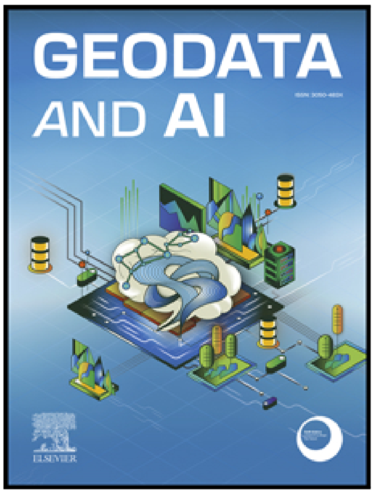
Journal Pre-proof



Perspectives: LLM agents reshaping the foundation of geotechnical problem-solving

Stephen Wu, Chao Shi, Andy Leung, Yu Otake, Chisato Konishi, Mingliang Zhou, Yuanqin Tao, Zijun Cao, Tomoka Nakamura

PII: S3050-483X(25)00035-8
DOI: https://doi.org/10.1016/j.geoai.2025.100036
Reference: GEOAI 100036

To appear in: *Geodata and AI*

Received date: 7 June 2025
Revised date: 14 August 2025
Accepted date: 27 August 2025

Please cite this article as: Stephen Wu, Chao Shi, Andy Leung, Yu Otake, Chisato Konishi, Mingliang Zhou, Yuanqin Tao, Zijun Cao, Tomoka Nakamura, Perspectives: LLM agents reshaping the foundation of geotechnical problem-solving, *Geodata and AI* (2025), doi: https://doi.org/10.1016/j.geoai.2025.100036

Highlights

Perspectives : LLM agents reshaping the foundation of geotechnical problem-solving

Stephen Wu,Chao Shi,Andy Leung,Yu Otake,Chisato Konishi,Mingliang Zhou,Yuanqin Tao,Zijun Cao,Tomoka Nakamura

- Comment of Agentic AI application in Geotechnics

Perspectives : LLM agents reshaping the foundation of geotechnical problem-solving

Stephen Wu[a,b], Chao Shi[c], Andy Leung[d], Yu Otake[e], Chisato Konishi[f], Mingliang Zhou[g], Yuanqin Tao[h], Zijun Cao[i] and Tomoka Nakamura[e]

[a]Research Organization of Information and Systems, The Institute of Statistical Mathematics, 10-3 Midori-cho, Tachikawa, Tokyo, 190-8562, Japan
[b]The Graduate University for Advanced Studies (SOKENDAI), 10-3 Midori-cho, Tachikawa, Tokyo, 190-8562, Japan
[c]School of Civil and Environmental Engineering, Nanyang Technological University, 639798, Singapore
[d]Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong S.A.R.
[e]Department of Civil and Environmental Engineering, Tohoku University, Sendai, Miyagi, 980-8579, Japan
[f]OYO Corporation, Chiyoda, Tokyo, 101-8486, Japan
[g]College of Civil Engineering, Tongji University, Shanghai , 200092, China
[h]Zhejiang University of Technology, Hangzhou, Zhejiang, 310023, China
[i]MOE Key Laboratory of High-Speed Railway Engineering, Institute of Smart City and Intelligent Transportation, Southwest Jiaotong University, Chengdu, Sichuan, 611756, China

ARTICLE INFO

Keywords:
Agentic AI
LLMs
Foundation models

ABSTRACT

This paper explores the transformative potential of Large Language Model (LLM)-based agentic artificial intelligence (AI) in addressing longstanding challenges in geotechnical engineering. It begins by highlighting the significant growth and increasing interest in applying machine learning (ML) and AI techniques across various geotechnical domains, such as soil classification, slope stability analysis, and foundation design. Emphasizing the Gartner Hype Cycle, the authors reflect on the transition from initial enthusiasm toward realistic appraisal and adoption, highlighting current barriers like limited foundational understanding, skepticism about AI reliability, and a lack of standardized practices. The authors then introduce LLM agents as promising solutions for automating the extraction, interpretation, and quantification of qualitative and semi-quantitative geotechnical data. Drawing insights from the 1st GeoTechathon event, an international collaboration involving engineers, data scientists, and AI practitioners, the paper demonstrates practical applications in geotechnical site planning, landslide investigations, liquefaction analysis, and shield tunnel safety evaluation. Each project leveraged basic techniques, including Retrieval-Augmented Generation (RAG), multimodal data integration, and prompt engineering, achieving improvements in efficiency, accuracy, and decision-making processes. The paper concludes by discussing broader implications for interdisciplinary collaboration, ethical considerations, and future directions, emphasizing the necessity for standardized practices, rigorous validation, and enhanced AI literacy to sustainably integrate LLM technologies within the geotechnical engineering community.

1. Introduction

Over the past two decades, the integration of machine learning (ML) and artificial intelligence (AI) techniques into geotechnical engineering has received significant attention. A growing body of literature reflects this trend, with numerous studies exploring ML applications across various geotechnical domains, including site-specific soil property prediction, slope stability analysis, and foundation design [1, 2, 3, 4, 5]. Figure 1a illustrates the rapid increase in the number of publications related to ML in geotechnical engineering. This surge in research highlights the engineering community's strong interest in leveraging data-driven approaches to address complex subsurface challenges; however, this momentum may not continue indefinitely.

According to the Gartner Hype Cycle [6], a graphical model illustrating how emerging technologies evolve over time, technologies typically progress through five phases: an initial surge of excitement (Technology Trigger), inflated expectations (Peak of Inflated Expectations), subsequent disappointment when technologies fail to meet unrealistic expectations (Trough of Disillusionment), gradual understanding and realistic use (Slope of Enlightenment), and

stewu@ism.ac.jp (S. Wu)
ORCID(s):

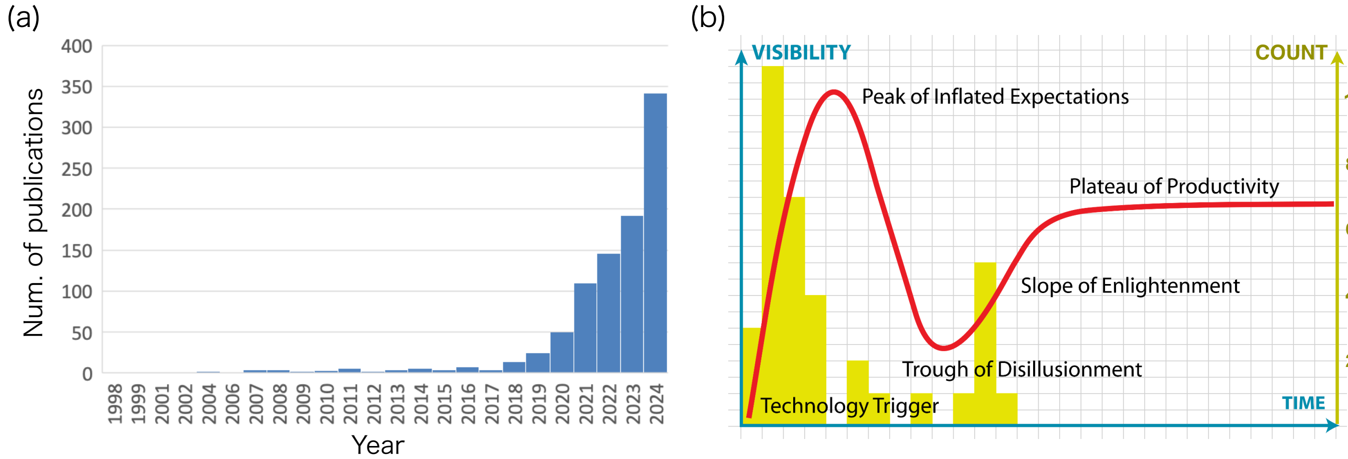


Figure 1: Survey on ML applications in geotechnical engineering. (a) Number of publications found in Web of Science containing the keywords "machine learning" and "geotechnical." (b) Number of respondents (yellow bars) selecting each stage of the Gartner Hype Cycle, represented by the red curve and cyan axis. The plot is modified from graphics by Jeremykemp at English Wikipedia, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=10547051.

finally, widespread adoption and productivity (Plateau of Productivity). Recently, we conducted an informal survey within the geotechnical community in Japan to better understand engineers' perspectives on ML and AI technologies (see Figure 1b for the result summary). Among the 36 respondents, who included junior and senior industry engineers as well as academic researchers and government agents, most expressed positive expectations for ML and AI applications in geotechnical engineering. Several respondents suggested that these technologies are currently transitioning from initial hype toward a more realistic appraisal phase. While professionals widely acknowledge AI's potential to enhance infrastructure monitoring, predictive analysis, and automation, practical adoption has lagged due to limited foundational understanding, skepticism regarding reliability and explainability, and a lack of standardized practices. Respondents emphasized that reaching the subsequent "Slope of Enlightenment" requires deeper AI literacy, strategic integration, and accessible datasets to enable the effective, sustainable implementation of AI-driven solutions in geotechnical engineering.

While there are numerous challenges associated with moving beyond the "Trough of Disillusionment" in Gartner Hype Cycle, a key issue in geotechnical engineering relates to the inherent nature of the problems engineers commonly face. Problems suited to physical or statistical modeling typically have well-defined quantitative input features and clearly established quantitative evaluation metrics, allowing for systematic optimization based on observed data. In contrast, many geotechnical engineering challenges inherently include qualitative or semi-quantitative features, such as subjective soil classifications, visual assessments of geological conditions, or expert-driven interpretations of subsurface variability. These qualitative characteristics often compel engineers to rely on empirical methods or expert judgment rather than standardized quantitative frameworks, complicating systematic optimization and hindering the direct adoption of ML and AI technologies [7]. To effectively employ ML and AI approaches, these qualitative features must first be translated into quantitative metrics—a task that usually involves substantial human effort and significant labor costs. This requirement highlights the critical importance of automation. Recently emerging technologies, such as Large Language Model (LLM) agents, offer potential for automating the extraction, interpretation, and quantification of traditionally qualitative information, thereby bridging this gap (Figure 2). Furthermore, LLM agents can provide human-understandable explanations for model outputs, a crucial capability for geotechnical applications where risk-averse decision-making demands transparency and interpretability.

Exploring the potential of LLM agents in geotechnical applications can hardly be accomplished by any single individual; instead, it requires a community-wide effort involving collective discussion, experimentation, and iterative trial-and-error. This paper reviews insights gained from the 1st GeoTechathon, an international hackathon held between May and September 2024, which brought together geotechnical engineers, data scientists, and AI practitioners to collaboratively investigate the applicability of LLM agents in addressing real-world geotechnical problems. By analyzing the approaches and outcomes of four participating teams, each focusing on distinct geotechnical topics, this paper aims to evaluate the viability of LLM agents for enhancing geotechnical problem-solving and to identify

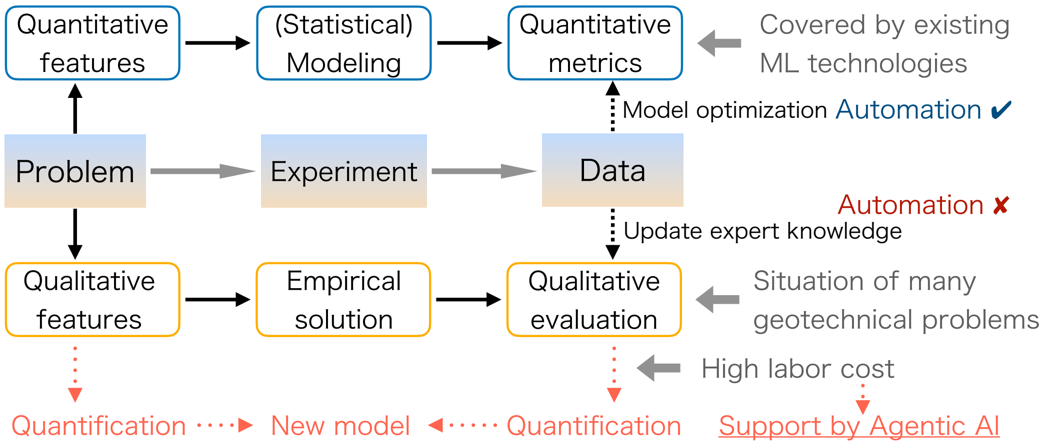


Figure 2: Overview of the challenge and potential of AI technologies in geotechnical problem solving.

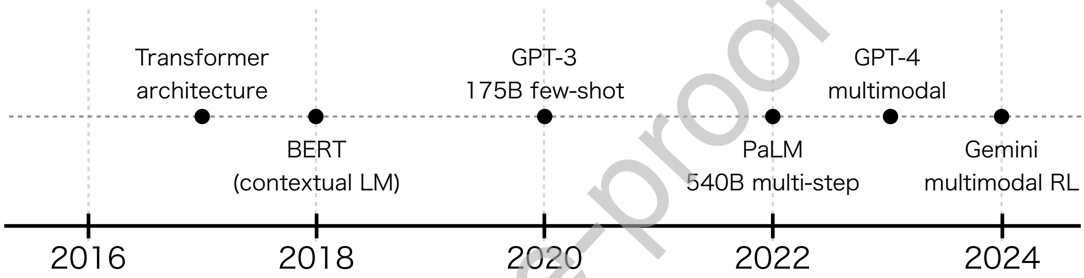


Figure 3: Selected important milestones for the advancement of multimodal foundation models.

practical pathways for integrating these technologies into engineering practice. The paper is organized as follows: First, we briefly introduce key concepts related to LLM agents. Second, we provide a general discussion of the potential of LLM agents in geotechnical engineering, supported by a review of existing literature. Third, we summarize the findings from the 1st GeoTechathon event, followed by reflections and concluding remarks.

2. Emergence of foundation models, LLMs, and agentic AI

2.1. From deep learning to LLMs and multimodal foundation models

The past decade has seen a shift in machine learning from task-specific deep architectures toward large-scale, pre-trained "foundation models" [8]. These models are trained on broad and diverse datasets, enabling adaptation to many downstream tasks. Early examples include deep image features from ImageNet [9] and word embedding models such as Word2Vec [10].

A major milestone in NLP was BERT [11], a 340M-parameter Transformer trained with masked language modeling, which achieved state-of-the-art results across multiple benchmarks [12]. GPT-3 [13] expanded this paradigm to 175B parameters, demonstrating emergent "few-shot" capabilities that established LLMs as general problem-solvers rather than single-task models. The term "foundation model" [8] captures their central yet adaptable nature. Figure 3 highlights key milestones leading to the current generation of multimodal foundation models.

More recent systems such as Google's PaLM [14], OpenAI's GPT-4 [15], and Google's Gemini [16] have pushed scale and capability further, incorporating advanced reasoning, multimodal inputs, and reinforcement learning from human feedback [17]. These developments underpin the ability of modern LLMs to process heterogeneous data sources — a critical feature for geoscience and geotechnics domains where textual, visual, and numerical data often need to be integrated.

Recent work has begun to harness LLMs and foundation models for geoscientific tasks, such as, GeoGalactica — a 30-billion-parameter LLM further pretrained and instruction-tuned using a large geoscience-specific corpus—demonstrated state-of-the-art performance on geoscience examinations and open-domain questions, as well as improved retrieval of domain knowledge [18]. The specialized geoscience LLM K2 was developed by fine-tuning LLaMA-7B on a corpus of over 5.5B geoscience tokens, paired with the GeoSignal dataset and benchmarked against GeoBench; it showed strong geoscience knowledge understanding and utility [19]. Meanwhile, GeoRSMLLM extended multimodal language models to vision-language tasks in remote sensing, addressing complex referring-expression comprehension, segmentation, and change detection through a unified point-based data representation and self-augmentation strategies [20]. These studies underscore the growing deployment of LLMs for integrating heterogeneous geoscientific data and supporting domain-specific reasoning—approaches highly relevant to geotechnical analysis.

2.2. Rise of LLM agents

Building on these advances, LLMs have evolved into agentic AI systems capable of multi-step reasoning, tool use, and autonomous task execution. Techniques such as chain-of-thought prompting [21] improve problem-solving by encouraging explicit intermediate reasoning steps. Frameworks like ReAct [22] interleave reasoning with actions such as API calls, enabling closed-loop decision-making. Systems including Toolformer and the ChatGPT plugin ecosystem allow LLMs to augment their knowledge with external software tools, expanding their ability to perform domain-specific analyses.

In parallel, LLMs have gained long-term memory [23] and planning abilities [24], enabling them to retain context between interactions, formulate subgoals, and iteratively refine solutions. Self-reflection techniques allow models to detect and correct errors, while structured approaches can use one LLM to oversee or debug another's outputs, analogous to a human supervisor. These capabilities support complex, multi-turn problem-solving in which high-level tasks are decomposed into smaller steps, executed—often via external tools or code generation—and then assembled into final results. Such auto-regressive planning underlies systems like AutoGPT [25], where an LLM cycles through generating, executing, and evaluating tasks until a goal is reached.

Language serves as a universal interface for these agentic behaviors: LLMs can follow natural language instructions from humans and also communicate with other agents to coordinate actions. This has led to collaborative multi-agent systems, with numerous applications emerging since late 2024. In science, research teams have combined multiple LLM agents with external tools to automate research workflows [26, 27], while multi-agent simulations have advanced social science studies [28].

LLM-based agentic systems tailored for geoscience are also emerging. In GeoPredict-LLM, a knowledge-graph-enhanced multimodal LLM was developed to facilitate advanced geological predictions, integrating structured ontology with geospatial data to support subsurface forecasting [29]. The GeoFactory framework demonstrated how LLM enhancement methods — such as fine-tuning, prompt engineering, and retrieval augmentation — can boost geoscience performance even with smaller models; for example, Mistral-7B enhanced through GeoFactory outperformed GPT-4 on average across various geoscience tasks [30]. These frameworks illustrate early yet promising steps toward deploying LLM agents capable of reasoning, integrating multimodal inputs, and delivering expert-level geoscientific outputs, directly aligning with the demands of geotechnical engineering workflows.

2.3. Relevance to geotechnical engineering

Geotechnical engineering presents challenges well-suited to LLM agents.

- **Qualitative-to-quantitative conversion:** Automating the extraction of structured parameters from narrative site investigation reports or borehole logs.
- **Multimodal integration:** Combining imagery (e.g., core photos, slope inspections), tabular monitoring data, and text-based guidelines into unified analyses.
- **Interdisciplinary coordination:** Acting as a communication bridge between geologists, structural engineers, and regulators by harmonizing domain-specific terminology.
- **Transparent decision support:** Producing interpretable reasoning traces for risk assessments in safety-critical projects.

- **Rapid scenario testing:** Iteratively running design or hazard simulations through tool-assisted agent workflows.

These capabilities address persistent barriers in applying AI to geotechnics, such as unstructured data formats, fragmented standards, and the need for explainable outputs, thereby motivating the applications explored in the 1st GeoTechathon (more details in Sections 3 and 4).

3. Potential Impacts of LLM Agents on Geotechnical Engineering

3.1. Accelerated Problem-Solving

LLM agents hold significant potential for accelerating the problem-solving process in geotechnical engineering by automating traditionally manual tasks. Automated data interpretation, particularly of subsurface conditions, can significantly reduce the time and effort required for preliminary analysis. For example, interpreting geotechnical investigation reports, soil classification logs, and field testing results, which traditionally involve extensive expert effort, can be systematically managed by LLM agents trained to recognize and quantify qualitative features into structured quantitative datasets. This automation not only accelerates decision-making but also reduces the potential for human error in repetitive tasks. Some preliminary evidences include a study on automating geotechnical simulation using ChatGPT [31] and building LLM agent to generate two-dimensional geological cross-sections from sparse site investigation data [32].

Additionally, LLMs can rapidly perform literature reviews and knowledge extraction from large repositories of academic and industry publications. Engineers traditionally spend significant amounts of time gathering relevant information from extensive technical literature to support project-specific analyses. LLM-driven agents can streamline this process by summarizing current research trends, extracting essential insights, and identifying knowledge gaps efficiently. The evidence of such potential can be seen in the study by Chai et al. (2025) [33], who showed the ability of customized LLMs to extract useful geotechnical information.

3.2. Enhanced Decision-Making and Risk Management

LLM agents can significantly enhance decision-making and risk management processes in geotechnical engineering. One primary benefit is their ability to conduct comprehensive scenario analyses and uncertainty quantification rapidly. By integrating historical data, expert knowledge, and real-time monitoring data, these agents can simulate numerous scenarios, evaluating potential risks and providing probabilistic assessments that are crucial for informed decision-making in projects involving slope stability, foundation design, and tunneling.

Furthermore, continuous learning capabilities of LLM agents allow them to assimilate insights from historical case studies and incidents systematically. For example, lessons learned from past geotechnical failures, such as slope collapses or foundation settlement incidents, can be continuously integrated into their predictive models. Machine learning techniques to achieve this goal include periodically adding feedback information to the "memory" of an agentic AI system by expanding the vector database in Retrieval-Augmented Generation (RAG) [34], performing supervised fine-tuning to existing LLM, or applying Reinforcement Learning from Human Feedback (RLHF) [17]. Such methods improve reliability and safety in decision-making and risk management for future projects. Xu et al. (2025) [35, 36] took the first step to explore the ability of LLM agents to support decision-making by building a LLM-based system for automating geotechnical design.

3.3. Collaboration and Knowledge-Sharing

The collaborative potential of multi-agent LLM systems significantly benefits interdisciplinary coordination, which is crucial in complex geotechnical projects. By facilitating clearer communication among diverse stakeholders, including geologists, structural engineers, construction managers, and regulators, LLM agents can minimize misunderstandings and promote effective interdisciplinary collaboration. The GeoTechathon event introduced in the later section was designed with such mindset specifically.

Moreover, the inherent transparency and reproducibility afforded by LLM agents can transform knowledge-sharing practices within the geotechnical community. These agents provide explainable reasoning behind their decisions, enabling reproducibility of analyses and facilitating better peer reviews and knowledge transfer, especially for critical infrastructure projects that demand high transparency standards. Pang et al. (2025) [37] showed an early attempt in this direction, where LLM agent is used to summarize important geotechnical information from multiple reports with citations for engineers to proofread. They also demonstrated the benefit of include a statistical analysis agent to reduce the variability of LLM outputs in the task of landslide geometry estimation.

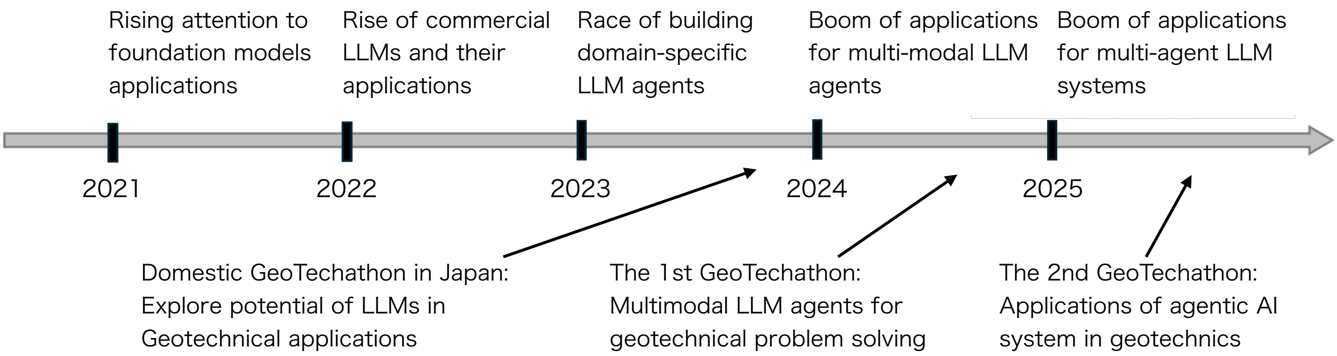


Figure 4: Timeline of GeoTechathon (bottom) matches with the different stages of LLM agent advancement (top).

3.4. Accessibility and Democratization of Expertise

One of the most profound impacts of LLM agents is their potential to democratize expert-level geotechnical knowledge, making sophisticated problem-solving capabilities accessible even in resource-constrained environments. Regions with limited access to geotechnical specialists can leverage these advanced AI agents to perform essential evaluations, ensuring consistent quality and safety in geotechnical practices worldwide. Wu et al. (2025) [38] has made a preliminary investigation of such potential through four different geotechnical applications of LLM agents, and concluded their study with a positive attitude.

Additionally, AI-driven guidelines and workflows powered by LLM agents enable the standardization of methodologies across projects, industries, and geographical regions. This standardization is crucial in geotechnical engineering, where consistency in methods and interpretations significantly influences project outcomes. LLM-driven frameworks can establish and propagate best practices, creating a standardized, globally applicable geotechnical knowledge base. Such potential is yet to be explored by the geotechnics community.

4. The 1st GeoTechathon

4.1. Overview of the event

The 1st (international) GeoTechathon was initiated in response to the growing recognition of the potential for LLM agents to overcome longstanding challenges in geotechnical engineering, as explained in previous sections. Recognizing the gap between the theoretical potential of LLM agents and their practical implementation in geotechnics, the event aimed to stimulate community-driven exploration and innovation. It is motivated from a domestic workshop in Japan to explore the same topic through a group of 20 geotechnical engineers from a highly diverse background [38]. Figure 4 shows the timeline of GeoTechathon along with the different stages of LLM agent advancement.

Held in two distinct phases from March to August 2024, the 1st (international) GeoTechathon attracted a group of participants comprising geotechnical representatives from both academia and industry. The first phase, spanning March to May, was dedicated to team formation and ideation, as well as providing participants with essential training through tutorials focused on utilizing LLM agents. This phase emphasized collaborative brainstorming to identify practical geotechnical problems that could benefit from AI-driven solutions.

The second phase, occurring from June to August, involved remote collaboration and problem-solving, facilitated by regular online meetings. Teams focused on four distinct geotechnical challenges: 3D geological modeling, slope stability assessment, liquefaction risk evaluation, and tunneling projects. Each team harnessed a variety of LLM agent technologies, such as Retrieval-Augmented Generation (RAG) [34] and multimodal data integration [39], to develop innovative and practical solutions.

Results from these efforts were showcased at a dedicated special session during the 2nd Workshop on Future of Machine Learning in Geotechnics (2FOMLIG) held from October 11–13, 2024, in Chengdu, China. At this venue, each team presented its findings and demonstrated practical implementations, highlighting both the capabilities and current limitations of LLM agents in geotechnical applications.

The outcomes of the GeoTechathon not only highlight the considerable promise of LLM agents in transforming geotechnical problem-solving but also provided valuable insights into practical integration challenges. The event

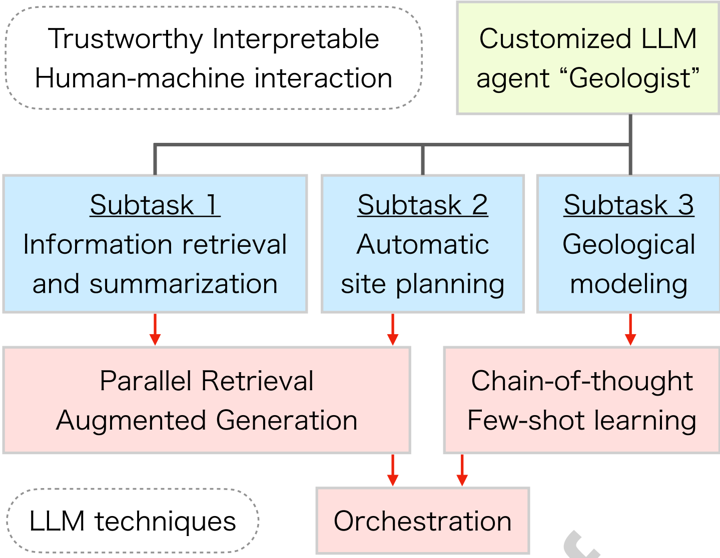


Figure 5: Outline of the agentic AI system for the geotechnical site planning team. The workflow utilizes various LLM techniques to achieve trustworthy interpretable human-machine interaction.

fostered interdisciplinary collaboration and reinforced the importance of a community-based approach in harnessing the full potential of advanced AI technologies in geotechnical engineering. The following subsections provide a brief description of each of the four studied topics. The success of this event has motivated a continuation of the GeoTechathon event, for which the 2^{nd} GeoTechathon is jointly organized with the 3rd Workshop on Future of Machine Learning in Geotechnics (3FOMLIG) held from October 15–17, 2025, in Florence, Italy.

4.2. Geotechnical site planning team

This team addressed the challenge of streamlining geotechnical site planning and geological interpretation, which traditionally involve extensive manual effort and complex decision-making processes. They developed a customized LLM-based agent named "Geologist" designed specifically to optimize these tasks.

Their approach utilized retrieval-augmented generation (RAG) techniques to accurately extract relevant design clauses from multilingual geotechnical guidelines. These clauses guided the strategic placement of boreholes, optimizing site planning processes. Additionally, they leveraged LLM agents within a sequential, interactive workflow to automate the generation of geological cross-sections from multimodal borehole data, significantly reducing manual effort and enhancing interpretability. Figure 5 shows the outline of their resulting system.

Key outcomes from this initiative included the development of a multihop-RAG system that demonstrated higher accuracy and fewer errors (hallucinations) compared to conventional methods [32]. The exercise revealed that complex geotechnical tasks could effectively be decomposed into smaller, manageable subtasks, each executed by specialized LLM agents, benefiting from ongoing human-machine interactions.

The team's results indicate substantial potential for LLM-based automation in routine geotechnical operations, significantly enhancing efficiency and accuracy. Moreover, adopting a chain-of-thought approach using LLM agents promises improved decision-making support and interpretability in handling complex geotechnical scenarios, paving the way for further advancements in the field.

4.3. Landslide investigation team

The second team focused on improving the investigation and analysis of landslide events, which typically involve understanding the causes, impacts, and specific processes by interviewing eyewitnesses and analyzing post-event imagery. This task is challenging as it involves translating layman observations into technical insights and accurately interpreting complex imagery data [37].

To tackle this problem, the team utilized Retrieval Augmented Generation (RAG) coupled with careful prompt engineering to effectively distill relevant technical details from eyewitness accounts. The LLM summarized these extracted details, providing clear, professional-grade summaries of landslide processes. Concurrently, the team

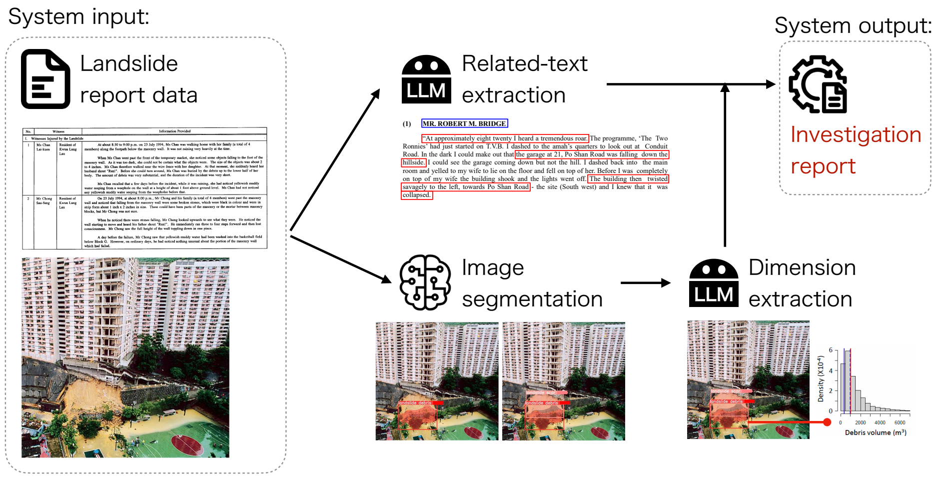


Figure 6: Outline of the agentic AI system for the landslide investigation team. Two typical modalities from post-event landslide investigation are analyzed using different AI agents. The text samples shown here are for illustration purpose only.

retrained the YOLO object-detection algorithm on a curated database of landslide images, enabling it to identify nuanced landslide features such as scars and debris. YOLO used identifiable objects within images (e.g., buildings or sports courts) as references to quantify landslide geometry accurately. Figure 6 shows the outline of the AI system developed for landslide investigation.

Key achievements of this project include the demonstration that LLM-driven RAG techniques could reliably transform eyewitness narratives into technically sound summaries consistent with professional assessments. Additionally, the team successfully showcased a workflow for image analysis that effectively quantified landslide debris geometry, though variability due to the probabilistic nature of LLM required precise prompt engineering. Here, landslide debris geometry refers to the geometry of the displaced landslide material, including the spatial extent, depth, and volume of debris deposits, as opposed to the pre-failure slope mass geometry.

This experience significantly benefited the participants, particularly research students, by providing practical exposure to applying LLM technologies to real-world problems within tight deadlines. The insights gained from this exercise establish a foundation for further explorations of LLM applications in other disaster scenarios such as earthquake reconnaissance. Despite publication challenges due to the geotechnical community's current skepticism, this approach offers substantial promise for future development and adoption.

4.4. Simulation automation team

This team aimed to automate the traditionally manual and expert-dependent estimation of parameters required for liquefaction analysis, particularly when utilizing advanced constitutive models. These models typically rely on interpreting soil test data, such as stress-strain curves and effective stress paths, frequently represented as images.

Their innovative approach leveraged LLM agents to extract nuanced and qualitative geotechnical insights from these soil test images (Figure 7). Two sets of features that represent the geotechnical insights were extracted through different AI agents: (1) using traditional image processing models to extract image features that are summarized in text, and (2) using LLM to directly extract features by including engineers' opinion in the prompts. The extracted features were then utilized to estimate model parameters, embedding the tacit and experiential knowledge traditionally provided by geotechnical experts directly into the analysis process. This methodology represented a significant departure from conventional manual interpretation, utilizing high-dimensional visual inputs guided by embedded engineering knowledge within the LLM.

Key outcomes included the development and demonstration of a novel framework capable of processing high-dimensional image data for geotechnical analysis. Among various methodologies tested, the approach employing specific engineering-focused attention points in guiding the LLM proved most effective, yielding notably superior performance in numerical simulations.

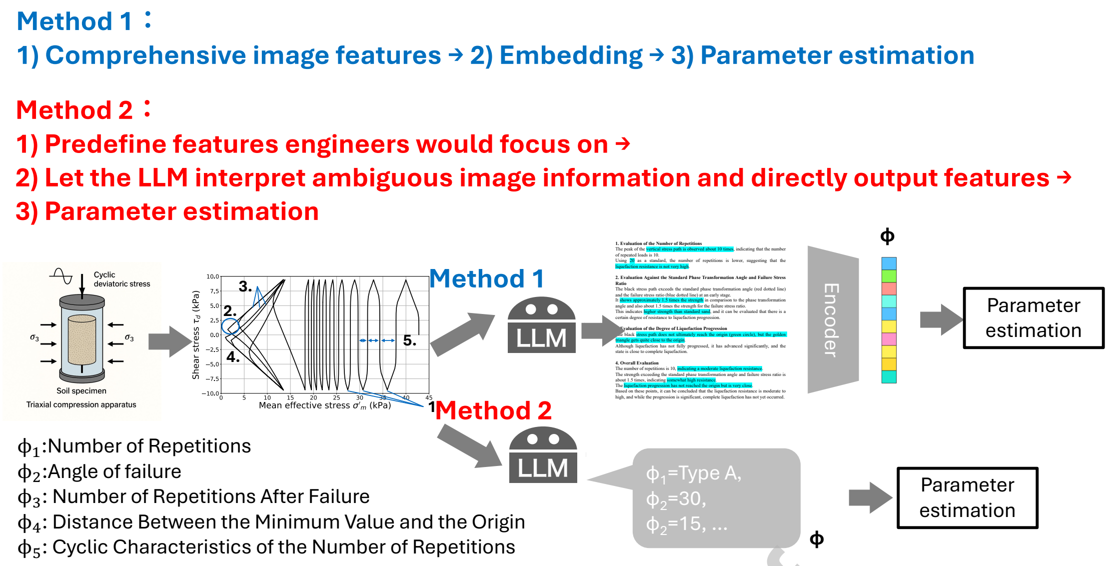


Figure 7: Outline of the agentic AI system for the simulation automation team. Two sets of features are used to build different AI agents for estimating the simulation parameters. The text sample shown here is for illustration purpose only.

The work undertaken by this team significantly advances automation in geotechnical numerical analysis, enhancing efficiency and accuracy. It also supports the evolution from traditional specification-based designs to performance-based, robust, and resilience-focused geotechnical engineering paradigms. This initiative underscores the potential of LLMs to substantially transform the future landscape of geotechnical analysis and design.

4.5. Shield tunnel safety evaluation team

The fourth team tackled the critical challenge of ensuring shield tunnel safety, a task traditionally dependent on expert-driven data collection and analysis. Recognizing the limitations and potential errors inherent in this manual process, the team pursued an innovative approach utilizing LLM-based agentic AI to automate and streamline the entire safety assessment workflow, from data acquisition to comprehensive analysis.

Their approach involved the development of a structured, LLM-based agent employing a multi-stage process. Initially, they focused on prompt engineering to efficiently retrieve relevant shield tunnel safety data with precision and context-awareness. Subsequently, this data was tokenized to facilitate semantic processing within the LLM. Lastly, a hierarchical prioritization mechanism was established, allowing the LLM-based agent to automate responses and safety evaluations effectively. The key idea to form an effective hierarchical structure was to make use of the natural hierarchical structure of the table-of-context and combine with a RAG system.

Key achievements of the project included successfully demonstrating end-to-end automation of the safety evaluation process. The developed LLM-based agent could seamlessly manage diverse data inputs, providing automated safety assessments complete with annotated risk evaluations and compliance checks. Additionally, the team created a specialized safety evaluation corpus, significantly enhancing the agent's capability to accurately identify key safety parameters and maintain full-process traceability.

This innovative solution profoundly impacts traditional safety evaluation practices, significantly improving efficiency and reliability in shield tunnel safety assessments. It supports a strategic shift towards data-driven safety models, enhancing accuracy and enabling lifecycle-wide applications. Looking forward, integrating real-time IoT sensor data and UAV inspection imagery with the semantic processing capabilities of LLMs promises further advancements, transitioning from reactive evaluations to proactive risk management and predictive safety interventions.

5. Reflections and Perspectives

The inaugural GeoTechathon demonstrated a significant step forward in applying LLM-based agentic AI to geotechnical engineering, marking a promising evolution in how classical problems are approached within the community. The event successfully generated widespread interest, evident through strong audience engagement during

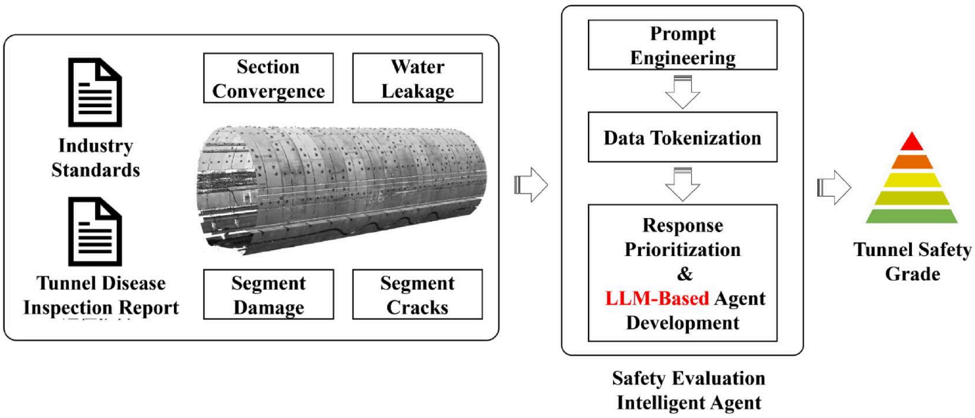


Figure 8: Outline of the agentic AI system for the shield tunnel saftey evaluation team. Multiple sources of documents are used to provide a better safety evaluation scheme.

presentations at the 2FOMLIG special session. It effectively showcased how LLMs could automate traditionally expert-driven tasks, thereby bringing efficiency, accuracy, and novel capabilities into geotechnical problem-solving.

Participants gained valuable insights into the practical integration of generative AI, recognizing its unexpected potential in complex scenarios typically dominated by human judgment. Teams were surprised to discover how effectively technologies like YOLO could be trained with relatively small datasets to differentiate nuanced geotechnical features. Additionally, the importance of clearly defining LLM roles, proper task design, and the necessity of carefully structuring prompts emerged as key learnings, underscoring the potential of LLM-based agents to substantially support or even substitute traditional human decision-making.

However, teams identified areas for improvement, such as the need for more structured pre-event training, provision of worked examples, and clearer datasets. Interdisciplinary friction and the rapid evolution of AI technologies posed additional challenges, highlighting the necessity for continuous learning and adaptation. The suggestions included reducing barriers through standardized data, clear role definitions, and structured verification frameworks, alongside nurturing a sustainable innovation ecosystem to foster continued growth and implementation.

Overall, the GeoTechathon represented an impactful and groundbreaking initiative, effectively demonstrating LLM's transformative potential and setting the foundation for future explorations in integrating cutting-edge AI into geotechnical engineering. It is clear that the extremely rapid advancement of LLM-based agentic AI technologies is within reach for geotechnical engineering when experts from diverse fields actively collaborate and engage in meaningful discussions.

6. Future Outlook and Recommendations

The past decade's progress in machine learning — from deep learning to foundation models to multi-agent systems – has vastly expanded the scope of what AI can do. We have moved from training single networks for single tasks, to building general-purpose models (BERT, GPT-3, PaLM, etc.) that serve as the foundation for myriad tasks, and now to arranging multiple such models into autonomous agentic systems capable of tackling complex, multi-faceted problems. The increasingly sophisticated capabilities of LLM agents (reasoning, tool use, planning, and collaboration) open up new engineering possibilities. Today's LLM-based agents can draft detailed plans, write and debug software, negotiate in natural language, use external tools, and work together towards a common goal – all through the flexible interface of language. This represents a significant shift toward more general intelligence in AI: rather than being limited to narrow domains, foundation models and LLM agents show promise in approaching open-ended challenges that require a mix of knowledge, reasoning, and interaction. As research continues, we expect multi-agent AI to become an increasingly standard approach for solving the toughest engineering problems – much like human experts collaborating as a team. The evolution is ongoing, but the trajectory is clear: foundation models have given AI a broad base of knowledge and skills, LLM agents are equipping it with reasoning and action, and agentic multi-agent systems are assembling these pieces to meet complex, real-world tasks in a robust and adaptive way.

These technologies represent the transformative future of geotechnical engineering. By automating traditionally qualitative and labor-intensive tasks, agentic AI systems offer unprecedented opportunities to enhance efficiency, accuracy, and consistency in geotechnical assessments through effective information extraction from multimodal data. Nonetheless, it is crucial to acknowledge that the current generation of LLM agents faces significant technical limitations. Reliability issues such as hallucinations and biases remain prevalent, and their performance is highly sensitive to prompt engineering, dataset quality, and task specificity. Moreover, ethical concerns regarding accountability, transparency, and explainability must be rigorously addressed. Decisions driven by AI inherently involve complexities that can affect public safety, infrastructure stability, and environmental protection. The geotechnical community must prioritize developing frameworks to ensure accountability, interpretability, and ethical considerations in the deployment of these technologies.

Regulatory and acceptance barriers also pose considerable challenges. Industry adoption of LLM-based agents requires not only technical validation but also a cultural shift towards accepting AI-driven methodologies. It is imperative for industry stakeholders and regulatory bodies to collaborate closely in creating comprehensive guidelines, standards, and best practices. Effective strategies for validation, verification, and responsible deployment include rigorous benchmarking against traditional methods, transparent reporting of failures and limitations, and systematic peer reviews. Pilot studies and controlled deployments are essential to evaluate the effectiveness and reliability of AI agents, particularly under diverse geotechnical scenarios.

Envisioning the future, advancements in LLM-based agentic AI technologies could integrate multi-modal data fusion, real-time data processing from IoT-enabled sensors, and seamless interaction among multiple specialized agents. Coupled with the rapid development and increasing accessibility of cloud and high-performance computing resources, heavy computational workloads could be offloaded to these backends while lightweight local agents handle on-site data collection and decision-making. These capabilities would enhance real-time decision-making, predictive maintenance, and proactive risk management in geotechnical engineering. To facilitate widespread adoption, the geotechnical community should actively pursue several key steps:

(1) Conduct focused pilot studies and collaborative research to build confidence in AI capabilities.

(2) Engage the community through workshops, hackathons, and knowledge-sharing events, promoting interdisciplinary collaboration.

(3) Develop and disseminate standardized datasets and clear guidelines for validation and benchmarking.

(4) Establish dedicated forums for dialogue between researchers, practitioners, and regulatory bodies to address ethical, regulatory, and practical considerations.

(5) Advocate for educational initiatives and training programs to improve AI literacy within the engineering community.

By systematically addressing these challenges and actively promoting interdisciplinary and community-wide collaboration, the geotechnical engineering field can realize the full potential of agentic AI technologies, substantially transforming its practices and achieving robust and reliable outcomes.

7. Disclaimer

During the preparation of this work the author(s) used ChatGPT-4.5 in order to proofread the manuscript. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Declaration of interests

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Stephen Wu and Zijun Cao hold editorial roles, as guest editor and managing editor, respectively, with the journal to which this manuscript has been submitted, which is considered a potential competing interest. Given their role as Geodata and AI, Stephen Wu and Zijun Cao had no involvement in the peer review of this article and had no access to information regarding its peer review. Full responsibility for the editorial process for this article was delegated to another journal editor. All other authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

8. Acknowledgments

This research was supported by “Strategic Research Projects” grant from ROIS (Research Organization of Information and Systems). Wu and Otake would like to thank Prof. Akiyoshi Kamura, the Vice Secretary-General of the Tohoku brunch of the Japanese geotechnical society, for his support on the Gartner Hype Cycle survey. Cao would like to express sincere appreciation for the financial support from the Fundamental Research Funds for the Central Universities (Project No. 2682025CG002).

References

[1] K.-K. Phoon, J. Ching, T. Shuku, Challenges in data-driven site characterization, Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards 16 (1) (2022) 114–126. arXiv:https://doi.org/10.1080/17499518.2021.1896005, doi:10.1080/17499518.2021.1896005.
URL https://doi.org/10.1080/17499518.2021.1896005

[2] Y. Wang, H. M. Tian, Digital geotechnics: from data-driven site characterisation towards digital transformation and intelligence in geotechnical engineering, Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards 18 (1) (2023) 8–32. doi:10.1080/17499518.2023.2278136.
URL https://doi.org/10.1080/17499518.2023.2278136

[3] M.-X. Wang, Y. Leung, D.-Q. Li, Neural network-assisted generic predictive models of safety factor and yield acceleration for seismic slope stability and displacement assessments, Canadian Geotechnical Journal 62 (2025) 1–19. doi:10.1139/cgj-2024-0106.

[4] Z.-H. Qian, C. Shi, Y. Wang, Z.-J. Cao, Nonparametric and continuous variable-based stratigraphic modelling from sparse boreholes using signed distance function and bayesian compressive sensing, Canadian Geotechnical Journal 62 (2025) 1–23. doi:10.1139/cgj-2024-0295.

[5] C. Tang, K.-K. Phoon, Model Uncertainties in Foundation Design, 1st Edition, CRC Press, 2021.

[6] Gartner, Hype Cycle for Emerging Technologies, Gartner Inc., https://www.gartner.com/en/research/methodologies/gartner-hype-cycle (2023).

[7] S. Wu, Y. Otake, Y. Higo, I. Yoshida, Pathway to a fully data-driven geotechnics: Lessons from materials informatics, Soils and Foundations 64 (3) (2024) 101471. doi:https://doi.org/10.1016/j.sandf.2024.101471.
URL https://www.sciencedirect.com/science/article/pii/S0038080624000490

[8] R. Bommasani, D. A. Hudson, E. Adeli, R. Altman, S. Arora, S. von Arx, M. S. Bernstein, J. Bohg, A. Bosselut, E. Brunskill, E. Brynjolfsson, S. Buch, D. Card, R. Castellon, N. S. Chatterji, A. S. Chen, K. A. Creel, J. Davis, D. Demszky, C. Donahue, M. Doumbouya, E. Durmus, S. Ermon, J. Etchemendy, K. Ethayarajh, L. Fei-Fei, C. Finn, T. Gale, L. E. Gillespie, K. Goel, N. D. Goodman, S. Grossman, N. Guha, T. Hashimoto, P. Henderson, J. Hewitt, D. E. Ho, J. Hong, K. Hsu, J. Huang, T. F. Icard, S. Jain, D. Jurafsky, P. Kalluri, S. Karamcheti, G. Keeling, F. Khani, O. Khattab, P. W. Koh, M. S. Krass, R. Krishna, R. Kuditipudi, A. Kumar, F. Ladhak, M. Lee, T. Lee, J. Leskovec, I. Levent, X. L. Li, X. Li, T. Ma, A. Malik, C. D. Manning, S. P. Mirchandani, E. Mitchell, Z. Munyikwa, S. Nair, A. Narayan, D. Narayanan, B. Newman, A. Nie, J. C. Niebles, H. Nilforoshan, J. F. Nyarko, G. Ogut, L. Orr, I. Papadimitriou, J. S. Park, C. Piech, E. Portelance, C. Potts, A. Raghunathan, R. Reich, H. Ren, F. Rong, Y. H. Roohani, C. Ruiz, J. Ryan, C. R’e, D. Sadigh, S. Sagawa, K. Santhanam, A. Shih, K. P. Srinivasan, A. Tamkin, R. Taori, A. W. Thomas, F. Tramèr, R. E. Wang, W. Wang, B. Wu, J. Wu, Y. Wu, S. M. Xie, M. Yasunaga, J. You, M. A. Zaharia, M. Zhang, T. Zhang, X. Zhang, Y. Zhang, L. Zheng, K. Zhou, P. Liang, On the opportunities and risks of foundation models, arXiv abs/2108.07258 (2021).

[9] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. Fei-Fei, Imagenet: A large-scale hierarchical image database, in: 2009 IEEE Conference on Computer Vision and Pattern Recognition, 2009, pp. 248–255. doi:10.1109/CVPR.2009.5206848.

[10] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, J. Dean, Distributed representations of words and phrases and their compositionality, in: C. Burges, L. Bottou, M. Welling, Z. Ghahramani, K. Weinberger (Eds.), Advances in Neural Information Processing Systems, Vol. 26, Curran Associates, Inc., 2013.
URL https://proceedings.neurips.cc/paper_files/paper/2013/file/9aa42b31882ec039965f3c4923ce901b-Paper.pdf

[11] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, BERT: Pre-training of deep bidirectional transformers for language understanding, in: J. Burstein, C. Doran, T. Solorio (Eds.), Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), Association for Computational Linguistics, Minneapolis, Minnesota, 2019, pp. 4171–4186. doi:10.18653/v1/N19-1423.

[12] C. Zhou, Q. Li, C. Li, J. Yu, Y. Liu, G. Wang, K. Zhang, C. Ji, Q. Yan, L. He, H. Peng, J. Li, J. Wu, Z. Liu, P. Xie, C. Xiong, J. Pei, P. S. Yu, L. Sun, A comprehensive survey on pretrained foundation models: a history from bert to chatgpt, International Journal of Machine Learning and Cybernetics (2024). doi:10.1007/s13042-024-02443-6.

[13] T. Brown, B. Mann, N. Ryder, M. Subbiah, J. D. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, D. Amodei, Language models are few-shot learners, in: H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, H. Lin (Eds.), Advances in Neural Information Processing Systems, Vol. 33, Curran Associates, Inc., 2020, pp. 1877–1901.

[14] A. Chowdhery, S. Narang, J. Devlin, M. Bosma, G. Mishra, A. Roberts, P. Barham, H. W. Chung, C. Sutton, S. Gehrmann, P. Schuh, K. Shi, S. Tsvyashchenko, J. Maynez, A. Rao, P. Barnes, Y. Tay, N. Shazeer, V. Prabhakaran, E. Reif, N. Du, B. Hutchinson, R. Pope, J. Bradbury, J. Austin, M. Isard, G. Gur-Ari, P. Yin, T. Duke, A. Levskaya, S. Ghemawat, S. Dev, H. Michalewski, X. Garcia, V. Misra, K. Robinson, L. Fedus, D. Zhou, D. Ippolito, D. Luan, H. Lim, B. Zoph, A. Spiridonov, R. Sepassi, D. Dohan, S. Agrawal, M. Omernick, A. M. Dai, T. S. Pillai, M. Pellat, A. Lewkowycz, E. Moreira, R. Child, O. Polozov, K. Lee, Z. Zhou, X. Wang, B. Saeta, M. Diaz, O. Firat, M. Catasta, J. Wei, K. Meier-Hellstern, D. Eck, J. Dean, S. Petrov, N. Fiedel, Palm: scaling language modeling with pathways, J. Mach. Learn. Res. 24 (1) (Jan. 2023).

[15] OpenAI, GPT-4 Technical Report, arXiv (2023) 2303.08774.
URL https://arxiv.org/abs/2303.08774

[16] Gemini Team, Google, Gemini: A family of highly capable multimodal models, arXiv (2023) 2312.11805.
URL https://arxiv.org/abs/2312.11805

[17] L. Ouyang, J. Wu, X. Jiang, D. Almeida, C. L. Wainwright, P. Mishkin, C. Zhang, S. Agarwal, K. Slama, A. Ray, J. Schulman, J. Hilton, F. Kelton, L. Miller, M. Simens, A. Askell, P. Welinder, P. Christiano, J. Leike, R. Lowe, Training language models to follow instructions with human feedback, in: Proceedings of the 36th International Conference on Neural Information Processing Systems, Vol. 2011 of NIPS '22, Curran Associates Inc., Red Hook, NY, USA, 2022, pp. 27730–27744.

[18] Z. Lin, C. Deng, L. Zhou, T. Zhang, Y. Xu, Y. Xu, Z. He, Y. Shi, B. Dai, Y. Song, B. Zeng, Q. Chen, Y. Miao, B. Xue, W. Wang, J. He, Y. Zhu, X. Wang, C. Zhou, Geogalactica: A scientific large language model in geoscience, arXiv preprint arXiv:2401.00434Preprint (2024).

[19] C. Deng, T. Zhang, Z. He, Q. Chen, Y. Shi, Y. Xu, L. Fu, W. Zhang, X. Wang, C. Zhou, Z. Lin, J. He, K2: A foundation language model for geoscience knowledge understanding and utilization, in: Proceedings of the 17th ACM International Conference on Web Search and Data Mining, WSDM '24, Association for Computing Machinery, New York, NY, USA, 2024, pp. 161—170. doi:10.1145/3616855.3635772.
URL https://doi.org/10.1145/3616855.3635772

[20] Z. Zhang, H. Shen, T. Zhao, B. Chen, Z. Guan, Y. Wang, X. Jia, Y. Cai, Y. Shang, J. Yin, Georsmllm: A multimodal large language model for vision-language tasks in geoscience and remote sensing, arXiv preprint arXiv:2503.12490Preprint (2025).

[21] J. Wei, X. Wang, D. Schuurmans, M. Bosma, B. Ichter, F. Xia, E. H. Chi, Q. V. Le, D. Zhou, Chain-of-thought prompting elicits reasoning in large language models, in: Proceedings of the 36th International Conference on Neural Information Processing Systems, NIPS '22, Curran Associates Inc., Red Hook, NY, USA, 2022.

[22] S. Yao, J. Zhao, D. Yu, N. Du, I. Shafran, K. Narasimhan, Y. Cao, React: Synergizing reasoning and acting in language models, CoRR abs/2210.03629 (2022).
URL https://doi.org/10.48550/arXiv.2210.03629

[23] K. Hatalis, D. Christou, J. Myers, S. Jones, K. Lambert, A. Amos-Binks, Z. Dannenhauer, D. Dannenhauer, Memory matters: The need to improve long-term memory in llm-agents, Proceedings of the AAAI Symposium Series 2 (1) (2024) 277–280.
URL https://ojs.aaai.org/index.php/AAAI-SS/article/view/27688

[24] K. Valmeekam, M. Marquez, S. Sreedharan, S. Kambhampati, On the planning abilities of large language models: a critical investigation, in: Proceedings of the 37th International Conference on Neural Information Processing Systems, no. 3320 in NIPS '23, Curran Associates Inc., Red Hook, NY, USA, 2023, pp. 75993–76005.

[25] H. Yang, S. Yue, Y. He, Auto-GPT for online decision making: Benchmarks and additional opinions, arXiv (2023) 2306.02224.
URL https://arxiv.org/abs/2306.02224

[26] C. Lu, C. Lu, R. T. Lange, J. Foerster, J. Clune, D. Ha, The ai scientist: Towards fully automated open-ended scientific discovery, arXiv (2024) 2408.06292.
URL https://arxiv.org/abs/2408.06292

[27] J. Gottweis, W.-H. Weng, A. Daryin, T. Tu, A. Palepu, P. Sirkovic, A. Myaskovsky, F. Weissenberger, K. Rong, R. Tanno, K. Saab, D. Popovici, J. Blum, F. Zhang, K. Chou, A. Hassidim, B. Gokturk, A. Vahdat, P. Kohli, Y. Matias, A. Carroll, K. Kulkarni, N. Tomasev, Y. Guan, V. Dhillon, E. D. Vaishnav, B. Lee, T. R. D. Costa, J. R. Penadés, G. Peltz, Y. Xu, A. Pawlosky, A. Karthikesalingam, V. Natarajan, Towards an ai co-scientist, arXiv (2025) 2502.18864.
URL https://arxiv.org/abs/2502.18864

[28] C. Gao, X. Lan, N. Li, Y. Yuan, J. Ding, Z. Zhou, F. Xu, Y. Li, Large language models empowered agent-based modeling and simulation: a survey and perspectives, Humanities and Social Sciences Communications 11 (1) (2024) 1259. doi:10.1057/s41599-024-03611-3.
URL https://doi.org/10.1057/s41599-024-03611-3

[29] Z. Xu, Z. Wang, S. Li, X. Zhang, P. Lin, Geopredict-llm: Intelligent tunnel advanced geological prediction by reprogramming large language models, Intelligent Geoengineering 1 (1) (2024) 49–57. doi:https://doi.org/10.1016/j.ige.2024.10.005.
URL https://www.sciencedirect.com/science/article/pii/S3050619024000053

[30] Z. Chen, X. Wang, X. Zhang, M. Lin, Y. Liao, J. Li, Y. Bai, Geofactory: an llm performance enhancement framework for geoscience factual and inferential tasks, Big Earth Data 0 (0) (2025) 1–33. doi:10.1080/20964471.2025.2506291.

[31] D. Kim, T. Kim, Y. Kim, Y.-H. Byun, T. S. Yun, A chatgpt-matlab framework for numerical modeling in geotechnical engineering applications, Computers and Geotechnics 169 (2024) 106237. doi:https://doi.org/10.1016/j.compgeo.2024.106237.
URL https://www.sciencedirect.com/science/article/pii/S0266352X24001733

[32] H. Li, C. Shi, Few-shot learning of geological cross-sections from sparse data using large language model, Geodata and AI 2 (2025) 100010. doi:https://doi.org/10.1016/j.geoai.2025.100010.
URL https://www.sciencedirect.com/science/article/pii/S3050483X25000097

[33] S. H. Chai, I. Chen, J. Huang, T. Yacoub, Large Language Model for Geotechnical Engineering Applications Using Retrieval Augmented Generation, 2025, pp. 1–10. arXiv:https://ascelibrary.org/doi/pdf/10.1061/9780784485989.001, doi:10.1061/9780784485989.001.
URL https://ascelibrary.org/doi/abs/10.1061/9780784485989.001

[34] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Küttler, M. Lewis, W.-t. Yih, T. Rocktäschel, S. Riedel, D. Kiela, Retrieval-augmented generation for knowledge-intensive nlp tasks, in: H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, H. Lin (Eds.), Advances in Neural Information Processing Systems, Vol. 33, Curran Associates, Inc., 2020, pp. 9459–9474.

[35] H.-R. Xu, N. Zhang, Z.-Y. Yin, P. Guy Atangana Njock, GeoLLM: A specialized large language model framework for intelligent geotechnical design, Computers and Geotechnics 177 (2025a) 106849. doi:https://doi.org/10.1016/j.compgeo.2024.106849.
URL https://www.sciencedirect.com/science/article/pii/S0266352X24007882

[36] H.-R. Xu, N. Zhang, Z.-Y. Yin, P. G. A. Njock, Multimodal framework integrating multiple large language model agents for intelligent geotechnical design, Automation in Construction 176 (2025b) 106257. doi:https://doi.org/10.1016/j.autcon.2025.106257.

[37] H. Pang, M. K. Lo, A. Y. F. Leung, S. Wu, Reconstruction of landslide events using llm-based agentic ai with multimodal data, SSRN (2025) 5340542.

[38] S. Wu, Y. Otake, D. Mizutani, C. Liu, K. Asano, N. Sato, T. Saito, H. Baba, Y. Fukunaga, Y. Higo, A. Kamura, S. Kodama, M. Metoki, T. Nakamura, Y. Nakazato, A. Shioi, M. Takenobu, K. Tsukioka, R. Yoshikawa, Future-proofing geotechnics workflows: accelerating problem-solving with large language models, Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards (2024) 1–18.

[39] J. Wu, W. Gan, Z. Chen, S. Wan, P. S. Yu, Multimodal large language models: A survey, in: 2023 IEEE International Conference on Big Data (BigData), 2023, pp. 2247–2256. doi:10.1109/BigData59044.2023.10386743.