ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/390481139

Leveraging the DeepSeek large model: A framework for AI-assisted disaster prevention, mitigation, and emergency response systems

Article *in* Earthquake Research Advances · April 2025

DOI: 10.1016/j.eqrea.2025.100378

CITATION
1

READS
93

6 authors, including:

Chenchen Xie
China University of Geosciences (Beijing)
10 PUBLICATIONS **246** CITATIONS

SEE PROFILE

Huiran Gao
National Institute of Natural Hazards, Ministry of Emergency Management of China
46 PUBLICATIONS **455** CITATIONS

SEE PROFILE

Yuandong Huang
university of chinese acadmic of science
41 PUBLICATIONS **651** CITATIONS

SEE PROFILE

Zhiwen Xue
Institute of Mechanics, Chinese Academy of Sciences
15 PUBLICATIONS **196** CITATIONS

SEE PROFILE

All content following this page was uploaded by Chenchen Xie on 25 May 2025.

The user has requested enhancement of the downloaded file.

ARTICLE IN PRESS

Earthquake Research Advances xxx (xxxx) xxx

Contents lists available at ScienceDirect

Earthquake Research Advances

journal homepage: www.keaipublishing.com/en/journals/earthquake-research-advances

Leveraging the DeepSeek large model: A framework for AI-assisted disaster prevention, mitigation, and emergency response systems

Chenchen Xie [a,b,d], Huiran Gao [a,b], Yuandong Huang [a,b,c], Zhiwen Xue [a,b,c], Chong Xu [a,b,*], Kebin Dai [a,b,c]

[a] Ministry of Emergency Management of China, National Institute of Natural Hazards, Beijing, 100085, China
[b] Key Laboratory of Compound and Chained Natural Hazards Dynamics, Ministry of Emergency Management of China, Beijing, 100085, China
[c] School of Emergency Management Science and Engineering, University of Chinese Academy of Sciences, Beijing, 100049, China
[d] School of Earth Sciences and Resources, China University of Geosciences (Beijing), Beijing, 100083, China

ARTICLE INFO

Keywords:
AI large language models
DeepSeek
System framework research
Natural disaster prevention and control
Emergency assistance

ABSTRACT

We proposes an AI-assisted framework for integrated natural disaster prevention and emergency response, leveraging the DeepSeek large language model (LLM) to advance intelligent decision-making in geohazard management. We systematically analyze the technical pathways for deploying LLMs in disaster scenarios, emphasizing three breakthrough directions: (1) knowledge graph-driven dynamic risk modeling, (2) reinforcement learning-optimized emergency decision systems, and (3) secure local deployment architectures. The DeepSeek model demonstrates unique advantages through its hybrid reasoning mechanism combining semantic analysis with geospatial pattern recognition, enabling cost-effective processing of multi-source data spanning historical disaster records, real-time IoT sensor feeds, and socio-environmental parameters. A modular system architecture is designed to achieve three critical objectives: (a) automated construction of domain-specific knowledge graphs through unsupervised learning of disaster physics relationships, (b) scenario-adaptive resource allocation using risk simulations, and (c) preserving emergency coordination via federated learning across distributed response nodes. The proposed local deployment paradigm addresses critical data security concerns in cross-border disaster management while complying with the FAIR principles (Findable, Accessible, Interoperable, Reusable) for geoscientific data governance. This work establishes a methodological foundation for next-generation AI-earth science convergence in disaster mitigation.

1. Introduction

In the era of rapidly advancing technology, the exploration and breakthroughs in the field of artificial intelligence (AI) are profoundly influencing the development patterns of numerous industries (Xu and Xue, 2024). Among these, the evolution of AI large language model technology is particularly eye-catching. Currently, the global technology of AI large language models has entered a crucial stage of multi-modality fusion. This transformation provides a broader space and stronger capability support for the expansion of AI applications (Huang et al., 2024a). The domestic DeepSeek series has shown remarkable performance among the numerous AI large language models. With its outstanding achievements in the optimization of attention mechanisms and algorithmic innovation, it has demonstrated extraordinary strength in core areas such as semantic understanding and knowledge reasoning. This undoubtedly represents a new step in domestic AI large language model technology (Gao et al., 2025). Its training cost has been significantly reduced compared to similar international models by several times or even more. This provides a more cost-effective choice for the application of various industries (Guo et al., 2024; Kotsis, 2025; Peng et al., 2025). However, in the important field of natural disaster prevention and control, which concerns the safety of people's lives and property, there are still pain points to be solved. Firstly, traditional natural disaster prevention and control rely on manual experience analysis. Based on the accumulated knowledge, skills, observations, and research over a long period, judgments and predictive assessments are made on current or future situations. This method is relatively slow. In the face of massive information, such as meteorological observation data and seismic-geological monitoring data (Huang et al., 2023; Xu et al., 2014a), it

* Corresponding author. Ministry of Emergency Management of China, National Institute of Natural Hazards, Beijing 100085, China.
E-mail address: chongxu@ninhm.ac.cn (C. Xu).

Peer review under the responsibility of Editorial Board of Earthquake Research Advances.

https://doi.org/10.1016/j.eqrea.2025.100378
Received 1 March 2025; Received in revised form 30 March 2025; Accepted 1 April 2025
2772-4670/© 2025 China Earthquake Networks Center. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Please cite this article as: Xie, C. et al., Leveraging the DeepSeek large model: A framework for AI-assisted disaster prevention, mitigation, and emergency response systems, Earthquake Research Advances, https://doi.org/10.1016/j.eqrea.2025.100378

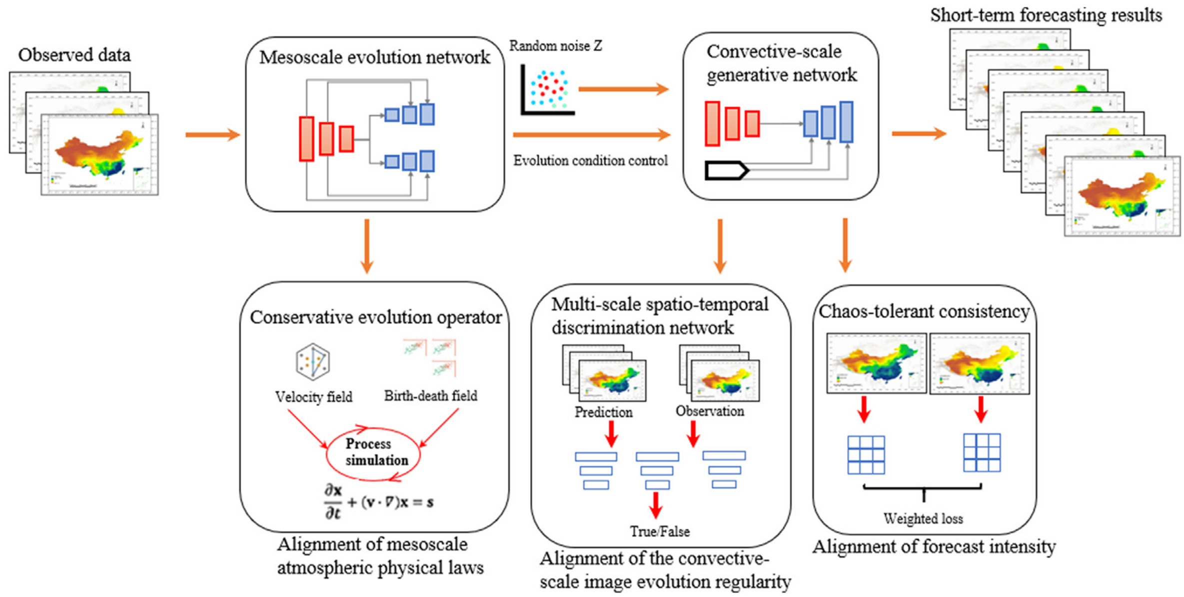

Fig. 1. Schematic diagram of the "Feng Lei" large model.

takes much time and manpower to integrate and analyze the data, making it hard to make real-time and comprehensive decisions on events. Also, due to different understandings of professional knowledge, different disciplines may have different evaluation criteria, leading to inconsistent analysis and decision-making results (Shirzaei et al., 2025). Extreme events can trigger multiple natural disasters, causing huge losses (Xie et al., 2025b; Xu et al., 2014b). As research on natural disasters deepens and their complexity is better understood, it's crucial to fully and accurately grasp the occurrence mechanisms, development trends, and impact scopes of disasters (Shao et al., 2024; Xiao et al., 2023). We need to build connections between different disciplines and use basic data from various disaster events for comprehensive analysis and evaluation of coupled disaster impacts. Moreover, promptly issuing precise and effective dynamic emergency plans in critical moments to deal with disasters and minimize losses is now an urgent issue to solve (Wu et al., 2025).

In view of this, we focus on closely combining the advanced technology of AI large language models with the actual needs of the natural disaster prevention and control field. We propose a research framework for an AI-assisted system for natural disaster prevention and control based on DeepSeek. The aim is to explore an integrated, intelligent, professional, and efficient new path for natural disaster prevention and control. It is expected to contribute to the protection of people's lives and property safety and provide beneficial references and examples for technological innovation and development in the field of natural disaster prevention and control.

2. The application level of AI large language models in natural disaster prevention and control

2.1. Existing application platforms

Based on their application domains, AI large models are categorized into general-purpose and industry-specific large models. In the field of natural disaster prevention and mitigation, a limited number of industry-specific large models have emerged (Bi et al., 2023; Chen et al., 2023; Han et al., 2024; Huang et al., 2024b). For instance, on June 18, 2024, the China Meteorological Administration released three AI meteorological large model systems. Among them, "Fengqing" enhances the timeliness and accuracy of medium-to short-term forecasts; "Fengshun" optimizes sub-seasonal to seasonal predictions and improves precipitation forecasting skills; and "Fenglei" enhances radar echo forecasting capabilities to achieve short-term forecasting. As shown in Fig. 1, observational data are processed through a mesoscale evolution network and a convective-scale generative network, combined with random noise to generate short-term forecasting results. These results are calibrated and optimized physically using conservative evolution operators, multi-scale spatiotemporal discriminative networks, and chaos-tolerant consistency modules to ensure forecasting accuracy and consistency. On August 7, 2024, the Ministry of Emergency Management officially released the "Jiu An" large model for the emergency management sector (News, 2024). The "Jiu An" large model

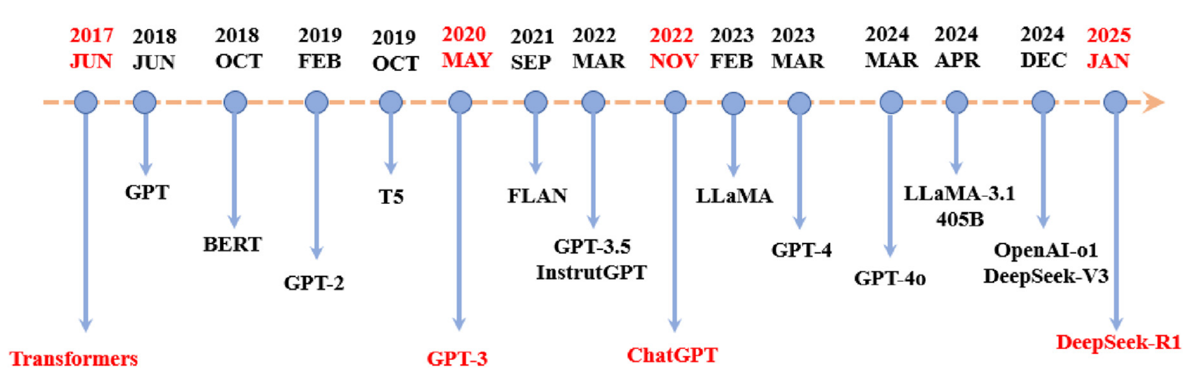

Fig. 2. Developmental history of large models(Xu et al., 2024).

drives industrial intelligence through five core innovations. It establishes China's first privately deployed industry-specific AI foundation, integrating domestic hardware and software via supervised fine-tuning, OP-RAG optimization, and multimodal routing for customized applications. The system synthesizes over 180 PB of cross-modal data and emergency domain corpora to build an expert-level knowledge framework, enabling multimodal query support spanning text, imagery, and video. By leveraging the model's visual comprehension capabilities, it innovates safety supervision in high-risk industries, rapidly identifying hazards through image/video analysis to enhance regulatory efficiency. The model further strengthens disaster preparedness by integrating communication tower metrics, meteorological data, and geospatial inputs for real-time hazard assessment and rescue strategy optimization. Its "Smart Emergency Brain" enables sub-second incident response plan generation, 30-second cross-departmental coordination, and one-click orchestration of 40+ operational systems, significantly improving rescue precision and emergency chain efficiency. These advancements collectively advance intelligent emergency management through tailored AI infrastructure, domain-specific knowledge fusion, and rapid decision-making workflows.

2.2. Existing technological pathways

The current technological approaches of AI large language models in the natural disaster field center on disaster simulation, multimodal integration, and intelligent decision-making (Fig. 2). They analyze satellite remote-sensing data, ground-sensor information, and historical disaster cases, employing deep learning and machine learning algorithms to process massive amounts of meteorological, geological, and historical disaster data from diverse and heterogeneous sources. This leads to the development of high-precision prediction models for forecasting typhoon paths, flood peaks, or seismically active areas (Adnan et al., 2025; Xiao et al., 2024). Based on pre-trained models, Domain-adaptive Fine-tuning is used to enhance disaster-scene adaptability (Gema et al., 2023). For instance, the "Jiu An" AI large language model uses a multi-task transfer learning framework with specialized datasets for floods, wildfires, etc. It dynamically adjusts loss function weights (such as using Focal Loss to ease class imbalance), enabling real-time identification of abnormal water accumulation and dust explosions in high-risk areas, and increasing early-warning accuracy by 18 %. Considering the unique nature of disaster tasks, multi-modal structured prompt templates are designed (e.g., "Based on [satellite images] and [historical precipitation data], predict flood areas in the next 24 hr"). Through reinforcement learning feedback mechanisms, prompt words are iteratively optimized to reduce model misjudgments from ambiguous instructions (Liao et al., 2024; Liu et al., 2023b). Integrating multi-modal data such as satellite images, weather radar, social-media text, knowledge graphs, and causal reasoning models helps assess disaster spread and consequences. Contrastive Learning aligns text descriptions with image features (Wang and Qi, 2022) (for example, linking the keyword "road collapse" with crack regions in satellite images), improving comprehensive disaster identification (Xue et al., 2023). In real-time disaster analysis, the Chain-of-Thought prompt guides the model to process information step-by-step (Wei et al., 2022) (e.g., "Step 1: Extract location keywords from social-media text; Step 2: Match with strong rainfall regions in weather radar; Step 3: Evaluate landslide probability"), significantly reducing errors in multi-source data fusion. Disaster impact and economic losses are quantified by leveraging reinforcement learning and edge computing.

2.3. Designing decision-making agents

Multimodal models, model fine-tuning, and prompt optimization are three key technologies for enhancing AI performance. Multimodal models (e.g., GPT-4, CLIP), by integrating data from multiple modalities such as text, images, and speech, endow AI with cross-modal understanding and generation capabilities, enabling it to handle more complex scenarios. Based on pre-trained models, model fine-tuning involves targeted training using domain-specific data (e.g., medical, legal) to optimize task adaptability. Prompt optimization enhances answer accuracy by introducing external knowledge bases through RAG (Retrieval-Augmented Generation) or guides model logical reasoning step-by-step using CoT (Chain of Thought), significantly improving the efficiency of solving complex problems. These three technologies complement each other and collectively drive AI accuracy, generalization, and interactivity breakthroughs.

Earthquakes are one of the most frequent natural disasters (Xu et al., 2025), and it is crucial to construct earthquake decision-making agents. Building such agents requires a deep integration of multimodal data, domain knowledge bases, and adaptive AI technologies (Naghshvarianjahromi et al., 2023). Real-time collection of seismic waveforms, surface deformation, and disaster-related text is achieved through seismometers, satellite remote sensing, and social media. A structured knowledge base is constructed with historical earthquake catalogs and geological models (Hou et al., 2024). Multimodal models are used to align textual, imaging, and temporal data features, and the capabilities for magnitude prediction and fault classification are fine-tuned using seismology-labeled data. RAG is employed to dynamically retrieve emergency plans and historical cases, and CoT is used to guide step-by-step reasoning (risk assessment-aftershock prediction-resource scheduling) to generate decision chains while incorporating uncertainty quantification and dynamic prompt adjustments. Ultimately, an edge-cloud collaborative architecture enables second-level early warnings and precise emergency plan outputs (Wu, 2024), enhancing earthquake response speed and decision-making scientificity through a multi-technology closed loop (Ghaffarian et al., 2025). Integrating existing DeepSeek large model technology can further optimize this agent. Its multimodal architecture deeply integrates heterogeneous data such as seismic waveforms, satellite images, and disaster text, and combines RAG technology to real-time access seismology knowledge bases and historical cases, addressing the "hallucination" issue of general-purpose large models. It outputs interpretable decision steps and confidence intervals by simulating expert logic step-by-step with CoT (e.g., risk assessment-aftershock prediction-resource scheduling). Simultaneously, the open-source and low-cost characteristics enable private deployment and dynamic knowledge updates, supporting a collaborative architecture between edge-device sub-second warnings and cloud-based complex analysis. This approach demonstrates significantly improved computational efficiency and domain-specific adaptability compared to traditional models, offering an AI solution for earthquake response that combines real-time responsiveness, accuracy, and scientific rigor.

To demonstrate the practicality of the above-mentioned earthquake decision-making agent, we posit its application in the Kumamoto Earthquake risk assessment. By integrating real-time seismograph waveforms, satellite deformation data, and historical earthquake catalogs, and leveraging Retrieval-Augmented Generation (RAG) technology to accurately retrieve the precursor patterns of the Hanshin Earthquake and Coulomb stress models, we drive a multimodal model to achieve three core decision-making steps. First, based on the characteristics of increased minor seismic frequency and fault stress accumulation reaching a critical value, we calculate the probability of seismic intensity exceeding 6.5 in the future year. Second, by combining population density, we predict the impact range of intensity exceeding 6 degrees and identify the number of people in Kumamoto City and surrounding areas who need evacuation. Finally, by dynamically retrieving post-disaster traffic data, we generate optimized evacuation routes that avoid high-risk bridges. This process simultaneously outputs confidence annotations, supporting expert review and real-time knowledge base updates, and through edge-end lightweight models, it enables minute-level warning alerts, showcasing the core advantages of multimodal correlation analysis, explainable reasoning chains, and low-cost deployment.

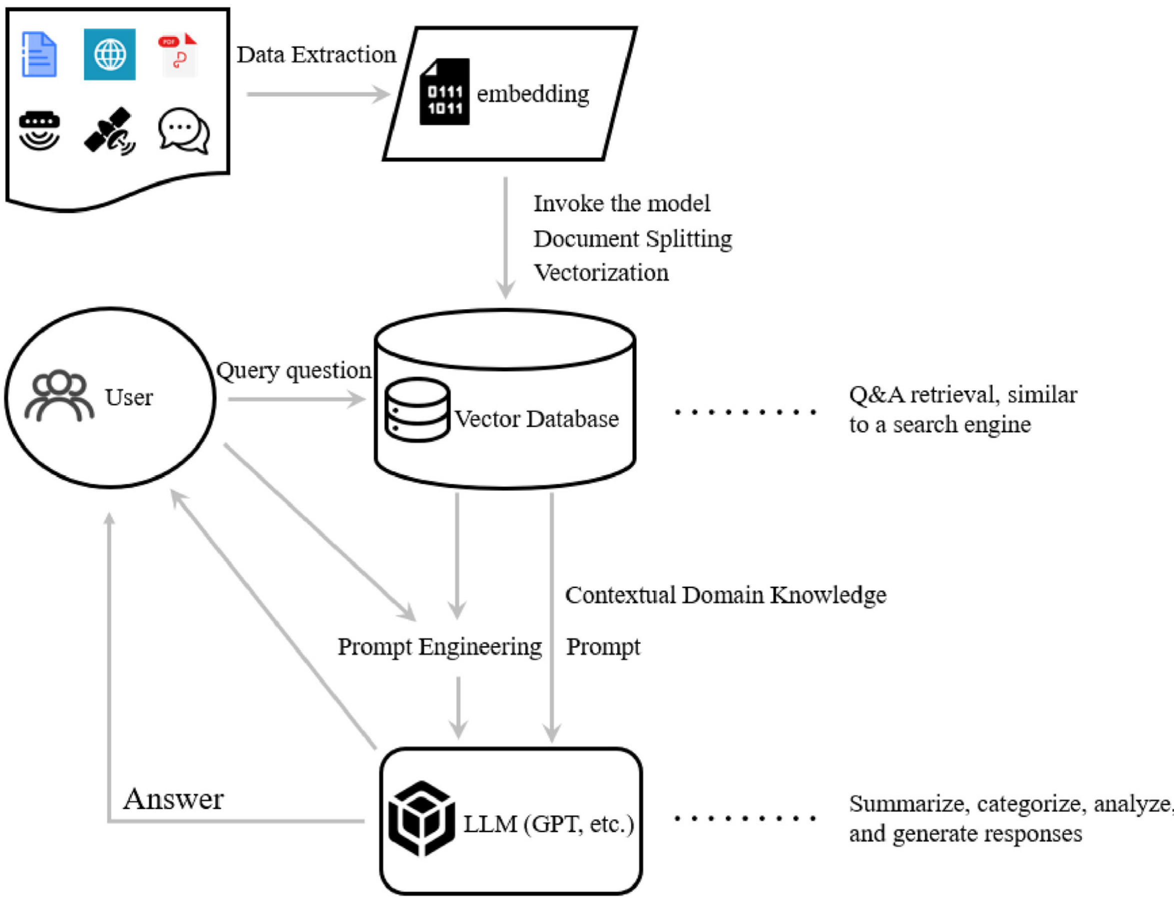

Fig. 3. Technology roadmap of large language models.

This compresses the traditional emergency response time from several hours to the minute level.

3. The development trend of AI large language models and the technological advantages of DeepSeek

3.1. Definition of large AI models and feasibility of DeepSeek

Large language AI models are those with massive parameters (usually billions to tens of billions) and complex structures (Xu, 2023). Relying on deep learning tech (like Transformer), they can process multimodal data including natural language, images, and audio. They're adapted to diverse tasks through pre-training and fine-tuning (Li et al., 2024; Tang et al., 2024; Yao et al., 2024). Both domestic and foreign large-language AI models now perform excellently. Take the LLaMA series from Meta for example. Based on the Transformer architecture, its parameter scale ranges from 7 billion to 70 billion. It's stable in general-purpose language tasks but has low inference efficiency and relies heavily on large-scale computing power (Inan et al., 2023; Yeom et al., 2024). The GPT series by OpenAI has parameters reaching the trillion level. It's known for production capacity and multimodal integration, leading in text generation and complex reasoning tasks. However, its training cost is extremely high, over $100 million for GPT-4o (Achiam et al., 2023; Baktash and Dawodi, 2023). Domestically, the Qwen series also has advanced performance. With 18 trillion tokens of pre-training data, it's good at multilingual tasks and complex logical reasoning. Yet, it needs improvement in resource consumption, understanding of complex instructions, structured output constraints, and training data bias (Bai et al., 2023; Yang et al., 2024). The Kimi series has technical highlights, supporting 128k long-context processing and optimizing multimodal task performance. However, it has shortcomings like dependence on computing power, average data processing ability, no multilingual support, and limited reasoning ability (Team et al., 2025). The DeepSeek series adopts a Mixture-of-Experts (MoE) architecture, dynamically activating sub-models (e.g., DeepSeek-V3 with 37 billion activated parameters out of a total of 67.1 billion). This reduces computing resource needs. It also uses Multi-Head Latent Attention (MLA) to optimize memory usage and long-sequence processing. FP8 mixed-precision training cuts training costs and boosts efficiency. But DeepSeek still faces major limitations and challenges like data purity, case dependency, optimization bottlenecks, and international competitiveness (Bi et al., 2024; Liu et al., 2024a, 2024b; Lu et al., 2024). Globally, models like GPT and LLaMA enhance performance by expanding parameter scales and have made significant progress in multimodal integration. Chinese models, however, show unique advantages in efficient training and reasoning optimization. DeepSeek performs well, especially in math and coding problems, reaching a leading level (Gao et al., 2025). Kimi stands out in multimodal and long-context tasks. Qwen has strong capabilities in multilingual and complex reasoning tasks. In particular, DeepSeek and Kimi offer higher cost-effectiveness for practical applications.

3.2. Technological evolution and development trends

The introduction and development of generative large models, which are constructed based on deep learning technology and feature a vast number of parameters and complex structures (Brown et al., 2020), have garnered significant attention. These models generate content such as text, images, and audio videos through learning probability distributions. The core evolution of generative large models has shifted from deepening multi-modal capabilities (such as Sora's dynamic video generation and

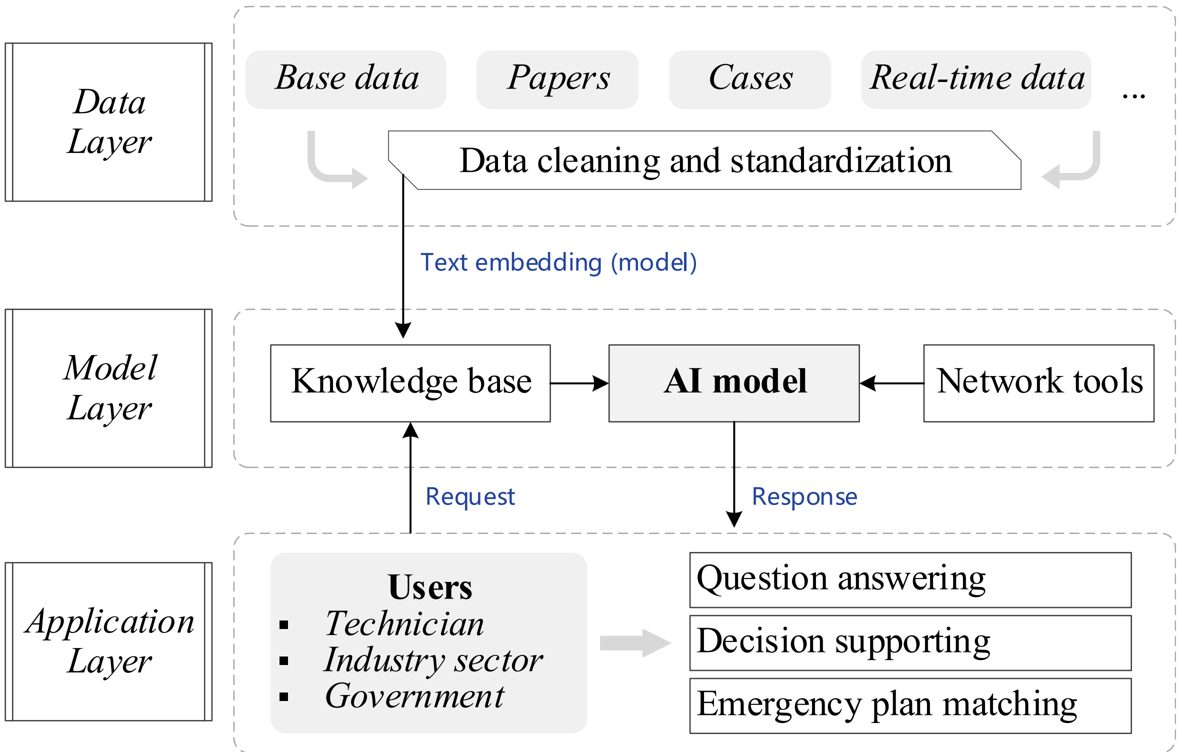

Fig. 4. Schematic diagram of the architecture for AI-based intelligent auxiliary system.

Suno's complete song creation) to accelerating open-source accessibility and cost reduction (Fig. 2). Despite mainstream models lowering technological barriers through open-source ecosystems (e.g., DeepSeek R1 and Alibaba's Qwen), and reducing API costs to 3 % of OpenAI's via techniques like model distillation and FP8 precision training, technological iteration shows no signs of slowing down. On one hand, vertical domain applications and long text processing capabilities continue to be optimized. On the other hand, DeepSeek has enhanced its reasoning abilities through a "deep thinking" mode, pioneered the MLATM mechanism to reduce memory usage by 30 %, and is building a global developer ecosystem with a fully open-source strategy (Liu et al., 2023a).

3.3. Key technological breakthroughs

First, based on the self-attention mechanism, long sequences are computed and processed in parallel. Next, the model automatically learns from massive amounts of data via masked language models and other methods. Then, both data-parallel and model-parallel computing speed up the training process. Furthermore, the efficiency of the model is improved by means of quantization, distillation, etc. As shown in Fig. 3, the process involves extracting data from multiple sources and performing vectorization, storing the data in a vector database. User queries retrieve relevant information from the vector database, and prompt engineering optimizes the input to a large language model (LLM) to generate answers, thereby enabling the functions of summarization, classification, analysis, and response generation. The technological breakthroughs of AI large models in natural disaster prevention and control revolve around the "perception-decision-universal benefit" chain. First, multi-modal data fusion integrates satellite remote sensing, ground sensors, and social media (such as dynamic flood range calculation and disaster-affected demand mining). Then, a real-time disaster perception network is constructed, combining DeepSeek's FP8 precision optimization to achieve second-level early warnings on edge devices (such as landslide prediction). Second, reinforcement learning-driven dynamic decision-making simulates disaster chain reactions (earthquake-secondary disaster-traffic disruption), generates multi-level emergency plans, and optimizes resource scheduling (such as real-time adjustment of typhoon path avoidance strategies). Open-source ecosystems and low-cost technologies, represented by DeepSeek-R1 and Alibaba's Qwen, reduce training costs to 1/70 of OpenAI's through model distillation, enabling the deployment of edge AI devices in remote areas (such as thermal imaging fire situation analysis by drones in offline environments). Synthetic data and architectural innovation break through the bottleneck of rare disaster data, with the MLATM mechanism reducing memory usage by 30 %, supporting the local operation of TB-level models for high-precision 3D flood simulation. Finally, knowledge graphs and generative AI work together to output post-disaster reconstruction plans (such as planning resettlement sites and prioritizing infrastructure repair).

4. Design and implementation of an intelligent system for natural disaster prevention and control based on AI large language models

4.1. Requirement analysis

Existing web-based or enterprise-level AI models have certain limitations in application. High concurrency can lead to server overload and reduced response speed, affecting the user experience. While commercial APIs can provide higher performance, their high cost increases the burden on users. Additionally, privacy and security issues during data transmission cannot be ignored. Given these considerations, the solution of deploying AI large models locally, supported by open-source frameworks such as the DeepSeek model, has gained favor with an increasing number of domestic institutions. Based on the localization of AI models, integrating the vast professional literature library of the natural disaster industry and developing intelligent auxiliary systems based on the API interfaces of local AI models is a feasible solution to ensure data security, get rid of external dependencies, and meet the new demands of natural disaster prevention, control, and emergency work.

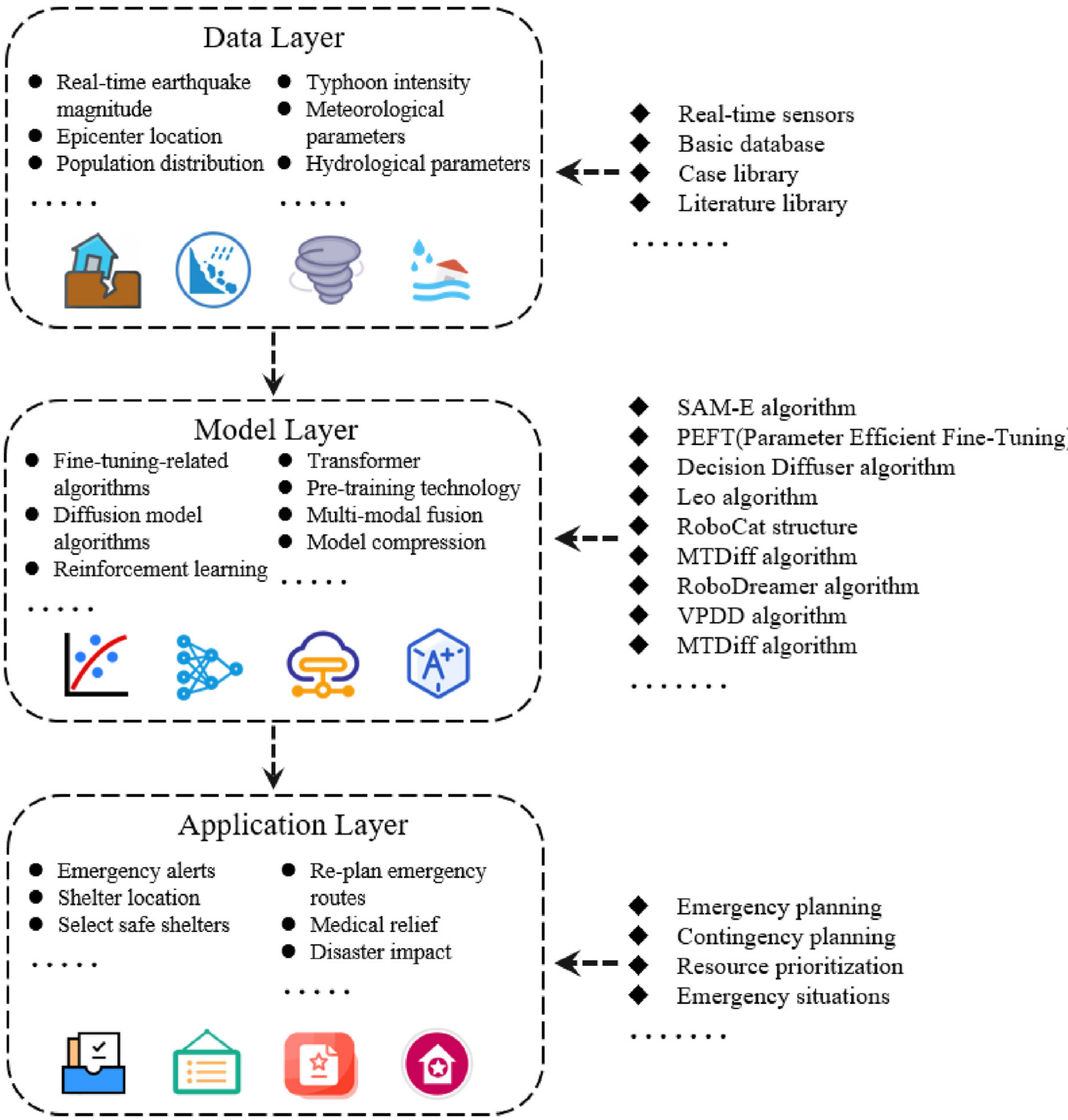

Fig. 5. Process of the data-model-application framework in natural disaster prevention and control.

4.2. System architecture and technical implementation

The AI intelligent auxiliary system follows the traditional Model-View-Controller (MVC) pattern (Pop and Altar, 2014). In terms of technical structure, it can adopt a three-layer architecture design of “data layer algorithm layer application layer”. It deeply integrates multi source heterogeneous data and domain specific knowledge to achieve the full process application and management of natural disaster knowledge graph construction, intelligent decision making, and emergency response (Fig. 4).

The data layer supports the construction of a multi modal natural disaster knowledge base, including basic databases of natural disasters, case libraries, literature libraries, and necessary real time data streams, etc. It cleanses and standardizes the structured and unstructured text, information, and data in multi modal data. By using technologies such as text embedding models (e.g., the nomic-embed-text model), it builds a natural disaster knowledge base to achieve the goal of knowledge distillation. The algorithm layer is mainly based on existing open-source AI large models. It integrates domain specific knowledge bases of natural disasters for vertical fine tuning, enhancing the AI models' understanding of complex scenarios and professional backgrounds. The application layer integrates and presents the intelligent and auxiliary functions of the system through user views. For example, it can quickly query, organize, and summarize relevant knowledge through natural language interaction, providing decision making support for natural disaster prevention and emergency response in complex environments. It can also automatically match emergency plans and dynamically adjust resource allocation plans. Through characteristics such as knowledge driven, rapid response, and dynamic adaptation, the intelligent auxiliary system based on AI large models can achieve full chain intelligent auxiliary services for natural disaster prevention and emergency response. As shown in Fig. 5, the data layer integrates real-time sensor, basic database, case-library, and literature-library data, covering seismic, typhoon-related data, and population distribution (Xie et al., 2025a). The model layer applies algorithms and technologies like fine-tuning, diffusion models, and reinforcement learning, combined with methods such as Transformer, pre-training, and multi-modal fusion, using specific algorithms like SAM-E, Leo, and RoboCat. The application layer fulfills functions like

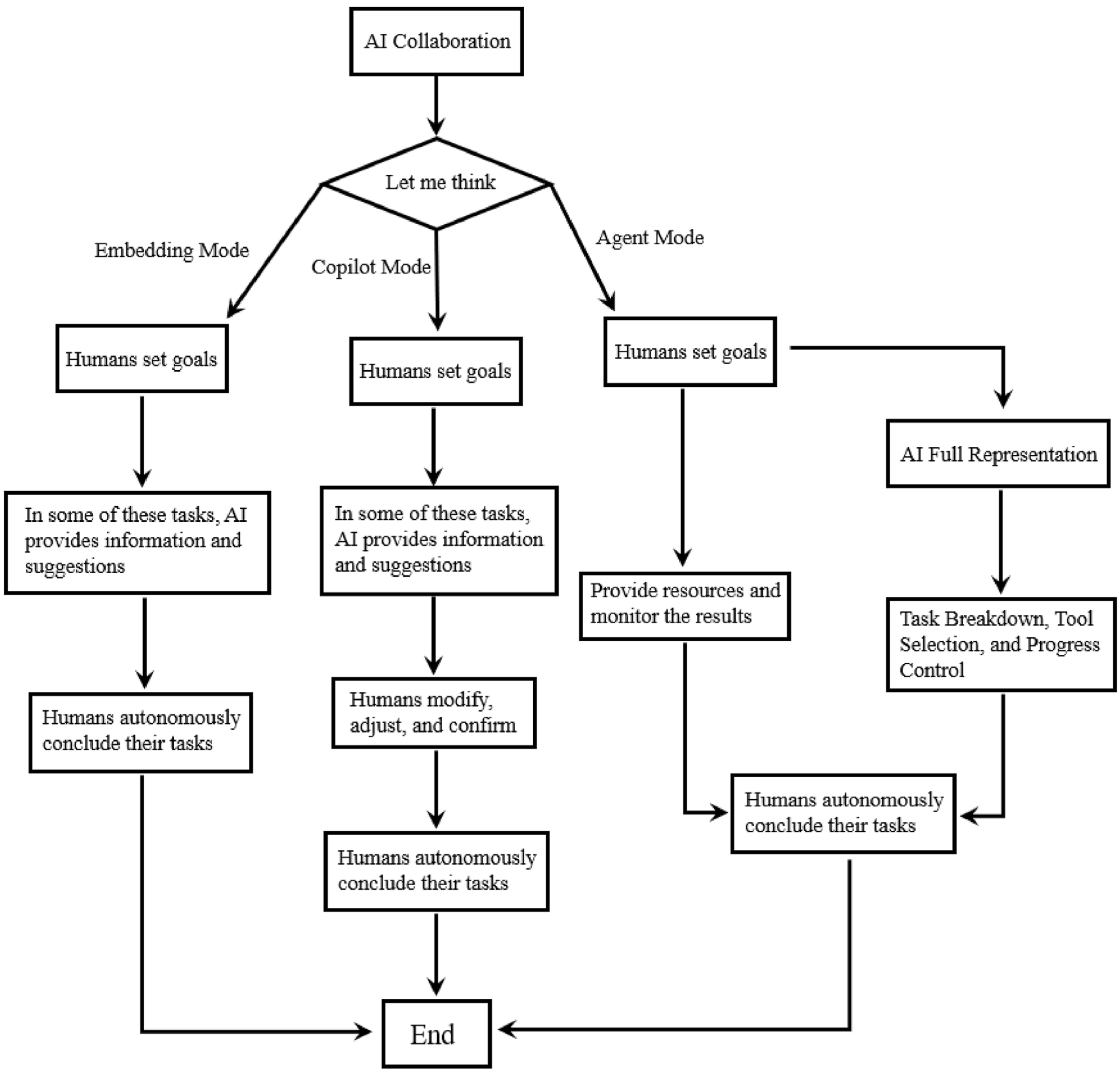

Fig. 6. Schematic representation of the envisioned synergy between AI and humans(Zhao et al., 2023).

emergency alerting, replanning escape routes, selecting shelters, medical rescue, and disaster-impact assessment. It is associated with concepts like emergency planning, crisis planning, resource prioritization, and contingency and supports disaster decision-making and emergency planning.

5. Conclusions and prospects

5.1. Conclusions

The research framework of the natural disaster prevention and emergency AI auxiliary system integrating the domestic DeepSeek large model proposed in this paper provides new ideas and methods for enhancing the intelligent level of natural disaster prevention. By deeply analyzing the current status of application platforms and technical paths of AI large models in natural disaster prevention, and summarizing the functions and technical paths of existing AI large models for natural disaster prevention, the development directions of key technologies such as disaster simulation, multi-modal data fusion, and intelligent decision making are clarified. Meanwhile, it elaborates on the definition and evolution of large AI models, examines the feasibility and advantages of the DeepSeek large model, and relates it to existing disaster prevention measures for integrative conceptualization. Then, solid technical support is provided for the system design and implementation by combining the technical advantages of the DeepSeek large model, such as reinforcement learning-driven reasoning model, architectural innovation, open-source deployment, and low cost. In terms of system design and implementation, a three-layer architecture of "data layer-algorithm layer-application layer" is adopted to achieve the full process application and management of natural disaster knowledge graph construction, intelligent decision making, and emergency response. This architecture can not only effectively integrate multi source heterogeneous data but also ensure data security through local deployment of AI large models, adapting to the new demands of natural disaster prevention and emergency work.

5.2. Prospects

AI large models are gaining traction across fields, yet their development faces multifaceted challenges. Balancing model and data scale is critical, with researchers exploring data augmentation, transfer learning, and model compression to reduce costs and enhance efficiency (Che et al., 2023). Innovations in network architecture are also key; the Transformer's limitations have spurred designs like the state-space-based Mamba model to boost computational efficiency and generalization (Chen et al., 2024). Prompt engineering, an emerging paradigm, improves model performance via specific prompts, though designing robust prompts remains an open research question. The emergence of abilities like contextual reasoning suggests models may internalize human-like cognitive mechanisms, yet their nature and controllability require further exploration. Knowledge updating, interpretability, privacy and security, and issues of data bias and misleading information are also significant challenges (Luo et al., 2023; Sun, 2024; Wang et al., 2023). Addressing these is vital for technological advancement, expanded applications, and broader societal impact. As shown in Fig. 6, future AI-human collaboration is envisioned in three modes: embedding, copilot, and agent (Zhao et al., 2023). In embedding mode, humans set goals, AI offers information and suggestions, and humans independently complete tasks. In copilot mode, after receiving AI's input, humans modify, adjust, and confirm tasks before independent completion. In agent mode, humans set goals, and AI fully participates in handling task decomposition, tool selection, progress control, resource provision, and result monitoring, while humans retain final task completion. These modes reflect varying degrees of AI involvement, from auxiliary support to full task execution.

CRediT authorship contribution statement

Chenchen Xie: Writing – original draft. **Huiran Gao:** Visualization, Formal analysis. **Yuandong Huang:** Formal analysis. **Zhiwen Xue:** Investigation, Formal analysis. **Chong Xu:** Writing – review & editing, Investigation, Formal analysis, Conceptualization. **Kebin Dai:** Formal analysis.

Author agreement and acknowledgment

All the authors who contributed to the study have approved the final version. We thank the anonymous reviewer and the editor, whose constructive and helpful comments improved this manuscript.This research was funded by the Chongqing Water Resources Bureau, China (Project No. CQS24C00836).

Declaration of competing interest

The authors declared that they have no conflicts of interest in this work. Professor Chong Xu is the deputy EIC of Earthquake Research Advances and is not involved in the peer review process.

References

Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F.L., Almeida, D., Altenschmidt, J., Altman, S., Anadkat, S., 2023. Gpt-4 technical report. arXiv preprint arXiv:2303.08774. https://doi.org/10.48550/arXiv.2303.08774.

Adnan, B., Miryala, S., Sambu, A., Vaidhyanathan, K., De Sanctis, M., Spalazzese, R., 2025. Leveraging LLMs for dynamic IoT systems generation through mixed-initiative interaction. arXiv preprint arXiv:2502.00689. https://doi.org/10.48550/arXiv.2502.00689.

Bai, J., Bai, S., Chu, Y., Cui, Z., Dang, K., Deng, X., Fan, Y., Ge, W., Han, Y., Huang, F., 2023. Qwen technical report. arXiv preprint arXiv:2309.16609. https://doi.org/10.48550/arXiv.2309.16609.

Baktash, J.A., Dawodi, M., 2023. Gpt-4: a review on advancements and opportunities in natural language processing. arXiv preprint arXiv:2305.03195. https://doi.org/10.48550/arXiv.2305.03195.

Bi, K., Xie, L., Zhang, H., Chen, X., Gu, X., Tian, Q., 2023. Accurate medium-range global weather forecasting with 3D neural networks. Nature 619 (7970), 533–538.

Bi, X., Chen, D., Chen, G., Chen, S., Dai, D., Deng, C., Ding, H., Dong, K., Du, Q., Fu, Z., 2024. Deepseek llm: scaling open-source language models with longtermism. arXiv preprint arXiv:2401.02954. https://doi.org/10.48550/arXiv.2401.02954.

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J.D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., 2020. Language models are few-shot learners. Adv. Neural Inf. Process. Syst. 33, 1877–1901.

Che, W., Dou, Z., Feng, Y., Gui, T., Han, X., Hu, B., Huang, M., Huang, X., Liu, K., Liu, T., Liu, Z., Qin, B., Qiu, X., Wan, X., Wang, Y., Wen, J., Yan, R., Zhang, J., Zhang, M., Zhang, Q., Zhao, J., Zhao, X., Zhao, Y., 2023. Towards a comprehensive understanding of the impact of large language models on natural language processing: challenges, opportunities and future directions. Sci. Sin. 53 (9), 1645–1687. https://doi.org/10.1360/SSI-2023-0113.

Chen, H., Liu, Z., Sun, M., 2024. The social opportunities and challenges in the era of large language models. J. Comput. Res. Dev. 61 (5), 1094–1103. https://doi.org/10.7544/issn1000-1239.202330700.

Chen, K., Han, T., Gong, J., Bai, L., Ling, F., Luo, J.-J., Chen, X., Ma, L., Zhang, T., Su, R., 2023. Fengwu: pushing the skillful global medium-range weather forecast beyond 10 days lead. arXiv preprint arXiv:2304.02948. https://doi.org/10.48550/arXiv.2304.02948.

Gao, T., Jin, J., Ke, Z.T., Moryoussef, G., 2025. A comparison of DeepSeek and other LLMs. arXiv preprint arXiv:2502.03688. https://doi.org/10.48550/arXiv.2502.03688.

Gema, A.P., Minervini, P., Daines, L., Hope, T., Alex, B., 2023. Parameter-efficient fine-tuning of llama for the clinical domain. arXiv preprint arXiv:2307.03042. https://doi.org/10.48550/arXiv.2307.03042.

Ghaffarian, S., Shafapourtehrany, M., Lagap, U., Batur, M., Özener, H., Kılcı, R.E., Karaman, H., 2025. Earthquake-based multi-hazard resilience assessment: a case study of Istanbul, Turkey (neighborhood level). npj Natural Hazards 2 (1), 15. https://doi.org/10.1038/s44304-025-00065-8.

Guo, D., Zhu, Q., Yang, D., Xie, Z., Dong, K., Zhang, W., Chen, G., Bi, X., Wu, Y., Li, Y., 2024. DeepSeek-Coder: when the large language model meets programming — the rise of code intelligence. arXiv preprint arXiv:2401.14196. https://doi.org/10.48550/arXiv.2401.14196.

Han, T., Guo, S., Ling, F., Chen, K., Gong, J., Luo, J., Gu, J., Dai, K., Ouyang, W., Bai, L., 2024. Fengwu-ghr: learning the kilometer-scale medium-range global weather forecasting. arXiv preprint arXiv:2402.00059. https://doi.org/10.48550/arXiv.2402.00059.

Hou, B., Dai, H., Song, J., Li, S., 2024. P-wave arrival picking using Chinese strong-motion acceleration records based on PhaseNet. World Earthq. Eng. 40 (4), 131–141. https://doi.org/10.19994/j.cnki.WEE.2024.0073.

Huang, D., Yan, C., Li, Q., Peng, X., 2024a. From large language models to large multimodal models: a literature review. Appl. Sci. 14 (12), 5068. https://doi.org/10.3390/app14125068.

Huang, X., Lin, Y., Xiong, W., Li, J., Pan, J., Zhou, Y., 2024b. Research on international development of AI large meteorological models in numerical forecasting. Transactions of Atmos. Sci. 47 (1), 46–54. https://doi.org/10.13878/j.cnki.dqkxxb.20231201001.

Huang, Y., Xie, C., Li, T., Xu, C., He, X., Shao, X., Xu, X., Zhan, T., Chen, Z., 2023. An open-accessed inventory of landslides triggered by the MS 6.8 Luding earthquake, China on September 5, 2022. Earthq Res Advances 3 (1), 100181. https://doi.org/10.1016/j.eqrea.2022.100181.

Inan, H., Upasani, K., Chi, J., Rungta, R., Iyer, K., Mao, Y., Tontchev, M., Hu, Q., Fuller, B., Testuggine, D., 2023. Llama guard: llm-based input-output safeguard for human-ai conversations. arXiv preprint arXiv:2312.06674. https://doi.org/10.48550/arXiv.2312.06674.

Kotsis, K.T., 2025. ChatGPT and DeepSeek evaluate one another for science education. EIKI Journal of Effective Teaching Methods 3 (1). https://doi.org/10.59652/jetm.v3i1.439.

Li, D., Jiang, B., Huang, L., Beigi, A., Zhao, C., Tan, Z., Bhattacharjee, A., Jiang, Y., Chen, C., Wu, T., 2024. From generation to judgment: opportunities and challenges of llm-as-a-judge. arXiv preprint arXiv:2411.16594. https://doi.org/10.48550/arXiv.2411.16594.

Liao, H., Shen, H., Li, Z., Wang, C., Li, G., Bie, Y., Xu, C., 2024. Gpt-4 enhanced multimodal grounding for autonomous driving: leveraging cross-modal attention with large language models. Commun Transp Res 4, 100116. https://doi.org/10.1016/j.commtr.2023.100116.

Liu, A., Feng, B., Wang, B., Wang, B., Liu, B., Zhao, C., Dengr, C., Ruan, C., Dai, D., Guo, D., 2024a. Deepseek-v2: a strong, economical, and efficient mixture-of-experts language model. arXiv preprint arXiv:2405.04434. https://doi.org/10.48550/arXiv.2405.04434.

Liu, A., Feng, B., Xue, B., Wang, B., Wu, B., Lu, C., Zhao, C., Deng, C., Zhang, C., Ruan, C., 2024b. Deepseek-v3 technical report. arXiv preprint arXiv:2412.19437. https://doi.org/10.48550/arXiv.2412.19437.

Liu, X., Hu, B., Chen, K., Zhang, M., 2023a. Key technologies and future development direction of large language models: insights form ChatGPT. Bull. Natl. Sci. Found. China 37 (5), 758–766. https://doi.org/10.16262/j.cnki.1000-8217.20231026.004.

Liu, Y., Han, T., Ma, S., Zhang, J., Yang, Y., Tian, J., He, H., Li, A., He, M., Liu, Z., 2023b. Summary of chatgpt-related research and perspective towards the future of large language models. Meta-radiology 1 (2), 100017. https://doi.org/10.1016/j.metrad.2023.100017.

Lu, H., Liu, W., Zhang, B., Wang, B., Dong, K., Liu, B., Sun, J., Ren, T., Li, Z., Yang, H., 2024. Deepseek-vl: towards real-world vision-language understanding. arXiv preprint arXiv:2403.05525. https://doi.org/10.48550/arXiv.2403.05525.

Luo, J., Sun, Y., Qian, Z., Zhou, L., Wang, J., 2023. Overview and prospect of artificial intelligence large models. Radiotehnika 53 (11), 2461–2472. https://doi.org/10.3969/j.issn.1003-3106.2023.11.001.

Naghshvarianjahromi, M., Kumar, S., Deen, M.J., 2023. Natural intelligence as the brain of intelligent systems. Sensors 23 (5), 2859. https://doi.org/10.3390/s23052859.

News, J., 2024. Emergency management "Jiuan" large model officially released. China Occupational Safety and Health 19 (8), 4. https://doi.org/10.20115/j.cnki.cn11-5404/x.2024.08.003.

Peng, Y., Malin, B.A., Rousseau, J.F., Wang, Y., Xu, Z., Xu, X., Weng, C., Bian, J., 2025. From GPT to DeepSeek: significant gaps remains in realizing AI in healthcare. J. Biomed. Inf. 163, 104791. https://doi.org/10.1016/j.jbi.2025.104791.

Pop, D.-P., Altar, A., 2014. Designing an MVC model for rapid web application development. Procedia Eng. 69, 1172–1179. https://doi.org/10.1016/j.proeng.2014.03.106.

Shao, X., Ma, S., Xu, C., Xie, C., Li, T., Huang, Y., Huang, Y., Xiao, Z., 2024. Landslides triggered by the 2022 Ms. 6.8 Luding strike-slip earthquake: an update. Eng. Geol. 335, 107536. https://doi.org/10.1016/j.enggeo.2024.107536.

Shirzaei, M., Vahedifard, F., Sadhasivam, N., Ohenhen, L., Dasho, O., Tiwari, A., Werth, S., Azhar, M., Zhao, Y., Nicholls, R.J., 2025. Aging dams, political instability, poor human decisions and climate change: recipe for human disaster. npj Natural Hazards 2 (1), 5. https://doi.org/10.1038/s44304-024-00056-1.

Sun, B., 2024. Review of large models. Comput. Simulat. 41 (1), 1–7+24. https://doi.org/10.3969/j.issn.1006-9348.2024.01.002.

Tang, R., Chuang, Y.-N., Hu, X., 2024. The science of detecting LLM-generated text. Commun. ACM 67 (4), 50–59. https://doi.org/10.1145/3624725.

Team, K., Du, A., Gao, B., Xing, B., Jiang, C., Chen, C., Li, C., Xiao, C., Du, C., Liao, C., 2025. Kimi k1. 5: scaling reinforcement learning with llms. arXiv preprint arXiv: 2501.12599. https://doi.org/10.48550/arXiv.2501.12599.

Wang, M., Yin, T., Yang, H., Hu, J., Chen, J., 2023. Knowledge graphs and large language models technology development and application. Cyber Security And Data Governance 42 (S1), 126–131. https://doi.org/10.19358/j.issn.2097-1788.2023.S1.022.

Wang, X., Qi, G., 2022. Contrastive learning with stronger augmentations. IEEE Trans. Pattern Anal. Mach. Intell. 45 (5), 5549–5560. https://doi.org/10.1109/TPAMI.2022.3203630.

Wei, J., Wang, X., Schuurmans, D., Bosma, M., Xia, F., Chi, E., Le, Q.V., Zhou, D., 2022. Chain-of-thought prompting elicits reasoning in large language models. Adv. Neural Inf. Process. Syst. 35, 24824–24837.

Wu, H., 2024. Integration of large-scale models and cloud, transition from informatization to digital intelligence. Journal of Chongqing University of Posts and Telecommunications(Natural Science Edition) 36 (1), 1–8. https://doi.org/10.3979/j.issn.1673-825X.202312250431.

Wu, S., Xu, C., Ma, J., Gao, H., 2025. Escalating risks and impacts of rainfall-induced geohazards. Nat. Hazards Res. https://doi.org/10.1016/j.nhres.2025.03.003.

Xiao, B., Kantarci, B., Kang, J., Niyato, D., Guizani, M., 2024. Efficient prompting for llm-based generative internet of things. IEEE Internet Things J. 12 (1), 778–791. https://doi.org/10.1109/JIOT.2024.3470210.

Xiao, Z., Xu, C., Huang, Y., He, X., Shao, X., Chen, Z., Xie, C., Li, T., Xu, X., 2023. Analysis of spatial distribution of landslides triggered by the Ms 6.8 Luding earthquake in China on September 5, 2022. Geoenvironmental Disasters 10 (1), 1–15. https://doi.org/10.1186/s40677-023-00233-w.

Xie, C., Xu, C., Huang, Y., Liu, J., Jin, J., Xu, X., Cheng, J., Wu, L., 2025a. Detailed inventory and initial analysis of landslides triggered by extreme rainfall in the northern Huaiji County, Guangdong Province, China, from June 6 to 9, 2020. Geoenvironmental Disasters 12 (1), 7. https://doi.org/10.1186/s40677-025-00311-1.

Xie, C., Xu, C., Huang, Y., Liu, J., Shao, X., Xu, X., Gao, H., Ma, J., Xiao, Z., 2025b. Advances in the study of natural disasters induced by the" 23.7" extreme rainfall event in North China. Nat. Hazards Res 5 (1), 1–13. https://doi.org/10.1016/j.nhres.2025.01.003.

Xu, C., 2023. An introduction to "application of novel high-tech methods to geological hazard research." Nat. Hazards Res 3 (2), 353–357. https://doi.org/10.1016/j.nhres.2023.05.001.

Xu, C., Xu, X., Yao, X., Dai, F., 2014a. Three (nearly) complete inventories of landslides triggered by the May 12, 2008 Wenchuan Mw 7.9 earthquake of China and their spatial distribution statistical analysis. Landslides 11, 441–461. https://doi.org/10.1007/s10346-013-0404-6.

Xu, C., Xue, Z., 2024. Applications and challenges of artificial intelligence in the field of disaster prevention, reduction, and relief. Nat. Hazards Res 4 (1), 169–172. https://doi.org/10.1016/j.nhres.2023.11.011.

Xu, L., Meng, X., Xu, X., 2014b. Natural hazard chain research in China: a review. Nat. Hazards 70, 1631–1659. https://doi.org/10.1007/s11069-013-0881-x.

Xu, X., Wang, S., Cheng, J., Wu, X., 2025. Shaking the Tibetan Plateau: insights from the Mw 7.1 Dingri earthquake and its implications for active fault mapping and disaster mitigation. npj Natural Hazards 2 (1), 16. https://doi.org/10.1038/s44304-025-00074-7.

Xu, Y., Hu, L., Zhao, J., Du, W., Wang, W., 2024. Technology application prospects and risk challenges of large language models. J. Comput. Appl. 44 (6), 1655–1662. https://doi.org/10.11772/j.issn.1001-9081.2023060885.

Xue, Z., Xu, C., Xu, X., 2023. Application of ChatGPT in natural disaster prevention and reduction. Nat. Hazards Res 3 (3), 556–562. https://doi.org/10.1016/j.nhres.2023.07.005.

Yang, A., Yang, B., Zhang, B., Hui, B., Zheng, B., Yu, B., Li, C., Liu, D., Huang, F., Wei, H., 2024. Qwen2. 5 technical report. arXiv preprint arXiv:2412.15115. https://doi.org/10.48550/arXiv.2412.15115.

Yao, Y., Duan, J., Xu, K., Cai, Y., Sun, Z., Zhang, Y., 2024. A survey on large language model (llm) security and privacy: the good, the bad, and the ugly. High-confid comput. 4 (2), 100211. https://doi.org/10.1016/j.hcc.2024.100211.

Yeom, J., Lee, H., Byun, H., Kim, Y., Byun, J., Choi, Y., Kim, S., Song, K., 2024. Tc-llama 2: fine-tuning LLM for technology and commercialization applications. J. Big Data 11 (1), 100. https://doi.org/10.1186/s40537-024-00963-0.

Zhao, C., Zhu, G., Wang, J., 2023. The inspiration brought by ChatGPT to LLM and the new development ideas of multi-modal large model. Data Anal. Knowl. Discov. 7 (3), 26–35. https://doi.org/10.11925/infotech.2096-3467.2023.0216.

View publication stats