Taylor & Francis
Taylor & Francis Group

International Journal of Digital Earth

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tjde20

GPT, large language models (LLMs) and generative artificial intelligence (GAI) models in geospatial science: a systematic review

Siqin Wang, Tao Hu, Huang Xiao, Yun Li, Ce Zhang, Huan Ning, Rui Zhu, Zhenlong Li & Xinyue Ye

To cite this article: Siqin Wang, Tao Hu, Huang Xiao, Yun Li, Ce Zhang, Huan Ning, Rui Zhu, Zhenlong Li & Xinyue Ye (2024) GPT, large language models (LLMs) and generative artificial intelligence (GAI) models in geospatial science: a systematic review, International Journal of Digital Earth, 17:1, 2353122, DOI: 10.1080/17538947.2024.2353122

To link to this article: https://doi.org/10.1080/17538947.2024.2353122

Published online: 20 May 2024.

INTERNATIONAL JOURNAL OF DIGITAL EARTH
2024, VOL. 17, NO. 1, 2353122
https://doi.org/10.1080/17538947.2024.2353122

REVIEW ARTICLE

OPEN ACCESS Check for updates

GPT, large language models (LLMs) and generative artificial intelligence (GAI) models in geospatial science: a systematic review

Siqin Wang[a,b,c*], Tao Hu[d*], Huang Xiao[e*], Yun Li[f*], Ce Zhang[g*], Huan Ning[h], Rui Zhu[g], Zhenlong Li[h] and Xinyue Ye[i]

[a]Spatial Sciences Institute, University of Southern California, Los Angeles, USA; [b]School of Environment, The University of Queensland, Brisbane, Queensland, Australia; [c]School of Science, Royal Melbourne Institute of Technology (RMIT), Melbourne, Australia; [d]Department of Geography, Oklahoma State University, Stillwater, Oklahoma, USA ; [e]Department of Environmental Sciences at Emory University, Atlanta, USA; [f]Department of Computer Science, Emory University, Atlanta, USA; [g]School of Geographical Sciences, University of Bristol, Bristol, UK; [h]Geoinformation and Big Data Research Lab, Department of Geography, The Pennsylvania State University, University Park, USA; [i]Department of Landscape Architecture and Urban Planning & Department of Computer Science & Engineering, Texas A&M University, College Station, TX, USA

ABSTRACT

The launch of large language models (LLMs) like ChatGPT in late 2022 and the anticipated arrival of future GPT-x iterations have marked the beginning of the generative artificial intelligence (GAI) era. We conducted a systematic review of how to integrate LLMs including GPT and other GAI models into geospatial science, based on 293 papers obtained from four databases of academic publications – Web of Science (WoS), Scopus, SSRN and arXiv – 26 papers were eventually included for analysis. We statistically outlined the share of domains where LLMs and other GAI models, the type of data that have been used for these models, and the modelling tasks and roles that they play. We also pointed out the challenges and future directions for the next research agenda – along with which we could better position ourselves in the mainstream of science and the cutting-edge research paradigm as others leverage insights from the growing data deluge.

ARTICLE HISTORY
Received 19 December 2023
Accepted 4 May 2024

KEYWORDS
GPT; generative AI (GAI); large language models (LLMs); geospatial science; GIS

1. Introduction

The launch of large language models (LLMs) like ChatGPT (Generative Pretrained Transformer) in later 2022, followed by the rapidly evolving versions such as GPT-3.5 and GPT-4 in early 2023, and the anticipated arrival of future GPT-x iterations, has marked the beginning of the generative artificial intelligence (GAI) era. ChatGPT is developed and trained by a technique called Reinforcement Learning from Human Feedback to enable it to autonomously learn from data and produce sophisticated and conversational writing (Juhász et al. 2023). This type of GAI technology has significant and far-reaching implications for the ways researchers work and teach, while also raising concerns about the nature of these changes (Kumar 2023). As geographers, we are particularly interested in

CONTACT Tao Hu tao.hu@okstate.edu; Department of Geography, Oklahoma State University, Oklahoma, USA
*These authors equally contribute to this paper.

integrating LLMs into geographic education and research. It's crucial for us to embrace the opportunities presented by artificial intelligence (AI), while concurrently tackling the challenges, mitigating the risks, and thinking critically.

We have observed an emerging and promising research field in the literature that integrates GAI with geospatial science to solve problems and issues of geographic nature since ChatGPT was initially launched in November 2022. There is no systematic review in the current scholarship to summarize where and how GPT or LLMs together with other GAI models have been immersed within geospatial science – a research gap that this paper aims to address. In this review, we define GAI as the artificial intelligence capable of generating new and realistic text, programming code, images, or other media that are nearly indistinguishable from human creations or reality (Baidoo-Anu and Ansah 2023) to make the scope of our review specific within the application in the geospatial domain.

As such, we conducted a systematic review based on 293 papers obtained from four databases of academic publications – Web of Science (WoS)[1], Scopus[2], SSRN[3] and arXiv[4] (detailed in Section 2) – 26 papers were eventually included for analysis after manual scanning (See Appendix). We statistically outlined the share of domains where LLMs and GAI models, the type of data that have been used for these models, and the modelling tasks and roles that they play. We also pointed out the challenges and future directions for the next research agenda – along with which we could better position ourselves in the mainstream of science and the cutting-edge research paradigm as others leverage insights from the growing data deluge.

2. Review method

We employed the standard systematic review methodology (Moher, Altman, and Tetzlaff 1996), known as the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) method (Figure 2), to collect, scan, and select appropriate papers within our research scope. We obtained papers relevant to GPT and GAI models used in geospatial science from four databases: WoS and Scopus as two of the most popular academic databases containing peer-reviewed articles, as well as SSRN and arXiv as two highly reputed database repository providing credible sources of research pre-prints/post-prints.

Our search was based on several sets of predefined syntax: (1) paper topics (including article titles, abstracts, and keywords) relevant to large pretrained models, including 'GPT', 'generative artificial intelligence', 'generative AI', 'GAI', 'artificial general intelligence', 'AGI', 'large language model' and 'foundation model'; (2) paper topics relevant to geospatial science, including 'geospatial', 'geograph*' (the asterisk expands the search to include variations of the key syntax, such as geography, geographic, and geographical), 'spatial' OR 'spatialtemporal' and 'spatiotemporal'; (3) for WOS and Scope, article types defined as 'peer-reviewed journal articles' and 'conference preceding papers' while for SSRN and arXiv, article types include 'preprint', 'conference papers' and 'ongoing papers'; (4) publication languages as 'English'; (5) publication timespan as '2022 and 2023' given the ChatGPT built upon GPT 3.5 was launched on Nov 30th, 2022 and become the fastest-growing consumer software application in history – which drive a large amount of publications since then.

With all the aforementioned settings, the initial search results display a total of 293 papers from four databases (detailed in Figure 1). After removing duplicates and excluding papers beyond geospatial science via scanning journal types (e.g. neuroscience and food security), the number of selected papers was reduced to 135 which would need to further filter out via human scanning. Then we defined the eligible criteria for human scanning – regular research papers containing GPT or GAI models, case studies, and demonstrations need to be included; while qualitative papers (e.g. editorial, review, correspondence and option papers) were excluded. Eventually, 26 papers were obtained with attributes including the publication year, author name, article title,

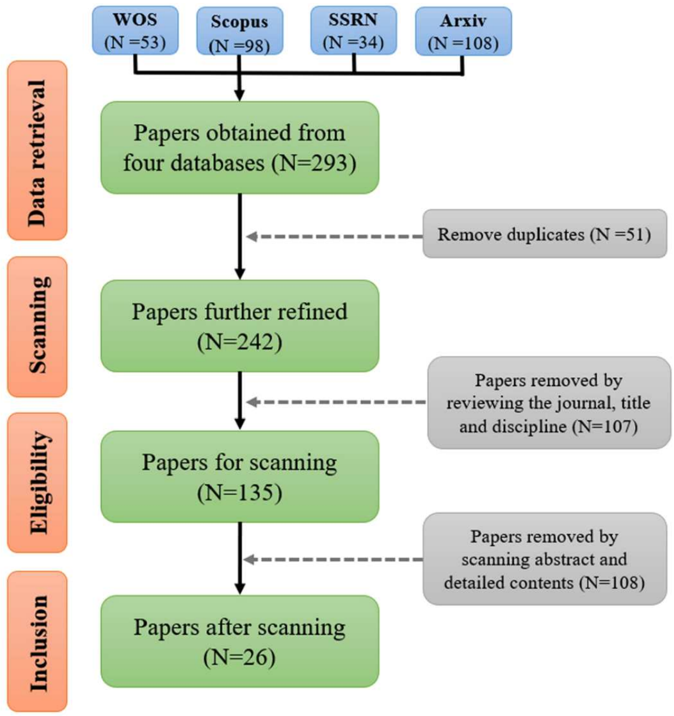

Figure 1. PRISMA workflow for paper collection.

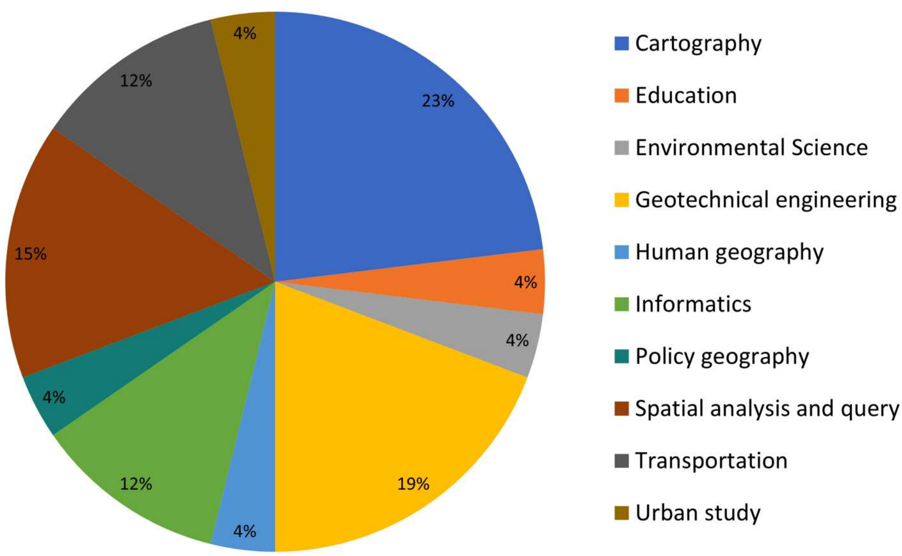

Figure 2. Domains where GPT and other GAI models have been integrated with geospatial science.

journal, keywords, abstract, domain (i.e. data used for modelling, method, modelling task and finding).

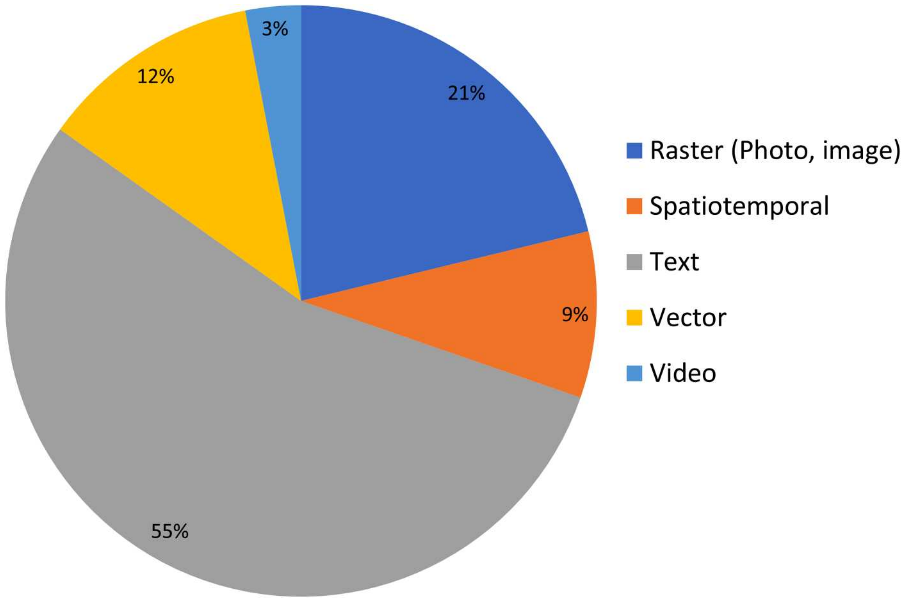

Figure 3. Data type used in geospatially empowered LLMs and GAI models.

3. Current trend of GAI-integrated geospatial studies

3.1. Summary statistics

Figure 2 illustrates the domains in which LLMs and GAI models have been incorporated into geospatial science. The delineation of such domains follows the domain category defined by WOS. Cartography stands out as the primary domain where LLMs and GAI models were used for making maps, accounting for 23% of all relevant publications. It is followed by geotechnical engineering (19%), spatial analysis and query (15%), informatics (12%), and transportation (12%). We can also observe that geospatially empowered LLMs and GAI models have been also used in education, environment science, policy geography, and urban study.

The types of data that have been used in the integration between LLMs and GAI models and geospatial science are categorized in Figure 3. The predominant type is text-based data (e.g. input of tweets, prompts, and image captions), accounting for 55% of all publications. It is followed by raster data (e.g. photos, images and land use maps; accounting for 21%), vector data (e.g. points, polygons and lines; 12%), spatiotemporal data (e.g. human mobility trajectory; 9%) and video (3%).

3.2. The diverse roles that large pre-trained GAI plays

Even before the emergence of GAI, AI algorithms, models, and techniques have been widely used in geospatial science. For example, AI models can help researchers filter or distil needed information from massive data, such as identifying suicide-related tweets (Wang et al. 2023) or marking buildings in remote sensing images (Ning et al. 2020). They also can be used as spatial models to absorb spatiotemporal information to model geographic dynamics for prediction (Yu et al. 2023; Lu et al. 2023; Yu, Masrur, and Blaszczak-Boxe 2023a; Ye et al. 2023). A recent research trend is to implement autonomous agents, such as autonomous GIS, to conduct intelligent tasks without human intervention (Li and Ning 2023). The emerging large pre-trained GAI models have brought more applications. The studies collected in this review show various innovative AI applications in geospatial science. GAI can provide embedded data or generate new data such as images; some are

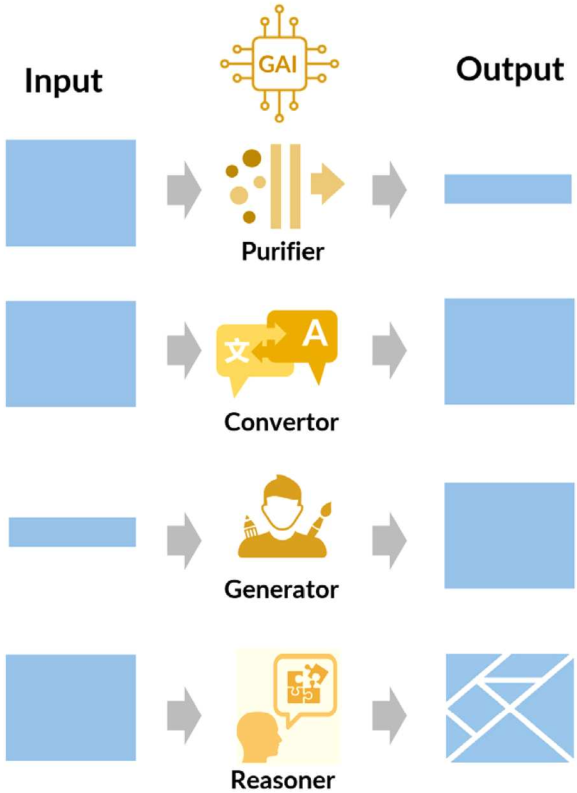

Figure 4. The role of LLMs and other GAI models in geospatial science from the information gain perspective (i.e. how information volume changes). Note: the size of the blue area indicates the information volume.

specialized for a single task, e.g. image segmentation (Kirillov et al. 2023); others are more general, such as GPT-4. To better analyze these applications and understand how GAI may motivate geospatial science's development, an appropriate framework clarifying GAI's roles is needed. Such a framework should incorporate recent fresh GAI while still potentially being compatible with new ones in the future, which may be beyond our current imagination. An appropriate framework is not only an axis where researchers locate their works but also a theoretical framework for systematically investigating possible improvements. Inspired by the innovative GAI applications reflected in the collected articles, we proposed a framework for GAI's roles in geospatial science according to information gain comparing the model output and input: purifier, convertor, generator, and reasoner. Note that we use the information gain, i.e. how the information volume changes, as its literal meaning, which is the volume of attained information or the change of information volume, rather than the specific term in the machine learning algorithm of decision tree (Nowozin 2012).

Figure 4 shows the four roles and their associated information gain. Geospatial science researchers expect that GAI's output contains interesting information only, such as place names in a tweet or land cover map from a remote sensing image; the expected information comes from the input data and training dataset. Based on this consideration, our framework helps geospatial professionals source the information in the GAI-generated output: the expected and sufficient information should come from either the input data or the training dataset. Failing to provide such information, GAI may generate unreliable results.

3.2.1. GAI as a purifier

In this role, GAI will only return information of interest, such as extracting location descriptions from tweets related to flood events (Hu et al. 2023). Most classification or clustering tasks can be considered information purification. GAI will emphasize the most relevant pieces of the input information while ignoring irrelevant pieces. Thus, the output is a part of the input information; the restoration is infeasible because of information loss. For example, a tweet mentioning an ongoing flood event onsite situation can be categorized as 'flood relevant,' then the original tweet details are missing because it is impossible to restore that tweet text using the 'flood relevant' tag or category. Similarly, GPT is good for extracting place names from text; these extracted names are a part of the input, while other information is missing. The GAI users for purification need to focus on amplifying the interested characteristics of the input data to facilitate the GAI extraction. Feature engineering may be a critical step for the GAI as a purifier. Since large GAI models often encrypt massive of information in the training data, users can directly fetch information from GAI by prompts without input data, e.g. gathering longitude and latitude for some cities, as demonstrated in Roberts et al. (2023)'s study probing the internal knowledge of the GPT-4. We still see such applications as purification: the input data is the training data.

3.2.2. GAI as a converter

The input information traverses through the GAI can be converted to other formats, forms, or modals; we expect GAI to keep most information of the input, at least at the aspect we want to emphasize, for example, transforming the interview audio input transcript (Radford et al. 2023), converting satellite images into land cover maps (Shang et al. 2023), or obtaining detailed captions of a street view image (Juhász et al. 2023). These applications demonstrate the transformations or projection of input information in different spaces; we expect such transformations, although usually projection, to keep the main body of the input information. GAI practitioners in geospatial science may need to pay more attention to the data embedding or representation in these information conversion scenarios.

3.2.3. GAI as a generator

GAI can generate new data with little input information or commands, such as creative writing, coding, or generating an image by prompts. In this scenario, GAI can provide more information volume than the input, such as data augmentation (Zheng, Wu, and Li 2023) and 3D reconstruction from a single image (Anciukevičius et al. 2023). The output of GAI largely depends on the model performance and the training data quality and volume because the only source of output information is the training data. Thus, the challenges for the generative application are the variety and sufficiency of the training data and the model capability of keeping the condensed training data information.

3.2.4. GAI as a reasoner

The most exciting progress of recent GAI is that they can reason like humans and explain their reasoning for human using natural language. In this review paper, we adopt a pragmatic rather than a metaphysical viewpoint to use the term 'reason' for LLMs or GAI. It refers to the generated content, such as text, *appearing* to meet human language, cognition, and thinking logic, and being nearly *indistinguishable* from human intelligence or reality.

Taking the flood depth estimation using onsite flood photos as an example, we showcase how to define GAI's role as a reasoner (Akinboyewa et al. 2024). Recall that our framework considers the information gain between the GAI generated out and its input. If the GAI model elaborates the reason for the estimated flood depth, for example, 'The flood inundates the adult male's knee, so I guess the depth is about 0.5 meters.' We think the GAI plays a reasoner role in this scenario. If GAI only provides a sample output as '0.5', indicating the flood depth without given any reason, then its role is purifier because no reasoning is presented to human users. Such purifier role can be replaced by an end-to-end model, which is trained on massive flood photos and the associated flood depth, and it can output a number of '0.5' for the same flooding photo fed into the GAI. Being used as a Reasoner, pre-trained multimodal GAI can be an explainable and universal method to estimate flood depth from geotagged onsite flooding photos. Previous visual approaches are also based on detecting submerged street objects (Chaudhary et al. 2020), e.g. vehicles (Park et al. 2021), bikes, people (Feng, Brenner, and Sester 2020; Meng, Peng, and Huang 2019; Quan et al. 2020), bikes, and stop signs (Kharazi & Behzadan, 2021; Song and Tuo 2021). However, these methods require training datasets and sophisticated algorithms for each submerged object; building such workflows is labor-intensive, uncertain, and less explainable.

In 2020, Janowicz (Janowicz et al. 2020) proposed a moonshot: Can we develop an artificial GIS analyst that passes a domain-specific Turing Test by 2030? This goal is ambitious, but we might achieve it years early using GAI as a reasoner. Currently, many autonomous digital agents adopt GPT as the reasoning core to accomplish tasks without human intervention. Autonomous GIS (Li and Ning 2023), GeoGPT (Zhang et al. 2023b), and MapGPT (Fernandez and Dube 2023) are some pioneer attempts to explore how to leverage GAI to analyze spatial data, provide solutions, and make decisions. Professionals in geospatial science and other domains have been waiting for the reasoning core for a long time. They expect GAI to develop reasonings, insights, and decisions according to input information; the information volume of output is usually less than the input, but the gain is less important compared with the thoughts in the output and the reason for those thoughts. Current GAI models obtained reasoning ability via massive text and code (Delétang et al. 2023; Yu et al. 2023; Zheng et al. 2023), while GAI users may have difficulty finetuning the reasoning ability; thus, they need to provide necessary and sufficient information in the input for GAI's further reasoning.

3.2.5. Implication for GAI applications from the information gain perspective

Our framework aims to identify the role of GAI, rather than the type of GAI; the implication for researchers is to remind them to trace back the expected information in the GAI output. Table 1 is an implication summary. The 'input data' in the table refers to the information (e.g. text, image, and audio), and the 'training data' means the data used for the model training, i.e. determining the model weights. Note that we do not consider the examples in the input data ('few-shot') as the training data since those examples aim to provide sufficient information rather than to determine or change the model weights.

The proposed framework can cover most collected GAI papers, but has no categories for those used GAI as numerical models, such as the population rank interpretation (Manvi et al. 2023) and spatial trajectory abnormal detections (Zhang et al. 2023). These applications try to use GAI as the traditional spatial algorithms or models. Although obtaining competitive results, GAI did not receive sufficient information to output accurate ranks. GAI may be a compressed representation of the training set (Delétang et al. 2023; Yu et al. 2023), thus may not be good at numerical analysis. Ji and Gao (2023) conducted experiment on GPT-2 and BERT also showed evidence of the GAI's weakness in numeric processing. They encode the well-known text format of geometries and then

Table 1. The implications for GAI applications in geospatial science from the information gain perspective.

Role	Information source in the GAI output	Implication
Purifier	Input data	Feature engineering may be helpful since it can amplify or salientize the interested characteristics of the input data to facilitate the GAI extraction.
Converter	Input data	Adopting appropriate embeddings or representations for input and output data.
Generator	Training data	the variety and sufficiency of the training data and the model capability
Reasoner	Input data, training data (especially text and code)	Providing necessary and sufficient information in the input for GAI's further reasoning.

feed embeddings into the classifier and regressor to restore the geometry area, centroid, and distance. The unsatisfied results revealed that the encoding techniques from GPT-2 and BERT have difficulty maintaining the numeric information, leading to unreliable further processing. Horikomi et al. (2023) converted people's daily trajectories into map grids and time slots represented by characters to avoid numeric input, reducing uncertainty for trajectory prediction. They trained GPT-2 from scratch then obtained better results than Markov chains. This is an interesting test to convert numeric problems to character prediction while adopting cutting-edge natural language processing embedding and models. Geospatial research could further explore such ideas to formulate research questions as sequence problems then try proofed language models such as transformer (Vaswani et al. 2017). For example, converting the 3D object generation to the mesh sequence prediction (Nash et al. 2020).

On the contrary, Juhász et al. (2023) fed necessary and sufficient information into GPT to generate street view image tags (e.g. road types). The researchers obtained formatted image captions by BLIP-2 (Lu et al. 2023) and then put the caption into GPT-3.5-turbo to determine tags. In this application, BLIP-2 served as a purifier to extract the information of interest (e.g. lighting, lane number, and road surface type), and the GPT served as another purifier to determine tags according to the information extracted. Collectively, we argue that GAI performs information purification, conversion, generation, or reasoning better than numeric analysis. The accurate numeric results cannot be generated merely based on the insufficient input information and the internal information embedded in the GAI model. GAI users may need to have a good sense of tracing back the source of the desired output information to coordinate various GAI models better to finish complex tasks.

4. Discussion: future directions and challenges

We summarize the challenges, issues, and future directions of implementing LLMs and GAI models in the current 26 papers that we reviewed. Such challenges and issues are articulated as below in four perspectives, further linking to the future directions.

4.1. Future GAI functions to be integrated with geospatial science

GAI models are demonstrating their worth in geospatial science, particularly for tasks such as extracting location data and conducting semantic analysis. Despite their promise, the full potential of their application in geospatial data management remains untapped. They encounter issues like reliability and inconsistent responses. This necessitates a specialized framework for reasonable evaluation in geospatial settings. Furthermore, GAI models' black-box nature in predictive tasks like time series forecasting poses challenges in understanding and trusting their decision-making processes.

4.1.1. Enhancing geospatial data query and recommendation

The advanced language understanding capabilities of LLMs as the subset of GAI have established them as potent semantic tools, offering significant aid in tasks related to geospatial information, such as gazetteer recognition, and semantic query understanding. This utility is evidenced by previous studies, like Hu et al. (2023), which have effectively utilized LLMs to extract location descriptions from social media messages. However, the application of tools like ChatGPT in the discovery and analysis of geospatial data remains underexplored. To facilitate this, a LLM-based recommendation engine for geospatial data discovery can be utilized (Jiang et al. 2018). For example, OpenAI's LLMs provides the embedding API to obtain the embedding vectors of metadata using their description information as shown in Figure 5. By measuring the distance between two metadata embedding vectors, the relatedness between each pair of metadata can be evaluated. Shorter distances represent greater similarity between metadata. In a metadata-based recommender, the embedding of the metadata of data a user views will be compared to the embedding vectors of

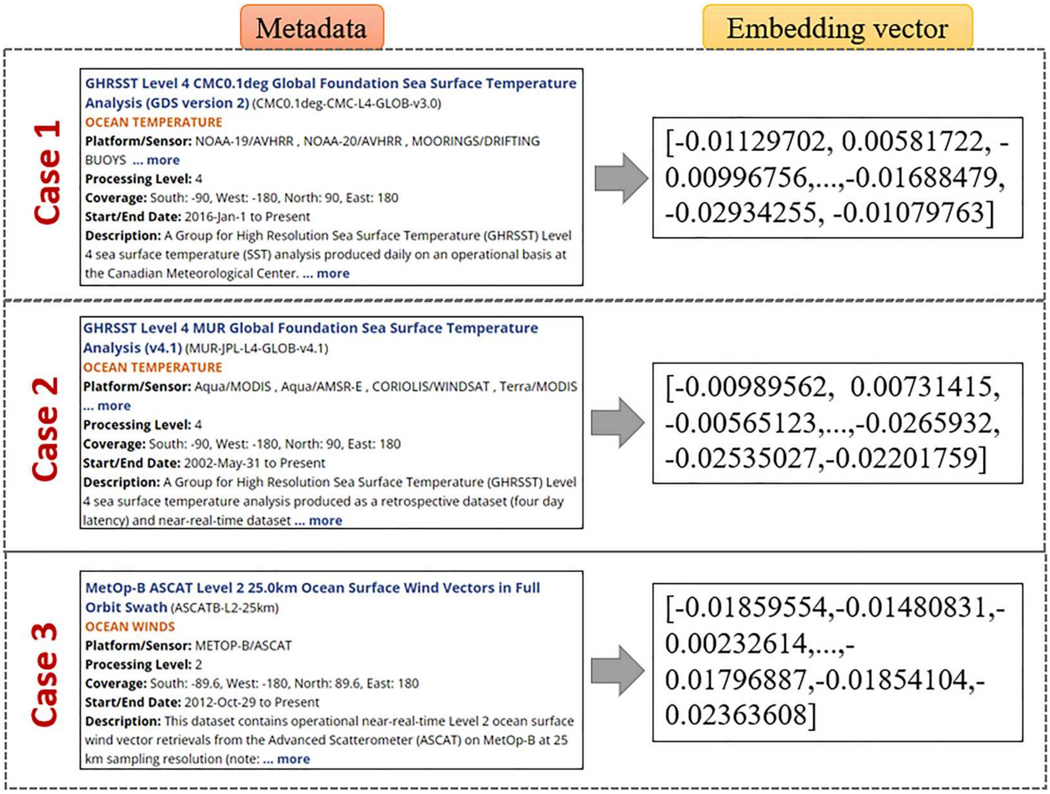

Figure 5. Embedding vectors computed by OpenAI API for the metadata in geospatial science.

all other metadata. This comparison will result in a list of the most similar data sets that can be recommended to the user. The metadata-based similarity measure can also be combined with embedding calculated from user queries and insights extracted from user behaviours data to make better recommendations. Our demonstrations showcase the strength of integrating OpenAI APIs in geospatial science applications, particularly for tasks related to natural language processing.

4.1.2. Building a robust LLM-based assessment framework

In the rapidly evolving domain of geospatial science, the integration of LLMs and GAI models presents distinct challenges that necessitate a specialized assessment framework. While these models exhibit a remarkable capacity for processing and generating complex information, their limitations in terms of reliability and response inconsistency become particularly critical in geospatial contexts. Given the iterative nature and version evolution of tools like ChatGPT, a unified LLMs assessment framework becomes imperative for systematic evaluation and validation within geospatial science. The criticality of precision in geospatial science such as locations and spatial relationships, underscores the need for a framework that rigorously assesses and ensures model accuracy. ChatGPT's current limitations, manifesting as occasional unreliable or inaccurate responses (Johnson et al. 2023), could lead to significant errors and unreliable outcomes in geospatial analyses. To mitigate these risks, a potential solution is to leverage the generative capabilities of large language models in combination with the robustness of existing databases and knowledge bases (Feng, Ding, and Xiao 2023). Meanwhile, a robust validation framework tailored for geospatial applications would enable systematic testing, aligning the model's output with established geospatial standards and factual accuracy.

Moreover, ChatGPT's tendency to generate varying responses to identical queries poses a unique challenge in maintaining consistency and coherence in geospatial analysis. For example, Li and Ning recently leveraged the ChatGPT's general abilities in natural language understanding, reasoning, and coding to build an autonomous geographical information system, but the success rate is only 80% when running the codes developed by ChatGPT (Li and Ning 2023). In a field where consistency and predictability are essential, an assessment framework designed for geospatial science is crucial. It should not only identify and evaluate patterns of inconsistency but also assess the model's ability to provide stable, reliable responses to geospatial queries. Such a framework is vital for aligning the model's performance with the stringent requirements of geospatial applications, ensuring that its contributions are both scientifically sound and practically reliable.

However, the potential limitations of LLMs worth attention. LLM-based predictions are not interpretable and explainable. Some studies turn to LLMs for conducting predictions due to its impressive capabilities. For example, Mai et al leveraged GPTs to conduct time series forecasting on the number of deaths by simply inputting historical data and predicting future data (Mai et al. 2023). However, it's essential to acknowledge that GPT and similar tools often operate as a complete black box when employed for predictive tasks. This means that while it can provide accurate predictions, the inner workings of the model remain largely hidden, making it challenging to fully comprehend how it arrives at its conclusions. The absence of a clear, interpretable rationale behind GPT and similar tools predictions can be a significant concern, particularly when stakeholders, policymakers, or end-users seek to understand the basis for these forecasts. It becomes difficult to provide meaningful insights or explanations about the factors or variables that contribute to the model's predictions, hindering its utility in decision-making processes.

4.2. Multimodal foundation models to develop highly effective GeoAI

GAI models represent the state-of-the-art development of deep learning, essentially a form of feature learning and feature representation, where relevant feature parameters are discovered automatically through deep neural networks (Sarker 2021). In both domains of natural language processing and computer vision, the existence of minimum feature granularity (e.g. letters: shapes, pronunciation, semantics; pixels: size and grayscale) allows the fusion of these features to form higher-level features (e.g. words, sentences, paragraphs, and contexts) (Iluz et al. 2023). In contrast to single-modal applications, geospatial science exhibits distinctive characteristics of multimodality: From the perspective of feature representation, it combines spatial, temporal, and semantic feature spaces. In terms of spatial and temporal dimensions, it encompasses two-dimensional (2D), three-dimensional (3D), and even four-dimensional (4D) representations of spatial and/or temporal objects; regarding scales, it involves spatial, temporal, and semantic dimensions, requiring their comprehensive integration and generalization across multiple scales. The multi-modality of geospatial science presents huge opportunities for multimodal foundation models to develop highly effective GeoAI, illustrated in the below two perspectives.

4.2.1. Feasibility of cross-scale geospatial science research

Scale is an inherent property of geographical information, and variations in scales can affect the level-of-details of geographic information about observation, analysis, representation, and exchange (Goodchild 1992). Traditional research is limited to specific scales due to constraints in information collection, data processing, and information expression. However, the development of technologies such as satellite remote sensing, LiDAR scanning, IoT sensors, etc., and the incorporation of video, audio, mobile signals, and other sensor information into the scope of geoscientific research, it is possible to further break constraints on temporal and spatial scales. Leveraging the analytical processing capabilities of large multimodal models supported by massive, multimodal data, it becomes feasible to conduct cross-scale, and even full-scale, geospatial research, thus transforming geospatial research and applications. For example, besides traditional (relatively coarse scale) geospatial data (e.g. satellite images, arial images, elevation models, and field investigation), other (spatiotemporal fine scale) data including various audios and videos, in-situ photos, wearable biosensor data, as well as various environmental sensing data, can all be applied for habitat and wildlife monitoring using multimodal models. The adoption of multimodal model can thus push the limits of scale crossing by aligning and mining the diverse scale spatiotemporal information, which would expand our spatiotemporal cognition in the study of habitat and animal behaviour monitoring.

4.2.2. Next-generation autonomous geospatial science

Researchers have already started utilizing large language models to explore research on autonomous, AI-powered geospatial science. Existing LLMs have effectively addressed challenges related to semantic reasoning in geoscientific research, laying a solid foundation for the fusion and transformation of geometric and semantic expressions (Li and Ning 2023). Multimodal models will further provide autonomous geospatial science with intelligent and high-effective solutions for the collection, processing, analysis and visualization of six multimodal data (e.g. text, image, audio, video, IMU data) relevant to semantic, spatial, and temporal information (Girdhar et al. 2023).

However, inherent challenges exist in applying multimodal models to geospatial science. First, it lacks massive, unstructured samples required for multimodal training. Currently, existing geospatial science primarily utilizes standardized, structured spatial data (vector data) produced by human intelligence, reflecting information that has undergone intelligent extraction or recognition by humans, without the information of minimum feature granularities that are suitable for deep learning models. It is therefore especially challenging for current deep learning models to use vector data (Mai et al. 2023). This falls far short of large-scale geospatial data samples required for deep learning model training. Therefore, solving the issue of lacking massive samples is a priority to achieve large data training multimodal models. Second, another challenge exists in the integration and innovation between geospatial science and AI. Existing large language models (deep learning) are mainly data-driven, establishing general paradigms through extensive sample training. When applying such models to geospatial science, challenges arise, including geographical fidelity, geographical bias, temporal bias, spatial scales, generality, and heterogeneity (Huang et al. 2023). Similar to ecology (Han et al. 2023), geospatial research involves substantial prior knowledge. Integrating knowledge based on traditional GIS theories into hierarchical features of deep learning or realizing artificial intelligence guided by knowledge poses a challenge and is one of the future development challenges for GeoAI.

4.3. From GeoAI to GeoAGI (artificial general intelligence): a promising way to go

4.3.1. Bridging human cognition and geospatial intelligence

The ascent of AGI marks an epochal transformation in the trajectory of machine intelligence, edging us closer to a future where machines are expected to perform with intellectual capacities paralleling those of humans. AGI is anticipated to emerge as a technological marvel, embodying the quintessential elements of human cognition – language comprehension, sophisticated reasoning, and a wellspring of creativity. This marks a profound departure from the capabilities of current AI systems like GPT, which, despite their impressive language processing capabilities, remain in the realm of ‘narrow AI’. Such systems specialize in specific tasks but do not possess the versatility required for complex cognitive functions that necessitate not just processing information but also forming connections, hypothesizing, and making judgments – functions that are intrinsic to human problem-solving and reasoning. The promise of AGI lies in its potential to transcend these limitations, offering solutions that are not bound by the constraints of predefined algorithms or limited scope of understanding.

In the field of geospatial intelligence, the delineation between narrow AI and AGI is particularly pronounced. While models like GPT show a remarkable ability to generate text that closely resembles human writing, they currently do not appear to possess the inherent capability to fully understand and address geospatial issues – these are complex tasks that would benefit from the more comprehensive cognitive abilities anticipated in AGI systems. This realization has spurred the emergence of GeoAGI, a cutting-edge interdisciplinary field that seeks to blend the rigor of artificial intelligence with the precision of geospatial analytics and the adaptability of machine learning. The ambition of GeoAGI is to forge systems that are not just capable of digesting vast amounts of spatial data but can also make intelligent decisions based on that data. Such systems would not only interpret geographical information but also anticipate trends, adapt to new spatial patterns, and interact with human users in a collaborative and intuitive manner.

4.3.2. Pioneering intelligent disaster management

GeoAI, to its credit, has significantly advanced the automation and enhancement of tasks, especially through the adoption of algorithms that are cognizant of spatial relationships. GeoAGI seeks to push the envelope further, aiming not only to automate processes but to pioneer new methods of spatial analysis and insight generation. This is where its transformative potential comes to the fore. For instance, in the context of flood management, AGI could act as an intelligent agent that does much more than simply processing data – it could provide a holistic response to disaster scenarios. By integrating language understanding capabilities similar to GPT models with spatial reasoning, a Geo-AGI system could efficiently identify flood-affected areas, estimate water levels from onsite social media photos, and delineate the extent of urban inundation with a high degree of precision. Furthermore, it could leverage this information to formulate and execute response strategies, such as the rapid deployment of emergency services, the strategic allocation of resources, and the management of evacuation procedures, all in real-time and with minimal human intervention. Such an AGI system could be the linchpin for a paradigm shift in how we manage disasters, propelling us toward a future where intelligent systems work alongside humans to mitigate risks and enhance resilience.

As we near the advent of the GeoAGI era, the imperative for targeted research to refine machine learning algorithms becomes apparent – these algorithms must adeptly parse and utilize geospatial data. The transition from GeoAI to GeoAGI transcends technical innovation and ventures into ethical territory, necessitating a thoughtful examination of its societal impacts, including workforce disruption and human autonomy. Establishing frameworks for the responsible development and application of AGI is essential to align these systems with societal norms and ensure their contribution to the public good. The initiation of the GeoAGI era signifies a seminal advancement in technology and a substantive shift in human engagement. This evolution compels the strategic utilization of AI, guiding AGI development towards outcomes that are advantageous and ethically responsible. The endeavor towards GeoAGI, with its considerable potential for innovative breakthroughs, is poised to augment our capacity for geospatial analysis and decision-making. Such a transformative progression is expected to fundamentally alter our engagement with global geography, culminating in a more profound and intricate understanding of and interaction with the spatial facets of our world.

4.4. Ethnical and privacy concerns: old problems in a new disguise

Researchers assert that 'GPT can improve equity in science' (Berdejo-Espinola and Amano 2023, 991) due to the diminished language barrier for non-native English researchers and students. While embracing this diminished inequity, we have to be aware of the ethnical considerations and privacy concerns of using GPT and other GAI models. First, intellectual property violations become a pressing concern, as the AI system might inadvertently generate content that closely resembles existing copyrighted works (Chang and Kidman 2023). Researchers need to acknowledge the geospatial content generated by GPT in publications and declare the use of AI tools for full transparency and responsibility. Second, it has been reported that AI-generated maps are possible to provide inaccurate or misleading information, possess unanticipated characteristics, and lack reproducibility (Kang, Zhang, and Roth 2023). It is the duty of cartographers, geographers, and geospatial scientists to cautiously bear in mind these concerns when developing AI-generated maps to alleviate the potential negative effects and ensure ethical and responsible usage. Third, as GAI models generates geospatial information based on the input it receives, it may inadvertently reveal sensitive information or be illegally used to track and profile individuals. This is particularly possible in geographic research where personal Internet Protocol (IP) address or locational information, if illegally released or hacked, can be used to track individual trajectories. Thus, it is crucial for both developers and users to have robust privacy and security measures in place to protect data and prevent misuse.

5. Conclusion

Our review work summarizes the latest trend in integrating GPT and GAI models with geospatial science research based on publications obtained from four academic databases – WoS, Scopus, SSRN and arXiv. We statistically outline the domains where GPT and GAI models have been employed in geospatial science, data and modelling task that have been implemented, and the role that GPT and GAI models play. Key findings include that LLMs and GAI models have been mainly used for making maps, geotechnical and geospatial engineering, spatial analysis and query, informatics and transportation. The types of data that have been used in the integration between LLMs and GAI models and geospatial science include text, raster and vector data as the top three data sources. The roles that LLMs and GAI models played in the existing studies can be categorized as a purifier, convertor, generator, and reasoner. We also point out the challenges and future directions for the next research agenda, including (1) to integrate future GAI functions with geospatial science through enhancing geospatial data query and recommendation as well as building a robust LLM-based assessment framework; (2) to develop multimodal foundation models to improve highly effective GeoAI through promoting the feasibility of cross-scale geospatial science research and next-generation autonomous geospatial science; (3) to advance from GeoAI to GeoAGI (artificial general intelligence) through bridging human cognition and geospatial intelligence and pioneering intelligent disaster management. In the end, we reiterate the ethnical and privacy concerns embedded in the domain – which needs wide attention and regulation at the system level.

Along with these future directions, we could better position ourselves in the mainstream of science and the cutting-edge research paradigm as others leverage insights from the growing data deluge. As geospatial researchers, we have a responsibility to ensure the honest and responsible use of LLMs (including GPT) and GAI models in geospatial research by embracing the benefits of GAI while judiciously addressing issues related to bias, provenance, and inaccuracies specific to research practices. We call for future efforts to incorporate the nature of geospatial science into GAI models and to leverage their strength to benefit geospatial research and education – starting from the road map we outlined here to make geospatial work one of the most visual, interactive, and exploratory endeavors in the age of AI.

Notes

1. https://www.webofscience.com/wos/
2. https://www.scopus.com/search/form.uri?display=basic#basic
3. https://www.ssrn.com/index.cfm/en/
4. https://arxiv.org/

Acknowledgements

All the contents in this paper were originally written by authors.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This study was funded by the star-up fund from Oklahoma State University.

Data availability statement

Data were not used, nor created for this research.

Competing interests

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

Authors' contributions

Siqin Wang: Data curation, investigation, formal analysis, visualization, software, writing – original draft, writing – review & editing, project administration.
Tao Hu: Data curation, investigation, Funding acquisition, Writing – original draft, Writing – review & editing.
Xiao Huang: Data curation, investigation, Writing – original draft, Writing – review & editing.
Yun Li: Data curation, investigation, Writing – original draft, visualization.
Ce Zhang: Data curation, investigation, Writing – original draft, Writing – review & editing.
Huan Ning: Writing – original draft, Writing – review & editing.
Rui Zhu: Writing – original draft, Writing – review & editing.
Zhenlong Li – Writing – review & editing.
Xinyue Ye: Writing – review & editing.

References

Aghzal, M., E. Plaku, and Z. Yao. 2023. Can large language models be good path planners? a benchmark and investigation on spatial-temporal reasoning. arXiv preprint arXiv:2310.03249.

Akinboyewa, T., H. Ning, M. N. Lessani, and Z. Li. 2024, February 26. Automated Floodwater Depth Estimation Using Large Multimodal Model for Rapid Flood Mapping. https://doi.org/10.48550/arXiv.2402.16684

Anciukevičius, T., Z. Xu, M. Fisher, P. Henderson, H. Bilen, N. J. Mitra, and P. Guerrero. 2023. "Renderdiffusion: Image Diffusion for 3d Reconstruction, Inpainting and Generation." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 12608–12618.

Baidoo-Anu, D., and L. O. Ansah. 2023. "Education in the era of Generative Artificial Intelligence (AI): Understanding the Potential Benefits of ChatGPT in Promoting Teaching and Learning." *Journal of AI* 7 (1): 52–62. https://doi.org/10.61969/jai.1337500.

Berdejo-Espinola, V., and T. Amano. 2023. "AI Tools Can Improve Equity in Science." *Science* 379 (6636): 991–991. https://doi.org/10.1126/science.adg9714.

Chang, C. H., and G. Kidman. 2023. "The Rise of Generative Artificial Intelligence (AI) Language Models-Challenges and Opportunities for Geographical and Environmental Education." *International Research in Geographical and Environmental Education*, 1–5.

Chaudhary, P., S. D'Aronco, J. P. Leitão, K. Schindler, and J. D. Wegner. 2020. "Water Level Prediction from Social Media Images with a Multi-Task Ranking Approach." *ISPRS Journal of Photogrammetry and Remote Sensing* 167: 252–262. https://doi.org/10.1016/j.isprsjprs.2020.07.003.

Dao, X. Q., and N. B. Le. 2023. Chatgpt is good but bing chat is better for Vietnamese students. arXiv preprint arXiv:2307.08272.

Delétang, G., A. Ruoss, P.-A. Duquenne, E. Catt, T. Genewein, C. Mattern, … J. Veness. 2023, September 19. *Language Modeling Is Compression*. https://doi.org/10.48550/arXiv.2309.10668

Feng, Y., C. Brenner, and M. Sester. 2020. "Flood Severity Mapping from Volunteered Geographic Information by Interpreting Water Level from Images Containing People: A Case Study of Hurricane Harvey." *ISPRS Journal of Photogrammetry and Remote Sensing* 169: 301–319. https://doi.org/10.1016/j.isprsjprs.2020.09.011.

Feng, Y., L. Ding, and G. Xiao. 2023. "GeoQAMap-Geographic Question Answering with Maps Leveraging LLM and Open Knowledge Base (Short Paper)." In *12th International Conference on Geographic Information Science (GIScience 2023)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik.

Fernandez, A., and S. Dube. 2023. *Core Building Blocks: Next Gen Geo Spatial GPT Application* (arXiv:2310.11029). https://doi.org/10.48550/arXiv.2310.11029

Girdhar, R., A. El-Nouby, Z. Liu, M. Singh, K. V. Alwala, A. Joulin, and I. Misra. 2023. "Imagebind: One Embedding Space to Bind Them All." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 15180–15190.

Goodchild, M. F. 1992. Geographical Information Science." *International Journal of Geographical Information Systems* 6 (1): 31–45. https://doi.org/10.1080/02693799208901893.

Han, B. A., K. R. Varshney, S. LaDeau, A. Subramaniam, K. C. Weathers, and J. Zwart. 2023. "A Synergistic Future for AI and Ecology." *Proceedings of the National Academy of Sciences* 120 (38): e2220283120. https://doi.org/10.1073/pnas.2220283120.

Horikomi, T., S. Fujimoto, A. Ishikawa, and T. Mizuno. 2023. *Generating Individual Trajectories Using GPT-2 Trained from Scratch on Encoded Spatiotemporal Data* (arXiv:2308.07940). https://doi.org/10.48550/arXiv.2308.07940

Hu, Y., G. Mai, C. Cundy, K. Choi, N. Lao, W. Liu, G. Lakhanpal, R. Z. Zhou, and K. Joseph. 2023. "Geo-Knowledge-Guided GPT Models Improve the Extraction of Location Descriptions from Disaster-Related Social Media Messages." *International Journal of Geographical Information Science* 37 (11): 2289–2318. https://doi.org/10.1080/13658816.2023.2266495.

Huang, C., O. Mees, A. Zeng, and W. Burgard. 2023, May. "Visual Language Maps for Robot Navigation." In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, 10608–10615. IEEE.

Iluz, S., Y. Vinker, A. Hertz, D. Berio, D. Cohen-Or, and A. Shamir. 2023. Word-as-Image for Semantic Typography. *arXiv:2303.01818*.

Jang, K. M., J. Chen, Y. Kang, J. Kim, J. Lee, and F. Duarte. 2023. Understanding Place Identity with Generative AI. *arXiv:2306.04662*.

Janowicz, K., S. Gao, G. McKenzie, Y. Hu, and B. Bhaduri. 2020. "GeoAI: Spatially Explicit Artificial Intelligence Techniques for Geographic Knowledge Discovery and Beyond." *International Journal of Geographical Information Science* 34 (4): 625–636. https://doi.org/10.1080/13658816.2019.1684500.

Jha, A., A. Davani, C. K. Reddy, S. Dave, V. Prabhakaran, and S. Dev. 2023. SeeGULL: A stereotype benchmark with broad geo-cultural coverage leveraging generative models. arXiv preprint arXiv:2305.11840.

Ji, Y., and S. Gao. 2023. *Evaluating the Effectiveness of Large Language Models in Representing Textual Descriptions of Geometry and Spatial Relations* (arXiv:2307.03678). http://arxiv.org/abs/2307.03678.

Jiang, Y., Y. Li, C. Yang, F. Hu, E. M. Armstrong, T. Huang, D. Moroni, L. J. McGibbney, and C. J. Finch. 2018. "Towards Intelligent Geospatial Data Discovery: A Machine Learning Framework for Search Ranking." *International Journal of Digital Earth* 11 (9): 956–971. https://doi.org/10.1080/17538947.2017.1371255.

Johnson, D., R. Goodman, J. Patrinely, C. Stone, E. Zimmerman, R. Donald, … L. Wheless. 2023. "Assessing the Accuracy and Reliability of AI-Generated Medical Responses: An Evaluation of the Chat-GPT Model." *Research Square*, https://doi.org/10.21203/rs.3.rs-2566942/v1.

Juhász, L., P. Mooney, H. H. Hochmair, and B. Guan. 2023. "ChatGPT as a Mapping Assistant: A Novel Method to Enrich Maps with Generative AI and Content Derived from Street-Level Photographs." *Spatial Data Science Symposium 2023*, https://doi.org/10.25436/E2ZW27.

Kang, Y., Q. Zhang, and R. Roth. 2023. The Ethics of AI-Generated Maps: A Study of DALLE 2 and Implications for Cartography. *arXiv:2304.10743*.

Kharazi, B. A., and A. H. Behzadan. 2021. "Flood depth mapping in street photos with image processing and deep neural networks." *Computers, Environment and Urban Systems* 88: 101628.

Kharazi, Alizadeh, B. Behzadan, and A. H. 2021. "Flood Depth Mapping in Street Photos with Image Processing and Deep Neural Networks." *Computers, Environment and Urban Systems* 88: 101628. https://doi.org/10.1016/j.compenvurbsys.2021.101628.

Kirillov, A., E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, et al. 2023. *Segment Anything* (arXiv:2304.02643). https://doi.org/10.48550/arXiv.2304.02643

Kumar, K. 2023. Geotechnical Parrot Tales (GPT): Harnessing Large Language Models in Geotechnical Engineering. arXiv-2304.

Li, J., R. Cai, Y. Tan, H. Zhou, A.-M. Sadick, W. Shou, and X. Wang. 2023a. "Automatic Detection of Actual Water Depth of Urban Floods from Social Media Images." *Measurement* 216: 112891. https://doi.org/10.1016/j.measurement.2023.112891.

Li, J., D. Li, S. Savarese, and S. Hoi. 2023b. *BLIP-2: Bootstrapping Language-Image Pre-Training with Frozen Image Encoders and Large Language Models* (arXiv:2301.12597). https://doi.org/10.48550/arXiv.2301.12597

Li, Z., and H. Ning. 2023. “Autonomous GIS: The Next-Generation AI-Powered GIS.” *International Journal of Digital Earth* 16 (2): 4668–4686. https://doi.org/10.1080/17538947.2023.2278895.

Ling, G., X. Mu, C. Wang, and A. Xu. 2023. “Enhancing Chinese Address Parsing in Low-Resource Scenarios Through In-Context Learning.” *ISPRS International Journal of Geo-Information* 12 (7): 296. https://doi.org/10.3390/ijgi12070296.

Lu, M., Y. Li, M. Yu, Q. Zhang, Y. Zhang, B. Liu, and M. Wang. 2023. “Spatiotemporal Prediction of Radar Echoes Based on ConvLSTM and Multisource Data.” *Remote Sensing* 15 (5): 1279. https://doi.org/10.3390/rs15051279.

Mai, G., W. Huang, J. Sun, S. Song, D. Mishra, N. Liu, … N. Lao. 2023. On the Opportunities and Challenges of Foundation Models for Geospatial Artificial Intelligence. arXiv:2304.06798.

Manvi, R., S. Khanna, G. Mai, M. Burke, D. Lobell, and S. Ermon. 2023. GeoLLM: Extracting Geospatial Knowledge from Large Language Models (arXiv:2310.06213). https://doi.org/10.48550/arXiv.2310.06213

Manvi, R., S. Khanna, G. Mai, M. Burke, D. Lobell, and S. Ermon. 2023. Geollm: Extracting geospatial knowledge from large language models. arXiv preprint arXiv:2310.06213.

Meng, Z., B. Peng, and Q. Huang. 2019. “Flood Depth Estimation from Web Images.” In *Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Advances on Resilient and Intelligent Cities*, 37–40. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3356395.3365542

Moher, D., D. G. Altman, and J. Tetzlaff. 1996. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses). Guidelines for Reporting Health Research: A User’s Manual, 1999, 250.

Nash, C., Y. Ganin, S. M. A. Eslami, and P. W. Battaglia. 2020, February 23. PolyGen: An Autoregressive Generative Model of 3D Meshes. https://doi.org/10.48550/arXiv.2002.10880

Ning, H., Z. Li, C. Wang, and L. Yang. 2020. “Choosing an Appropriate Training set Size When Using Existing Data to Train Neural Networks for Land Cover Segmentation.” *Annals of GIS* 26 (4): 329–342. https://doi.org/10.1080/19475683.2020.1803402.

Nowozin, S. 2012. Improved Information Gain Estimates for Decision Tree Induction. *arXiv:1206*.4620.

Park, S., F. Baek, J. Sohn, and H. Kim. 2021. “Computer Vision–Based Estimation of Flood Depth in Flooded-Vehicle Images.” *Journal of Computing in Civil Engineering* 35 (2): 4020072. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000956.

Quan, K.-A. C., V.-T. Nguyen, T.-C. Nguyen, T. V. Nguyen, and M.-T. Tran. 2020. “Flood Level Prediction via Human Pose Estimation from Social Media Images.” In *Proceedings of the 2020 International Conference on Multimedia Retrieval*, 479–485. New York, NY, USA: Association for Computing Machinery.

Radford, A., J. W. Kim, T. Xu, G. Brockman, C. McLeavey, and I. Sutskever. 2023, July. “Robust Speech Recognition via Large-Scale Weak Supervision.” In *International Conference on Machine Learning*, 28492–28518. PMLR.

Roberts, J., T. Lüddecke, S. Das, K. Han, and S. Albanie. 2023. GPT4GEO: How a Language Model Sees the World’s Geography (arXiv:2306.00020). https://doi.org/10.48550/arXiv.2306.00020

Sarker, I. H. 2021. “Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions.” *SN Computer Science* 2 (6): 420. https://doi.org/10.1007/s42979-021-00815-1.

Shang, K., L. Xu, X. Liu, Z. Yin, Z. Liu, X. Li, … W. Zheng. 2023. “Study of Urban Heat Island Effect in Hangzhou Metropolitan Area Based on SW-TES Algorithm and Image Dichotomous Model.” *Sage Open* 13 (4): 21582440231208851. https://doi.org/10.1177/21582440231208851.

Shuklin, A., D. Parygin, A. Gurtyakov, O. Savina, and N. Rashevskiy. 2022, October. “Synthetic News as a Tool for Evaluating Urban Area Development Policies.” In *2022 International Conference on Engineering and Emerging Technologies (ICEET)*, 1–6. IEEE.

Song, Z., and Y. Tuo. 2021. “Automated Flood Depth Estimates from Online Traffic Sign Images: Explorations of a Convolutional Neural Network-Based Method.” *Sensors* 21 (16): 5614. https://doi.org/10.3390/s21165614.

Stoewer, P., C. Schlieker, A. Schilling, C. Metzner, A. Maier, and P. Krauss. 2022. “Neural Network Based Successor Representations to Form Cognitive Maps of Space and Language.” *Scientific Reports* 12 (1): 11233. https://doi.org/10.1038/s41598-022-14916-1.

Tao, R., and J. Xu. 2023. “Mapping with Chatgpt.” *ISPRS International Journal of Geo-Information* 12 (7): 284. https://doi.org/10.3390/ijgi12070284.

Vaswani, A., N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, … I. Polosukhin. 2017. “Attention is all you Need.” *Advances in Neural Information Processing Systems*, 5998–6008.

Wang, J., D. Chen, Z. Wu, C. Luo, L. Zhou, Y. Zhao, … L. Yuan. 2022. “Omnivl: One Foundation Model for Image-Language and Video-Language Tasks.” *Advances in Neural Information Processing Systems* 35: 5696–5710.

Wang, S., H. Ning, X. Huang, Y. Xiao, M. Zhang, E. F. Yang, Y. Sadahiro, et al. 2023. “Public Surveillance of Social Media for Suicide Using Advanced Deep Learning Models in Japan: Time Series Study from 2012 to 2022.” *Journal of Medical Internet Research* 25: e47225. https://doi.org/10.2196/47225.

Weng, M. H., S. Wu, and M. Dyer. 2022. “Identification and Visualization of Key Topics in Scientific Publications with Transformer-Based Language Models and Document Clustering Methods.” *Applied Sciences* 12 (21): 11220. https://doi.org/10.3390/app122111220.

Yang, J., H. Jang, and K. Yu. 2023. “Analyzing Geographic Questions Using Embedding-Based Topic Modeling.” *ISPRS International Journal of Geo-Information* 12 (2): 52. https://doi.org/10.3390/ijgi12020052.

Ye, X., J. Du, Y. Han, G. Newman, D. Retchless, L. Zou, … Z. Cai. 2023. “Developing Human-Centered Urban Digital Twins for Community Infrastructure Resilience: A Research Agenda.” *Journal of Planning Literature* 38 (2): 187–199. https://doi.org/10.1177/08854122221137861.

Yu, Y., S. Buchanan, D. Pai, T. Chu, Z. Wu, S. Tong, … Y. Ma. 2023, November 24. White-Box Transformers via Sparse Rate Reduction: Compression Is All There Is? https://doi.org/10.48550/arXiv.2311.13110

Yu, M., A. Masrur, and C. Blaszczak-Boxe. 2023a. “Predicting Hourly PM2.5 Concentrations in Wildfire-Prone Areas Using a SpatioTemporal Transformer Model.” *Science of The Total Environment* 860: 160446. https://doi.org/10.1016/j.scitotenv.2022.160446.

Yu, M., S. Zhang, K. Zhang, J. Yin, M. Varela, and J. Miao. 2023b. “Developing High-Resolution PM2.5 Exposure Models by Integrating low-Cost Sensors, Automated Machine Learning, and big Human Mobility Data.” *Frontiers in Environmental Science*, https://doi.org/10.3389/fenvs.2023.1223160.

Zhang, Z., H. Amiri, Z. Liu, A. Züfle, and L. Zhao. 2023a. Large Language Models for Spatial Trajectory Patterns Mining (arXiv:2310.04942). http://arxiv.org/abs/2310.04942.

Appendix

Title	Authors	Journal	Keywords	PaperType	Domain	Data	Method	Role
The Ethics of AI-Generated Maps: A Study of DALLE 2 and Implications for Cartography	Kang, Zhang, and Roth (2023)	-	ChatGPT; DALL·E; place identity; generative artificial intelligence; sense of place	Preprint	Cartography	Image	DALL·E 2 to generate AI-maps and compare with real maps	Generator
Understanding Place Identity With Generative AI	Jang et al. (2023)	Leibniz International Proceedings in Informatics (LIPIcs)	Place identity · Generative AI · ChatGPT · GeoAI	Conference paper	Cartography	Text	transformer BERT model	Purifier
Visual Language Maps for Robot Navigation	Huang et al. (2023)	Proceedings – IEEE International Conference on Robotics and Automation	Computational linguistics; HTTP; Navigation; Robots; Visual languages; Geometric maps; Grounding language; Image caption; Language description; Map representations; Natural languages; Robot navigation; Spatial maps; Visual language model; Visual observations; Three dimensional computer graphics	Conference paper	Cartography	photos	computing dense pixel-level embeddings from an existing visuallanguage model and by backprojecting them onto the 3D surface of the environment via VLMaps	Purifier, Convertor
Mapping with ChatGPT	Tao and Xu (2023).	ISPRS International Journal of Geo-Information	ChatGPT; cartography; thematic maps; mental maps; spatial thinking	Article	Cartography	Text	conducted a pilot study exploring making maps with ChatGPT,	Generator
ChatGPT as a mapping assistant: A novel method to enrich maps with generative AI and content derived from street-level photographs	Juhász et al. (2023)	-	ChatGPT; OpenStreetMap; Mapillary; LLM; volunteered geographic information; mapping	Preprint	Cartography	Vector	Prompt engineering (use information extracted from images)	Generator
Neural network based successor representations to form cognitive maps of space and language	Stoewer et al. (2022).	Scientific Reports	Spatial representation; hippocampal-formation; grid cells; memory; microstructure; mechanisms; dorsal; world; rats	Article	Cartography	Raster	neural network based approach for classification	Convertor
ChatGPT is Good but Bing Chat is Better for Vietnamese Students	Dao and Le (2023)	-	-	Preprint	Education	Text	Exam purpose	Generator, Reasoner
Geo-knowledge-guided GPT models improve the extraction of location descriptions from disaster-related social media messages	Hu et al. (2023)	International Journal of Geographical Information Science	Location description; social media; disaster; GPT; GeoAI	Article	Environmental Science	Text	Geospatial prompt engineering (add more geographic context to prompt)	Generator

(*Continued*)

Appendix Continued.

Title	Authors	Journal	Keywords	PaperType	Domain	Data	Method	Role
Enhancing Chinese Address Parsing in Low-Resource Scenarios through In-Context Learning	Ling et al. (2023).	Isprs International Journal Of Geo-Information	Chinese address parsing; low-resource scenarios; In-Context Learning; GPT; BERT; K-Nearest Neighbor	Article	Informatics	Text	Few-shot learning methods, language models (gpt plus bert),(RNN)-based model	Convertor
GeoLLM: Extracting Geospatial Knowledge from Large Language Models	Manvi et al. (2023)	ICLR 2024	–	Conference paper	Geotechnical engineering	Text	Geospatial prompt engineering (add more geographic context to prompt)	Purifier
GeoQAMap – Geographic Question Answering with Maps Leveraging LLM and Open Knowledge Base	Feng, Ding, and Xiao (2023).	Leibniz International Proceedings in Informatics, LIPIcs	Computational linguistics; Knowledge based systems; Knowledge management; Visual languages; Geographic question answering; Geographics; Knowledge base; Language model; Large language model; Natural languages; Question Answering; Research fields; SPARQL; Wikidata; Natural language processing systems	Conference paper	Geotechnical engineering	Text	Generating standardized queries	Generator, Purifier
Analyzing Geographic Questions Using	Yang, Jang, and Yu (2023).	ISPRS International Journal of Geo-Information	GeoQA; GeoQA dataset; KBQA; semantic parsing; topic modeling	Article	Geotechnical engineering	Text	Embedding	Convertor
Geotechnical Parrot Tales (GPT): Harnessing Large Language Models in geotechnical engineering	Kumar (2023)	–	–	Preprint	Geotechnical engineering	Text	Code/file generation, Establish a workflow	Generator, Reasoner
Evaluating the Effectiveness of Large Language Models in Representing Textual Descriptions of Geometry and Spatial Relations	Ji and Gao (2023).	Leibniz International Proceedings in Informatics, LIPIcs	Computational linguistics; Domain Knowledge; Embeddings; Embeddings; Foundation models; GeoAI; Geometric attributes; Language model; Large language model; Research focus; Spatial relations; Text format; Textual description; Geometry	Conference paper	Geotechnical engineering	Vector	Embedding	Convertor
Large Language Models for Spatial Trajectory Patterns Mining	Zhang et al. (2023a)	AAAI 2024	–	Conference paper	Human geography	Spatiotemporal data	Prompt engineering (with/ without hints)	Reasoner

(*Continued*)

Appendix Continued.

Title	Authors	Journal	Keywords	PaperType	Domain	Data	Method	Role
OmniVL: One Foundation Model for Image-Language and Video-Language Tasks	Wang et al. (2022).	Advances in Neural Information Processing Systems	Character recognition; Modeling languages; Natural language processing systems; Text processing; Visual languages; Action recognition; Foundation models; Image texts; Images classification; Language model; Performance; Pre-training; Spatial dimension; Temporal dimensions; Text images; Image classification	Conference paper	Informatics	Video	vision-language contrastive (UniVLC)	Convertor
Identification and Visualization of Key Topics in Scientific Publications with Transformer-Based Language Models and Document Clustering Methods	Weng, Wu, and Dyer (2022).	Applied Sciences-Basel	topic analysis; language model; document clustering; keyword extraction; bibliometric analysis	Article	Informatics	Text	GPT-3 embedding, topic detection and deploys the HDBSCAN (Hierarchical Density-based Spatial Clustering of Applications with Noise) clustering algorithm	Generator, Purifier
Autonomous GIS: the next-generation AI-powered GIS	Li and Ning (2023)	International Journal of Digital Earth	Spatial analysis; autonomous agent; artificial intelligence; large language models; ChatGPT	Article	Spatial analysis and query	Text	Establish a workflow	Reasoner, Generator
SeeGULL: A Stereotype Benchmark with Broad Geo-Cultural Coverage Leveraging Generative Models	Jha et al. (2023)	–	–	Preprint	Policy geography	Text	Prompt engineering (generate text from provided seeds)	Generator
GPT4GEO: How a Language Model Sees the World's Geography	Roberts et al. (2023)	–	–	Preprint	Spatial analysis and query	Text	prompt engineering (including geographical information in text format)	Generator, Purifier
GeoGPT: Understanding and Processing Geospatial Tasks through An Autonomous GPT	Zhang et al. (2023b)	–	Geospatial semantic understanding; AutoGPT; GeoAI, foundation model	Preprint	Spatial analysis and query	Text	Establish a workflow via LangChain	Generator, Reasoner
Core Building Blocks: Next Gen Geo Spatial GPT Application	Fernandez and Dube (2023)	–	Geo-Spatial Analytics; Spatial Vector Representation; GPT; Embeddings	Preprint	Spatial analysis and query	Text	Prompt engineering	Generator, Reasoner
Can Large Language Models be Good Path Planners? A Benchmark and	Aghzal, Plaku, Yao (2023)	–	–	Preprint	Transportation	Text	Geospatial prompt engineering (add more	Generator, Reasoner

(*Continued*)

Appendix Continued.

Title	Authors	Journal	Keywords	PaperType	Domain	Data	Method	Role
Investigation on Spatial-temporal Reasoning							geographic context to prompt)	
Semantic understanding and prompt engineering for large-scale traffic data imputation	Zhang et al. (2024).	Information Fusion	Computational linguistics; Cutting tools; Graph neural networks; Intelligent systems; Intelligent vehicle highway systems; Motor transportation; Search engines; Semantic Web; Semantics; Data imputation; Graph neural networks; Language model; Large language model; Large-scales; Prompt engineering; Semantics understanding; Traffic data; Traffic data imputation; Transformer; Roads and streets	Article	Transportation	Spatiotemporal data	prompt engineering (spatiotemporal semantic description)	Purifier
Generating Individual Trajectories Using GPT-2 Trained from Scratch on Encoded Spatiotemporal Data	Horikomi et al. (2023).	–	–	Preprint	Transportation	Spatiotemporal data	Embedding	Convertor
Synthetic News as a Tool for Evaluating Urban Area Development Policies	Shuklin et al. (2022)	8th International Conference on Engineering and Emerging Technologies, ICEET 2022	Artificial intelligence; Semantics; Urban growth; Area development; Gamification; News feeds; News generation; Social response insert; Sources of informations; Synthetic news; Text generations; Urban areas; Urban event; Decision support systems	Conference paper	Urban study	Text	GPT-3 text generation	Generator