve coverage, but they may also introduce bias, redundancy, temporal drift, and conflicting signals. Deeper models may improve representation capacity, but they may reduce interpretability and increase deployment complexity. Therefore, the central question is how to design a system architecture that can convert complex data into reliable decisions. This question requires attention to representation, prediction, causality, robustness, explainability, and adaptation as interdependent design goals rather than isolated evaluation metrics.

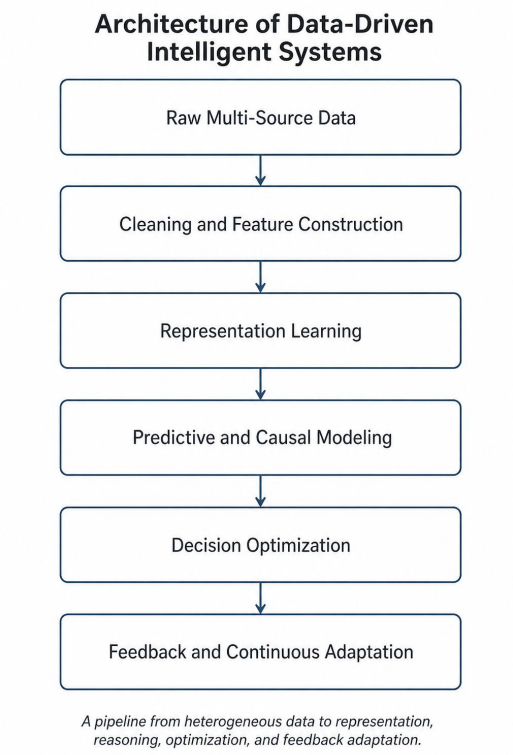


Fig. 1. General architecture of data-driven intelligent systems.

3 REPRESENTATION LEARNING FOR HETEROGENEOUS AND DYNAMIC DATA

Representation learning is one of the most important foundations of data-driven intelligent systems because it determines how raw data becomes usable for downstream reasoning. Traditional data-driven methods relied heavily on handcrafted variables designed by domain experts. Although such features can be effective in narrow contexts, they often fail when the data becomes high-dimensional, multi-source, graph-structured, or non-stationary. Modern representation learning addresses this limitation by learning compact and informative embeddings that capture hidden relations among observations, temporal states, entities, and events.

Recent studies illustrate several patterns in representation learning. Multivariate time-series modeling has combined graph neural networks and Transformer models to capture both variable-level dependencies and long-range temporal dynamics [12]. Self-supervised representation learning has been used to mine structured knowledge from heterogeneous multi-source data without relying entirely on manually labeled examples [28]. Spatiotemporal risk representation learning integrates Transformer architectures with graph structures to model dependencies across both space and time [30]. Shared representation learning has also been applied to high-dimensional multi-task forecasting under resource contention, suggesting that representation sharing can improve efficiency when multiple predictive tasks compete for limited computational resources [34].

Graph-based representation is especially important for data-driven systems because many real datasets contain relational structures. Dependencies among variables, services, users, components, or events cannot always be represented as independent feature vectors. Graph neural networks and graph-aware attention mechanisms provide a way to model interactions, propagation effects, and structural dependencies. In anomaly detection and risk modeling, for example, graph-based dependency learning can reveal abnormal relationships that may not be visible from individual feature distributions. Contrastive learning based dependency modeling further improves this process by encouraging the model to separate normal and abnormal relational patterns in latent space [40].

Transformer-based representation learning provides another major direction. Transformers can model long-range dependencies through attention mechanisms and can be adapted to sequences, graphs, tabular data, text, and multimodal inputs. Their value in data-driven systems lies not only in accuracy but also in flexibility. A Transformer can learn dependencies across time steps, entities, regions, or semantic units, making it useful for sequential forecasting, fine-grained discovery, and spatiotemporal modeling [7], [30], [35]. However, Transformer models also introduce challenges related to data scale, computational cost, and interpretability. A high-capacity representation may perform well on benchmark data but remain difficult to explain or deploy in resource-constrained environments.

Another emerging pattern is the use of self-supervised and contrastive objectives. In many data-driven settings, labeled data is limited, expensive, noisy, or delayed. Self-supervised learning provides a method for extracting general structure from unlabeled data by designing auxiliary tasks such as reconstruction, masked prediction, temporal ordering, graph neighborhood prediction, or instance discrimination. Contrastive learning extends this logic by pulling semantically similar observations closer and pushing dissimilar observations apart. These methods are valuable for building reusable representations that support anomaly detection, classification, retrieval, and decision optimization without requiring exhaustive manual annotation.

Overall, representation learning changes the role of data preprocessing in intelligent systems. Instead of treating preprocessing as a fixed manual step, modern systems increasingly treat representation construction as a trainable and adaptive component. This shift is important because the quality of downstream prediction, causal reasoning, and decision optimization depends heavily on the structure of the learned representation. A weak representation may hide relevant dependencies, while an overly complex representation may overfit noise or reduce interpretability. Future data-driven systems

will likely require representation modules that are not only accurate but also stable, interpretable, transferable, and aligned with decision objectives.

4 PREDICTIVE MODELING AND SEQUENTIAL INTELLIGENCE

Prediction remains a central component of data-driven intelligent systems, but the meaning of prediction has expanded considerably. Earlier predictive models often focused on static classification or regression. Recent systems increasingly operate in sequential and dynamic environments where observations arrive over time, distributions change, and decisions must be updated continuously. This requires models that can capture temporal dependencies, detect non-stationarity, estimate uncertainty, and support multi-step reasoning.

Sequential intelligence refers to the ability of a system to represent and reason over ordered observations. Session-level decision intelligence, multivariate time-series forecasting, adaptive anomaly detection, and spatiotemporal modeling all rely on sequential representations [4], [12], [25], [26], [41]. In these settings, a single observation is rarely sufficient. The system must infer the current state from historical patterns, recent changes, and long-range dependencies. Transformer models are particularly useful because they can attend to different segments of a sequence, while recurrent or hybrid models can capture local temporal continuity. Graph-based extensions further allow the system to model dependencies among multiple time-varying variables.

Non-stationarity is one of the major problems in sequential prediction. A model trained on historical data may become unreliable when the data-generating process changes. Recent adaptive anomaly detection frameworks address this issue by monitoring distributional changes and updating model behavior under non-stationary conditions [26]. Similarly, dependency drift awareness in spatiotemporal modeling reflects the idea that relationships among variables may evolve rather than remain fixed [41]. These studies suggest that data-driven systems should not assume a stable environment. Instead, they need drift detection, continual learning, and feedback adaptation as core components.

Uncertainty modeling is another important direction. In many sequential systems, the model must not only predict an expected value but also quantify the confidence or uncertainty of that prediction. Time-series approaches to uncertain power output, end-to-end weather prediction, and high-dimensional forecasting illustrate the importance of uncertainty-aware modeling in dynamic settings [7], [25], [34]. Uncertainty estimates can support more cautious decisions, trigger human review, or determine when a model should abstain from automated action. Without uncertainty modeling, data-driven systems may appear confident even when operating outside their training distribution.

Predictive modeling also increasingly involves multi-task learning. A single system may need to forecast several related outcomes, classify states, detect anomalies, and estimate risks simultaneously. Shared representation learning allows different tasks to benefit from common latent structures, but it also introduces resource contention and task interference. The design problem is therefore not only how to maximize performance on one task but also how to allocate representational and computational capacity across multiple objectives. This issue becomes more important in cloud-native and large-scale intelligent systems, where prediction must be efficient as well as accurate [34].

From a methodological perspective, the key evolution is from static prediction to adaptive sequential intelligence. Static prediction treats data as an independent sample set; sequential intelligence treats data as an evolving process. Static prediction often evaluates performance with accuracy, F1 score, or mean error; sequential intelligence also requires stability, drift sensitivity, calibration, latency, and decision utility. This broader evaluation view is necessary because a system may be accurate on average while failing during rare transitions, sudden distribution shifts, or high-risk states. Therefore, robust sequential prediction must be integrated with monitoring and adaptation rather than deployed as a fixed model.

5 CAUSAL REASONING AND COUNTERFACTUAL DECISION OPTIMIZATION

Although prediction is essential, prediction alone is insufficient for intelligent decision-making. A system may accurately estimate what is likely to happen while still failing to identify which action should be taken. This limitation motivates the growing importance of causal reasoning and counterfactual decision optimization. Causal approaches attempt to distinguish correlation from intervention effects, while counterfactual reasoning asks what would have happened under alternative actions. Together, they extend data-driven systems from passive prediction to active decision support.

Recent research on counterfactual guidance, calibrated ranking, exposure bias correction, and causal representation learning reflects this transition [1], [2], [6], [8]. In such work, the goal is not only to fit observed data but also to estimate how different choices might change outcomes. Counterfactual reasoning is especially important in interactive systems because the available data is shaped by previous decisions. A recommendation system observes only the items that were exposed; a ranking system observes feedback under a particular ordering; a training or guidance system observes behavior under previous instructions. Without causal correction, the system may reinforce historical bias rather than identify genuinely effective interventions.

Causal representation learning provides a bridge between representation learning and decision optimization. Instead of learning arbitrary latent embeddings, causal representation learning attempts to organize representations around variables or factors that have meaningful relationships with outcomes and interventions. This can improve interpretability and robustness because the model is encouraged to capture stable causal mechanisms rather than superficial correlations. In risk identification, causal representations may help distinguish indicators that are merely correlated with risk from mechanisms that contribute more directly to risk formation [6].

Multi-objective decision optimization further complicates the problem. Modern data-driven systems often balance accuracy, cost, fairness, latency, exposure, user utility, and robustness. A calibrated multi-objective ranking system, for example, must evaluate not only predicted relevance but also counterfactual utility under different ranking policies [2]. Robust denoising and sample reweighting methods similarly attempt to reduce the influence of biased, noisy, or semantically inconsistent samples [22]. These methods show that data-driven optimization increasingly requires explicit treatment of objective conflict and data bias.

Interpretability is also closely connected to causal reasoning. Fully data-driven but interpretable behavioral modeling demonstrates the need for models that combine flexible learning with transparent

decision mechanisms [11]. Policy-making frameworks similarly require decision support that can be explained to stakeholders rather than treated as a black-box output [37]. In high-stakes settings, the value of a data-driven model is limited if users cannot understand why a recommendation was generated, whether it is robust to confounding, or how it would change under alternative assumptions.

The main challenge is that causal reasoning usually requires stronger assumptions than predictive modeling. Observational data may be confounded, interventions may be unobserved, and counterfactual outcomes cannot be directly verified. Therefore, causal data-driven systems must combine statistical estimation with domain knowledge, sensitivity analysis, and careful evaluation. Even when causal claims are not fully proven, counterfactual thinking can improve system design by forcing researchers to consider how data was generated, which biases are present, and how decisions affect future observations.

6 ANOMALY DETECTION, RISK MODELING, AND ROBUST INTELLIGENCE

Anomaly detection and risk modeling represent another major layer of data-driven intelligent systems. These tasks are important because intelligent systems often operate in environments where rare events, abnormal patterns, or emerging risks matter more than average behavior. A model that performs well on common cases may still fail when facing fraud, system faults, distribution drift, adversarial behavior, or unusual combinations of features. Therefore, anomaly detection is not only a classification problem but also a form of robust intelligence.

Recent anomaly and risk modeling studies use ensemble learning, hybrid deep learning, hierarchical data fusion, retrieval-augmented detection, contrastive dependency modeling, meta-learning, and large language model based reasoning [9], [10], [20], [24], [32], [38], [40]. These methods reflect a shift away from simple threshold-based anomaly detection. Instead of assuming that anomalies are isolated deviations from a fixed normal distribution, modern approaches model context, dependencies, temporal evolution, semantic evidence, and few-shot generalization. This is important because abnormal events often appear as subtle changes in relationships rather than obvious outliers in individual variables.

Context-aware retrieval-augmented anomaly detection is a representative direction. Retrieval mechanisms can bring relevant historical cases, similar patterns, or supporting evidence into the detection process [10]. This allows the system to compare a current observation with contextual references rather than relying only on a fixed model representation. In complex tabular or business data, such contextual comparison may improve robustness when feature distributions evolve or when anomalies resemble rare but legitimate patterns. Retrieval also improves explainability because the system can provide examples or evidence paths that support its decision.

LLM-based risk reasoning extends anomaly detection by introducing semantic interpretation. Instead of treating all inputs as numerical features, large language models can process descriptions, documents, explanations, and multi-document evidence. Intelligent auditing and few-shot fraud detection studies illustrate how semantic mapping and reasoning can be combined with conventional detection pipelines [32], [38]. This is especially useful when abnormality depends on textual context, policy interpretation, or relationships across multiple documents. However, LLM-based anomaly detection also raises concerns about hallucination, consistency, and evaluation reliability.

Continual learning is another important requirement for robust anomaly detection. Many detection systems face non-stationary data, where normal behavior changes over time. A fixed anomaly detector may gradually classify legitimate new patterns as abnormal or fail to detect emerging risks. Continual learning frameworks with dynamic distribution monitoring address this issue by updating the model as data evolves [26]. Dependency drift awareness plays a similar role by recognizing that relationships among variables may change over time [41]. These approaches suggest that anomaly detection should be treated as a dynamic monitoring problem rather than a one-time model training task.

Risk modeling differs from anomaly detection in that it often estimates the probability or severity of future harm. Data-driven hedging with generative models, risk representation learning, and high-risk prioritization methods show how predictive and generative modeling can support risk-aware decision-making [21], [30], [31]. In these settings, uncertainty is central. A system must estimate not only whether an event is abnormal but also how serious it may be, what action should be taken, and how confident the system is. Robust intelligence therefore requires calibration, uncertainty estimation, and decision-sensitive evaluation.

The broader implication is that anomaly detection and risk modeling are moving toward integrated reasoning pipelines. Future systems will likely combine learned representations, temporal monitoring, retrieval, causal analysis, LLM-based explanation, and human review. Such systems may be more complex than conventional detectors, but they can provide richer evidence and more adaptable behavior. The challenge is to ensure that this complexity does not reduce transparency or create new failure modes.

7 LARGE LANGUAGE MODELS AND AGENTIC DECISION INTELLIGENCE

Large language models are increasingly changing the structure of data-driven intelligent systems. Their role is not limited to text generation. In decision-oriented systems, LLMs can serve as semantic interpreters, planners, evaluators, memory managers, attribution analyzers, and multi-agent coordinators. This creates a new form of agentic decision intelligence in which models can reason over natural language instructions, retrieve prior evidence, coordinate subtasks, and produce human-readable explanations.

Multi-agent LLM frameworks demonstrate how agentic systems can support decision-making through role decomposition and auditable attribution analysis [14]. Instead of relying on a single monolithic model, a multi-agent framework can assign different reasoning responsibilities to different agents, such as evidence collection, strategy generation, critique, verification, or attribution. This structure may improve transparency because intermediate outputs can be inspected, although it also introduces coordination complexity. The quality of the final decision depends not only on individual model capability but also on interaction protocols, memory design, and evaluation rules.

Memory-driven planning is another important direction. Long-horizon tasks require systems to maintain relevant information across many steps. Hierarchical encoding and dynamic retrieval provide mechanisms for selecting useful memory rather than storing all previous context equally [16]. This is important because unstructured memory can become noisy, redundant, or misleading. Effective memory systems must decide what to store, what to retrieve, how to update past information, and how to avoid

propagating outdated evidence. In data-driven intelligent systems, memory can connect historical observations with current decision contexts.

Constraint-guided policy learning also connects agentic intelligence with decision optimization. Structured state representation and constraint-guided policies provide a way to combine learned representations with explicit operational or logical constraints [18]. This is important because many real decisions cannot be optimized purely by maximizing a learned reward. They may need to respect budget limits, safety requirements, interpretability constraints, fairness considerations, or human oversight rules. LLM agents can potentially help express and reason about such constraints, but they must be grounded in reliable data and verifiable procedures.

LLM-based decision systems also raise new challenges. Unlike conventional predictive models, LLMs may generate fluent but unsupported explanations. They may be sensitive to prompt wording, context length, retrieved evidence quality, and agent interaction patterns. Therefore, agentic decision intelligence requires stronger evaluation methods than simple output accuracy. Researchers need to evaluate reasoning trace quality, evidence consistency, attribution reliability, uncertainty expression, and robustness under adversarial or ambiguous inputs. Without such evaluation, LLM-based decision systems may appear persuasive while producing unstable decisions.

The long-term significance of LLM agents lies in their ability to connect numerical prediction with semantic reasoning. Traditional data-driven systems often output scores, labels, or rankings. Agentic systems can additionally explain why an output matters, compare alternatives, summarize evidence, and communicate with human users. This makes LLMs useful as interface and reasoning layers above conventional data-driven models. However, the best future systems are unlikely to rely on LLMs alone. More robust architectures will combine LLM reasoning with structured data models, causal inference, retrieval mechanisms, uncertainty estimation, and human-in-the-loop review.

8 SYSTEM-LEVEL CHALLENGES

8.1 Data Heterogeneity and Quality Uncertainty

Data heterogeneity remains a fundamental challenge. Data-driven intelligent systems may need to combine numerical features, sequences, graphs, text, logs, images, and external metadata. These sources differ in scale, noise level, sampling frequency, missingness, and semantic meaning. If the system does not properly align and normalize them, representation learning may encode artifacts rather than meaningful signals. Quality uncertainty is especially problematic because model outputs may appear precise even when the underlying data is incomplete, biased, or outdated.

Another difficulty is that data quality problems often interact with modeling choices. A powerful model may hide data issues by fitting noise, while a simpler model may expose instability more clearly. Therefore, data-driven systems need data validation, anomaly screening, missing-data handling, bias analysis, and feedback monitoring as part of the system architecture. Data quality should be treated as an ongoing process rather than a preprocessing step completed before model training.

8.2 Distribution Shift and Non-Stationary Environments

Distribution shift occurs when the environment in which a model is deployed differs from the environment represented by the training data. This is common in sequential, operational, and interactive systems. Changes may occur in user behavior, system load, sensor conditions, policy rules, economic patterns, or external events. Non-stationarity reduces the reliability of static models and makes one-time evaluation insufficient.

Drift-aware modeling, continual learning, and monitoring frameworks address part of this problem, but they also introduce new questions. A system must distinguish meaningful drift from temporary noise, decide when to update model parameters, and avoid catastrophic forgetting. It must also maintain auditability when the model changes over time. This suggests that future data-driven systems need explicit lifecycle management rather than isolated training and deployment stages.

8.3 Interpretability and Causal Validity

Interpretability is essential because data-driven systems increasingly support decisions rather than merely produce predictions. Users need to understand which signals influenced a decision, how reliable the output is, and whether alternative actions would lead to different outcomes. Attention weights, feature importance, retrieved examples, counterfactual explanations, and natural-language rationales can all support interpretability, but none of them automatically guarantee causal validity.

Causal validity is more demanding than explanation. A model may provide a plausible explanation while relying on spurious correlations. Causal reasoning requires assumptions about the data-generating process, confounding, interventions, and counterfactual outcomes. Therefore, interpretable data-driven systems should avoid presenting correlational explanations as proven causal mechanisms. A mature system should communicate uncertainty and assumptions clearly, especially when used for decision optimization.

8.4 Scalability and Deployment Constraints

Scalability is another system-level challenge. High-capacity models can improve representation and prediction, but they may be expensive to train and deploy. Large Transformers, graph models, retrieval systems, and multi-agent LLM pipelines may require significant computation, storage, and latency budgets. In real deployments, decision systems must often operate under strict response-time and resource constraints.

This creates a tradeoff between model complexity and operational usability. A highly accurate model may be impractical if it is too slow, too expensive, or too difficult to maintain. Shared representation learning, model compression, efficient retrieval, modular system design, and selective human review are possible responses. Evaluation should therefore include not only accuracy but also inference cost, latency, stability, maintainability, and integration difficulty.

8.5 Governance, Auditability, and Human Oversight

As data-driven systems become more autonomous, governance becomes increasingly important. A decision pipeline should record what data was used, what model version produced the output, what evidence supported the result, what uncertainty was present, and whether a human reviewer was

involved. This is especially important for LLM-based and agentic systems because their reasoning processes may be long, multi-step, and partially stochastic.

Human oversight should not be treated as a superficial approval step. Effective human-in-the-loop design requires defining when human review is triggered, what information is presented to the reviewer, how reviewer feedback is incorporated, and how disagreements between humans and models are resolved. Governance-aware data-driven systems must therefore combine technical monitoring with organizational accountability.

9 FUTURE RESEARCH DIRECTIONS

Future research on data-driven intelligent systems is likely to move toward unified representation and reasoning frameworks. Current systems often separate feature extraction, prediction, causal inference, and decision optimization into different modules. While modularity is useful, it may also create mismatches between learned representations and decision objectives. A promising direction is to design representations that are simultaneously predictive, causally meaningful, uncertainty-aware, and useful for downstream optimization.

Another important direction is continual and drift-aware decision systems. As environments change, intelligent systems must update their knowledge without losing stability. Future research should develop methods for detecting distribution shift, distinguishing transient noise from structural change, and updating models with minimal disruption. Evaluation benchmarks should include temporal drift, dependency changes, and delayed feedback rather than only static train-test splits.

Explainable and auditable AI decision pipelines will also become more important. Future systems should provide evidence paths, uncertainty estimates, counterfactual comparisons, and reasoning traces that can be inspected by users. This is especially relevant for LLM agents, which can generate explanations but may not always provide reliable grounding. Research is needed on how to evaluate explanation faithfulness, evidence consistency, and attribution quality in multi-step decision systems.

Human-in-the-loop and multi-agent decision coordination represent another promising direction. Instead of viewing automation and human expertise as opposites, future systems can combine machine prediction with human judgment. Multi-agent frameworks may assign different reasoning tasks to specialized agents, while human reviewers provide oversight for uncertain or high-impact decisions. The challenge is to design coordination mechanisms that improve reliability without increasing complexity unnecessarily.

Finally, evaluation must move beyond accuracy. Data-driven intelligent systems should be assessed using multiple criteria, including robustness, calibration, interpretability, fairness, latency, cost, adaptability, and decision utility. A model that is slightly less accurate but more stable, explainable, and cheaper to deploy may be more valuable than a model with the highest benchmark score. Future benchmarks should therefore reflect system-level performance rather than isolated model performance.

10 METHODOLOGICAL EXTENSIONS

Recent research on parameter-efficient adaptation and representation learning reinforces the importance of selective knowledge incorporation in data-driven intelligent systems. Semantically guided

low-rank adaptation and adapter-based knowledge injection show how pre-trained models can be specialized while retaining compact trainable components [42], [43]. Robust small-target detection and self-supervised graph neural feature extraction further demonstrate that structured feature interaction can improve discrimination under complex observation conditions [44], [45].

Applied prediction studies also show that representation quality depends on how heterogeneous evidence is organized. Neural survival prediction across cancer types, interpretable event sequence construction for multidimensional time series, neural architecture search, cytopathology image classification, and transfer-learning-based breast image classification all rely on layered feature reuse and task-aware abstraction [46], [47], [48], [49], [50]. Public accessibility linked-data construction similarly highlights the value of well-structured external information [51].

Temporal and semantic modeling becomes especially important in evolving data environments. Incremental distributed time-series learning, multi-feature recurrent sequence modeling, and causal consistency-aware recommendation retain complementary temporal and contextual signals rather than depending on a single static representation [52], [53], [54]. Calibrated cloud resource planning, temporal-structural fault discrimination, and active data curation further demonstrate how uncertainty, dependency structure, and evidence quality can support more dependable decisions [55], [56], [57].

Retrieval and verification mechanisms provide another route to trustworthy intelligence. Dual retrieval with adaptive reranking, long-text reasoning with semantic conflict detection, spatiotemporal attention for forecasting, and microstructure-aware risk modeling connect predictions to relevant evidence at different operational scales [58], [59], [60], [61]. Causal graph-augmented root-cause analysis and requirement-aware API testing make this evidence trail usable for diagnosis and system validation [62], [63].

Deployment quality also depends on resource-aware execution and post-generation verification. Adaptive compression across the edge-cloud continuum, workload-aware database configuration, and causal factual correction for summarization address efficiency and reliability simultaneously [64], [65], [66]. KV-cache-aware routing, safe resilience configuration, and retrieval-augmented operational twins extend these controls to backend services that must balance response time, recovery, and traceable diagnosis [67], [68], [69].

Graph-based system studies further emphasize the importance of relational structure. Topology dependency modeling for microservice anomalies, graph-based cloud security detection, and topology-aware fault identification distinguish abnormal behavior from normal variation through explicit dependencies [70], [71], [72]. Structure-aware financial summarization, contrastive financial anomaly detection, and verifiable graph reasoning illustrate how representation constraints can improve traceability in high-stakes analysis [73], [74], [75].

Decision systems increasingly combine optimization, routing, and collaborative reasoning. Risk-constrained cloud-native scaling, semantic-aware multi-model routing, and privacy-preserving multi-agent collaboration offer complementary controls for service quality and partial information [76], [77], [78]. Evidence-gated memory writing and knowledge-graph-enhanced sentiment classification show

how language-enabled systems can preserve useful context while controlling spurious associations [79], [80].

Finally, privacy-preserving federated anomaly detection, resource-contention classification, and reinforcement-learning scheduling show how shared infrastructure can be managed under privacy, efficiency, and multi-tenant constraints [81], [82], [83]. Similar structured integration is useful in biomedical prediction, where adverse drug reaction modeling combines knowledge graph embeddings with deep learning [84]. Together, these studies strengthen the methodological view of data-driven intelligent systems as adaptive pipelines that coordinate representation, evidence, decision-making, and deployment.

11 CONCLUSION

This survey reviewed data-driven intelligent systems from a methodological perspective, focusing on representation learning, predictive modeling, causal reasoning, anomaly detection, LLM agents, and decision optimization. The reviewed literature shows that data-driven intelligence is evolving from isolated predictive modeling toward integrated decision architectures. Modern systems must learn from heterogeneous data, model temporal and structural dependencies, reason under uncertainty, detect abnormal patterns, evaluate counterfactual alternatives, and support adaptive decisions.

The central research problem is no longer simply how to train a high-accuracy model on a fixed dataset. Instead, the key problem is how to construct robust, interpretable, adaptive, and auditable AI systems that can convert complex data into reliable decisions. Representation learning provides the foundation for extracting useful structure; sequential modeling supports dynamic prediction; causal reasoning connects prediction with intervention; anomaly detection provides robust monitoring; and LLM agents introduce semantic reasoning and planning capabilities. Together, these components define the emerging architecture of data-driven intelligent systems.

At the same time, major challenges remain. Data quality, distribution shift, interpretability, causal validity, scalability, governance, and evaluation all require further research. Future systems should combine predictive power with transparent reasoning, continuous adaptation, human oversight, and system-level evaluation. Such development will be necessary for data-driven intelligent systems to become reliable components of modern AI-enabled decision environments.

References

[1] J. Bae, H. Nam, K. Ryu, J. Lee, J. Kim, H. Chun, J. Han, and J. Choi, "Data-Driven Driver Training via Counterfactual and Language-Based Guidance in Racing Scenarios," IEEE Access, vol. 13, pp. 170181-170199, 2025, doi: 10.1109/ACCESS.2025.3615128.

[2] X. Yang, S. Sun, Y. Li, Y. Xing, M. Wang, and Y. Wang, "CaliCausalRank: Calibrated Multi-Objective Ad Ranking with Robust Counterfactual Utility Optimization," arXiv preprint arXiv:2602.18786, 2026.

[3] L. Huber, T. Kucera, S. Hollerer, et al., "Data-Driven Protease Engineering by DNA-Recording and Epistasis-Aware Machine Learning," Nature Communications, vol. 16, Art. no. 5466, 2025, doi: 10.1038/s41467-025-60622-7.

[4] Y. Wang, Y. Zhang, C. Tech, Y. Sun, Y. Yang, Z. Pan, and W. Wang, "A Data-Centric Sequential Modeling Framework for Session-Level Decision Intelligence," 2026.

[5] T. C. Vieira, A. Bonamigo, H. A. Ferenhof, and V. S. Rodrigues, "Data-Driven Decision Support for Dairy Producer Selection," IFAC-PapersOnLine, vol. 59, no. 10, pp. 685-690, 2025, doi: 10.1016/j.ifacol.2025.09.117.

[6] C. Chen, R. Fang, and J. Lai, "Causal Representation Learning for Robust and Interpretable Audit Risk Identification in Financial Systems," in Proc. 2025 7th Int. Conf. Economic Management and Model Engineering (ICEMME 2025), Springer Nature, Mar. 2026, p. 454.

[7] A. Allen et al., "End-to-End Data-Driven Weather Prediction," Nature, vol. 641, pp. 1172-1179, 2025, doi: 10.1038/s41586-025-08897-0.

[8] Y. Xing, "Enhancing Advertising Recommendation Performance via Integrated Causal Inference and Exposure Bias Correction," Journal of Computer Technology and Software, vol. 2, no. 3, 2023.

[9] M. O. Butt, A. Mir, and A. Sujjada, "A Novel Data-Driven Framework for Automated Migraines Classification Using Ensemble Learning," Engineering Proceedings, vol. 107, no. 1, Art. no. 25, 2025, doi: 10.3390/engproc2025107025.

[10] N. Chen, Y. Zhang, W. Wang, Z. Pan, Y. Wang, and Y. Lu, "CoReAD: Context-Aware Retrieval-Augmented Deep Anomaly Detection for Evolving Business Tabular Data," 2026.

[11] F. Makinoshima et al., "Fully Data-Driven but Interpretable Human Behavioural Modelling with Differentiable Discrete Choice Model," IEEE Access, vol. 13, pp. 117420-117429, 2025, doi: 10.1109/ACCESS.2025.3585172.

[12] J. Wang, "Multivariate Time Series Forecasting and Classification via GNN and Transformer Models," Journal of Computer Technology and Software, vol. 3, no. 9, 2024.

[13] K. Spanaki et al., "Data-Driven Digital Transformation in Operations and Supply Chain Management," International Journal of Production Economics, vol. 284, Art. no. 109599, 2025.

[14] Q. Liu, H. Cui, Y. Wang, C. S. Lee, S. Chen, and L. Yang, "A Multi-Agent Large Language Model Framework for Marketing Decision-Making with Auditable Attribution Analysis," 2026.

[15] L. Attia, J. W. Burns, P. S. Doyle, et al., "Data-Driven Organic Solubility Prediction at the Limit of Aleatoric Uncertainty," Nature Communications, vol. 16, Art. no. 7497, 2025, doi: 10.1038/s41467-025-62717-7.

[16] Y. Wang, R. Yan, Y. Xiao, J. Li, Z. Zhang, and F. Wang, "Memory-Driven Agent Planning for Long-Horizon Tasks via Hierarchical Encoding and Dynamic Retrieval," 2025.

[17] S. Z. S. Abdalla, "Data-Driven Innovations in Disaster Risk Management: Advancing Resilience and Sustainability Through Big Data Analytics," Progress in Disaster Science, vol. 27, Art. no. 100451, 2025, doi: 10.1016/j.pdisas.2025.100451.

[18] X. Liang, Q. Liu, S. Chen, H. Qiu, and H. Zhang, "Structured State Representation and Constraint-Guided Policy Learning for Intelligent Business Decision Systems," 2026.

[19] E. Chaves, A. Barontini, N. Mendes, and V. Compan, "Optimising Sensor Placement in Heritage Buildings: A Comparison of Model-Based and Data-Driven Approaches," Sensors, vol. 25, no. 13, Art. no. 4212, 2025, doi: 10.3390/s25134212.

[20] P. Feng, "Hybrid BiLSTM-Transformer Model for Identifying Fraudulent Transactions in Financial Systems," Journal of Computer Science and Software Applications, vol. 5, no. 3, 2025.

[21] R. Cont and M. Vuletic, "Data-Driven Hedging with Generative Models," Annals of Operations Research, 2025, doi: 10.1007/s10479-025-06867-3.

[22] X. Yang, Y. Wang, Y. Li, and S. Sun, "Semantics-Aware Denoising: A PLM-Guided Sample Reweighting Strategy for Robust Recommendation," arXiv preprint arXiv:2602.15359, 2026.

[23] M. Tufano, F. Duarte, M. Mazzarello, et al., "Data-Driven Nutritional Assessment of Urban Food Landscapes: Insights from Boston, London, and Dubai," Scientific Reports, vol. 15, Art. no. 24453, 2025, doi: 10.1038/s41598-025-08098-9.

[24] J. Wang, "Credit Card Fraud Detection via Hierarchical Multi-Source Data Fusion and Dropout Regularization," Transactions on Computational and Scientific Methods, vol. 5, no. 1, 2025.

[25] N. Talebpour and M. Ilbeigi, "Modeling and Predicting Uncertainty in Tidal Turbine Power Output: A Data-Driven Time-Series Approach," IEEE Access, vol. 13, pp. 82986-82994, 2025, doi: 10.1109/ACCESS.2025.3567884.

[26] Y. Ou, S. Huang, F. Wang, K. Zhou, and Y. Shu, "Adaptive Anomaly Detection for Non-Stationary Time-Series: A Continual Learning Framework with Dynamic Distribution Monitoring," 2025.

[27] B. Kivilcim, "A Data-Driven Approach for Urban Heat Island Predictions: Rethinking the Evaluation Metrics and Data Preprocessing," Urban Science, vol. 9, no. 5, Art. no. 151, 2025.

[28] S. Sun, R. Xu, L. Yang, J. Huang, and N. Chen, "Self-Supervised Representation Learning and Structured Knowledge Mining for Heterogeneous Multi-Source Data," in Proc. 2025 5th Int. Conf. Electronic Communication, Computer Science and Technology (ECCST), Dec. 2025, pp. 277-281.

[29] M. Yazdi, Artificial Intelligence in Workplace Health and Safety: Data-Driven Technologies, Tools and Techniques. Boca Raton, FL, USA: CRC Press, 2025.

[30] X. Liang, Y. Zhao, M. Chang, R. Zhou, K. Cao, and Y. Zheng, "Spatiotemporal Risk Representation Learning Using Transformers and Graph Structure," 2026.

[31] K. Demircan, J. Carrasco-Zanini, A. Williamson, et al., "Data-Driven Prioritization of High-Risk Individuals for Weight Loss Interventions," Nature Medicine, 2026, doi: 10.1038/s41591-026-04353-2.

[32] Q. Gan, "Large Language Model Framework for Multi-Document Financial Anomaly Detection in Intelligent Auditing via Semantic Mapping and Risk Reasoning," Transactions on Computational and Scientific Methods, vol. 4, no. 12, 2024.

[33] T. D. Bruno, E. B. Worup, A.-L. Andersen, R. Andersen, F. M. Assef, and K. Nielsen, "Data Driven Method for Product and Component Sorting for Remanufacturing," IFAC-PapersOnLine, vol. 59, no. 10, pp. 1868-1873, 2025, doi: 10.1016/j.ifacol.2025.09.314.

[34] Z. Huang, J. Yang, S. Li, C. Zhang, J. Chen, and C. Xu, "Shared Representation Learning for High-Dimensional Multi-Task Forecasting under Resource Contention in Cloud-Native Backends," arXiv preprint arXiv:2512.21102, 2025.

[35] A. J. Lee, A. Dubuc, M. Kunst, et al., "Data-Driven Fine-Grained Region Discovery in the Mouse Brain with Transformers," Nature Communications, vol. 16, Art. no. 8536, 2025, doi: 10.1038/s41467-025-64259-4.

[36] M. Wang, "Multi-Level Attention and Sequence Modeling for Dynamic User Interest Representation in Real-Time Advertising Recommendation," Transactions on Computational and Scientific Methods, vol. 3, no. 2, 2023.

[37] Y. Son, Y. Kim, and Y. Choi, "Data-Driven Policy Making Framework Utilizing TOWS Analysis," IEEE Access, vol. 13, pp. 93786-93797, 2025, doi: 10.1109/ACCESS.2025.3572908.

[38] N. Chen, S. Sun, Y. Wang, Z. Li, A. Zhu, and Y. Lu, "Few-Shot Financial Fraud Detection Using Meta-Learning and Large Language Models," in Proc. 2025 6th Int. Conf. Computer Science and Management Technology, Dec. 2025, pp. 822-826.

[39] J. A. Long, "Data-Driven Movement Analysis," International Journal of Geographical Information Science, vol. 39, no. 5, 2025.

[40] Y. Xing, Y. Deng, H. Liu, M. Wang, Y. Zi, and X. Sun, "Contrastive Learning-Based Dependency Modeling for Anomaly Detection in Cloud Services," arXiv preprint arXiv:2510.13368, 2025.

[41] J. Jiang, C. Shao, C. Zhang, N. Lyu, and Y. Ni, "Adaptive AI Spatiotemporal Modeling with Dependency Drift Awareness for Anomaly Detection in Large-Scale Clusters," 2025.

[42] H. Zheng, Y. Ma, Y. Wang, G. Liu, Z. Qi and X. Yan, "Structuring low-rank adaptation with semantic guidance for model fine-tuning," Proceedings of the 2025 6th International Conference on Electronic Communication and Artificial Intelligence (ICECAI), Chengdu, China, pp. 731-735, 2025.

[43] H. Zheng, L. Zhu, W. Cui, R. Pan, X. Yan and Y. Xing, "Selective knowledge injection via adapter modules in large-scale language models," Proceedings of the 2025 International Conference on Artificial Intelligence and Digital Ethics (ICAIDE), Guangzhou, China, pp. 373-377, 2025.

[44] X. Yan, J. Du, X. Li, X. Wang, X. Sun, P. Li and H. Zheng, "A Hierarchical Feature Fusion and Dynamic Collaboration Framework for Robust Small Target Detection," IEEE Access, vol. 13, pp. 92953-92964, 2025, doi: 10.1109/ACCESS.2025.3570669.

[45] J. Wei, Y. Liu, X. Huang, X. Zhang, W. Liu and X. Yan, "Self-Supervised Graph Neural Networks for Enhanced Feature Extraction in Heterogeneous Information Networks", 2024 5th International Conference on Machine Learning and Computer Application (ICMLCA), pp. 272-276, 2024.

[46] X. Yan, W. Wang, M. Xiao, Y. Li, and M. Gao, "Survival prediction across diverse cancer types using neural networks", Proceedings of the 2024 7th International Conference on Machine Vision and Applications, pp. 134-138, 2024.

[47] X. Yan, Y. Jiang, W. Liu, D. Yi and J. Wei, "Transforming Multidimensional Time Series into Interpretable Event Sequences for Advanced Data Mining," 2024 5th International Conference on Intelligent Computing and Human-Computer Interaction (ICHCI), pp. 126-130, 2024.

[48] X. Yan, J. Du, L. Wang, Y. Liang, J. Hu and B. Wang, "The Synergistic Role of Deep Learning and Neural Architecture Search in Advancing Artificial Intelligence", Proceedings of the 2024 International Conference on Electronics and Devices, Computational Science (ICEDCS), pp. 452-456, Sep. 2024.

[49] M. Xiao, Y. Li, X. Yan, M. Gao, and W. Wang, "Convolutional neural network classification of cancer cytopathology images: taking breast cancer as an example," Proceedings of the 2024 7th International Conference on Machine Vision and Applications, pp. 145–149, Singapore, Singapore,2024

[50] W. Wang, Y. Li, X. Yan, M. Xiao and M. Gao, "Breast cancer image classification method based on deep transfer learning," Proceedings of the International Conference on Image Processing, Machine Learning and Pattern Recognition, pp. 190-197, 2024.

[51] Y. Li, X. Yan, M. Xiao, W. Wang and F. Zhang, "Investigation of Creating Accessibility Linked Data Based on Publicly Available Accessibility Datasets", Proceedings of the 2023 13th International Conference on Communication and Network Security, pp. 77-81, 2024.

[52] Wang, Y. (2024). AI-Enhanced Distributed Time Series Modeling: Incremental Learning for Evolving Streaming Data. Transactions on Computational and Scientific Methods, 4(8).

[53] Su, C., Jiang, L., & Pi, S. (2021, August). Sentiment classification of chinese railway review text based on multi-feature fusion gated recurrent unit. In 2021 International Conference on Information Science, Parallel and Distributed Systems (ISPDS) (pp. 197-202). IEEE.

[54] Li, S., Wang, Y., Xing, Y., & Wang, M. (2025, December). Mitigating correlation bias in advertising recommendation via causal modeling and consistency-aware learning. In Proceedings of the 2025 6th International Conference on Computer Science and Management Technology (pp. 585-589).

[55] Ke, Z. (2026). Right-Sizing Cloud Data Warehouses Under Asymmetric Loss via One-Sided Conformal Prediction. IEEE Access, 14, 94748-94763.

[56] Xue, Y., Wang, Y., Zhu, A., Sun, X., & Zhang, C. (2026, March). Joint Temporal-Structural Representation Learning for Distributed Fault Discrimination in Microservice Architectures. In 2026 IEEE 8th International Conference on Communications, Information System and Computer Engineering (CISCE) (pp. 1581-1586). IEEE.

[57] S. Sun, Y. Wang, R. Yan, J. Wang and Y. Lu, "CurateDistill: Efficient Data Selection via Active Data Curation and Dataset Distillation," 2026, DOI: 10.13140/RG.2.2.27464.28166.

[58] Mao, L. (2024). Joint Optimization of Dual Retrieval and Adaptive Re-Ranking for Retrieval-Augmented Generation. Artificial Intelligence and Computing Innovations, 4(4).

[59] Huang, S. Y. (2024). Retrieval-Augmented Long-Text Reasoning with Semantic Conflict Detection and Cross-Paragraph Evidence Integration. Artificial Intelligence and Computing Innovations, 4(1).

[60] Xu, Y., Kou, J., Wang, W., Su, Z., Yan, X., & Wang, J. (2026, March). Graph spatiotemporal attention-based collaborative demand forecasting for multi-level supply chains. In Proceedings of the 2026 3rd International Conference on Innovation Management and Information Systems (pp. 214-220).

[61] Su, Z. (2024). Microstructure-Aware Risk Detection in High-Frequency Trading Using Embedding and Sequence Modeling. Transactions on Computational and Scientific Methods, 4(3).

[62] Kou, J., Qing, Z., Wei, R., Xue, F., Huang, W., & Zhuang, H. (2026, April). CausalRCA: Causal Graph-Augmented Retrieval for Root Cause Analysis in Microservice Systems. In 2026 5th International Conference on Electronic Information Engineering, Big Data and Computer Technology (EIBDCT) (pp. 440-444). IEEE.

[63] F. Xue, T. Ying, T. Ding, L. Wang, Z. Qing and P. Hu, "Requirement-Aware LLM Testing for Backend REST APIs," 2026, DOI: 10.13140/RG.2.2.21048.30729.

[64] Hu, P., Z. Qing, T. Ding, T. Ying, F. Xue, and L. Wang, "Energy-Efficient Distributed Inference across the Edge-Cloud Continuum Using Adaptive Model Compression," 2026, DOI: 10.13140/RG.2.2.19323.07209.

[65] Lyu, Y., Jiang, J., & Hu, J. (2026). QueryMind: Workload-Aware LLM-Guided Database Configuration Tuning for Backend Performance.

[66] Hu, J., Lyu, Y., & Jiang, J. (2026). High-Reliability LLM Summarization through Causal Cue Modeling and Factual Correction.

[67] S. Pi, S. Zhang and S. Shawulieti, “CacheRoute: KV-Cache-Aware Routing for LLM Backend Services on Kubernetes,” 2026, DOI: 10.13140/RG.2.2.34191.55206.

[68] Wang, S., B. Su, Y. Liu, S. Pi and A. Zhu, "Safe Adaptive Resilience Configuration for Microservice Backends," 2026, DOI: 10.13140/RG.2.2.30702.57921.

[69] Z. Liu, K. Wu, S. Dang and Y. Yang, "A Retrieval-Augmented Log-Metric-Trace Twin for Backend Failure Diagnosis," 2026, DOI: 10.13140/RG.2.2.34050.64969.

[70] Z. Liu, Y. Yang, C. Chen, S. Dang and K. Wu, "Topology Dependency Modeling and Discriminative Representation Learning Methods for Backend Microservice Call Anomaly Detection Task," 2026, doi: 10.13140/RG.2.2.18843.02082.

[71] Shawulieti, S. (2024). Enhancing Cloud Security with Graph Neural Network-Driven Anomaly Detection in Multi-Tenant Environments. Artificial Intelligence and Computing Innovations, 4(2).

[72] Xue, F., Wu, K., & Abdulla, A. (2025). Topology-Aware Graph Neural Networks for Distributed Node Fault Identification. Transactions on Computational and Scientific Methods, 5(3).

[73] Wang, M., Peng, P. L., & Chen, J. (2024). Financial Report Summarization via Structure-Aware Modeling and Information Fidelity Constraints. Transactions on Computational and Scientific Methods, 4(7).

[74] Peng, Z. (2024). Cross-Period Financial Statement Anomaly Detection via Contrastive Representation Learning. Journal of Computer Technology and Software, 3(8).

[75] Z. Su, J. Chen, X. Meng, N. Sang and Z. Peng, "Relation-Aware Graph Neural Networks with Verifiable Reasoning for Accounting Fraud Detection," ResearchGate Preprint, 2026, DOI: 10.13140/RG.2.2.13631.06565.

[76] Hu, P., Uziel, G., & Ma, X. (2024). Risk-Constrained Reinforcement Learning for SLO-Driven Cloud-Native Elastic Scaling. Journal of Computer Technology and Software, 3(1).

[77] Jiang, Z., Zhang, S., Huang, W., Pi, S., Shawulieti, S., & Xu, S. (2026). Semantic-Aware Intelligent Request Routing in Multi-Model LLM Backend Gateways. Semantic-Aware Intelligent Request Routing in Multi-Model LLM Backend Gateways.

[78] Y. Jiang, Z. Xiao, O. Lin, Z. Gu and Z. Yang, "PrivCollab-Agent: Privacy-Preserving Multi-Agent Collaboration under Partial Information," 2026, DOI: 10.13140/RG.2.2.33663.16800.

[79] Ma, L., Song, T., Long, R., Mao, L., & Gu, Z. (2026, May). Evidence-Gated Memory Writing for Personalized Large Language Model Agents. In 2026 3rd International Conference on Image Processing and Artificial Intelligence (ICIPAI) (pp. 333-338). IEEE.

[80] Gu, Z., & Nie, L. (2025). Knowledge Graph-Enhanced Large Language Models for Opinion Target Sentiment Classification. Journal of Computer Technology and Software, 4(7).

[81] S. Li, B. Chen, Y. Li, Z. Wang, Y. Xue and C. Xu, "Privacy-Preserving Anomaly Detection in Cloud Services Using Hierarchical Federated Learning With Differential Privacy," 2026 9th International Symposium on Big Data and Applied Statistics (ISBDAS), pp. 668-672, 2026.

[82] Lyu, N., Chen, F., Zhang, C., Shao, C., & Jiang, J. (2025, December). Deep temporal convolutional neural networks with attention mechanisms for resource contention classification in cloud computing. In Proceedings of the 2025 5th International Conference on Big Data, Artificial Intelligence and Risk Management (pp. 62-68).

[83] Zhang, C. (2024). Adaptive Multi-Tenant Resource Scheduling in Cloud Computing via Reinforcement Learning. Transactions on Computational and Scientific Methods, 4(3).

[84] Y. Li, W. Zhao, B. Dang, X. Yan, M. Gao, W. Wang, and M. Xiao, "Research on adverse drug reaction prediction model combining knowledge graph embedding and deep learning", Proceedings of the 2024 4th International Conference on Machine Learning and Intelligent Systems Engineering (MLISE), pp. 322-329, June 2024.