

EDITORIAL

Open Access



The generative revolution: AI foundation models in geospatial health—applications, challenges and future research

Bernd Resch^{1,2}, Polychronis Kolokoussis³, David Hanny¹, Maria Antonia Brovelli⁴ and Maged N. Kamel Boulos^{5*}

Abstract

In an era of rapid technological advancements, generative artificial intelligence and foundation models are reshaping industries and offering new advanced solutions in a wide range of scientific areas, particularly in public and environmental health. However, foundation models have previously mostly focused on understanding and generating text, while geospatial features, interrelations, flows and correlations have been neglected. Thus, this paper outlines the importance of research into Geospatial Foundation Models, which have the potential to revolutionise digital health surveillance and public health. We examine the latest advances, opportunities, challenges, and ethical considerations of geospatial foundation models for research and applications in digital health. We focus on the specific challenges of integrating geospatial context with foundation models and lay out the future potential for multimodal geospatial foundation models for a variety of research avenues in digital health surveillance and health assessment.

Keywords Generative AI, Large language models, AI agents, Geospatial health, Health surveillance, AI-powered public health, Geospatial foundation models, Multimodal data and analysis

Introduction

This vision paper outlines the opportunities and challenges in the integration of geospatial factors into health research, highlighting the potential of advanced AI methodologies and multimodal data analysis. The paper further explores the profound implications of geospatial foundation models (FM), examining its key applications,

challenges, and the path forward toward AI-powered geospatial insights in health and disease research. In this paper, we lay the groundwork for the development and implementation of FMs in health geography. We start out by providing an overview of existing FMs like Language Foundation Models, Geospatial and Vision FMs, and Multimodal FMs. Thereafter, we identify concrete challenges in current research involving FMs in geospatial health. Finally, we present opportunities for the effective use of FMs in health geography and provide an agenda for future research.

Overall, our contributions can be summarised as follows:

- We systematically examine the potential and effectiveness of large AI foundation models in health geography, providing the first assessment of the State of the Art (SotA) in this domain.

This editorial is the second in a pair of editorials. The first editorial is available online at <https://doi.org/10.1186/s12942-025-00392-z>.

*Correspondence:

Maged N. Kamel Boulos
mnboulos@ieee.org

¹ IT:U Interdisciplinary Transformation University, 4040 Linz, Austria

² Center for Geographic Analysis, Harvard University, Cambridge, MA 02138, USA

³ School of Rural, Surveying & Geoinformatics Engineering, National Technical University of Athens, 15780 Athens, Greece

⁴ Department of Civil and Environmental Engineering, Politecnico Di Milano, 20133 Milan, Italy

⁵ School of Medicine, University of Lisbon, 1649-028 Lisbon, Portugal



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

- We thoroughly discuss the challenges and risks of developing and integrating FMs in a health-related context.
- We present an overview of the concrete opportunities of FMs in geospatial health including clear future research avenues.

This paper is the second in an *Int J Health Geogr* two-article series (2025) on the ‘Generative Revolution’. The first article in the series entitled ‘The Generative Revolution: A Brief Introduction’ provides an additional brief introduction and context to the current article.

Geospatial digital health surveillance

Digital health surveillance and disease research have recently gained more and more interest, particularly through the rise of smartphones, artificial intelligence methods, and the widespread availability of digital data from various sources. Numerous research efforts have investigated the use of smartphone usage patterns, smartphone-based disease diagnostics, smart wearable sensors, and large-scale data sources such as Internet searches and social media posts, among others. These efforts aimed to complement and enhance traditional methods in health research and practice, thus generating meaningful insights, enabling timely interventions in vulnerable regions [1].

However, the integration of geospatial factors in health research has historically been underexplored and underutilised. Geospatial health research primarily focuses on the spatial dimensions of health issues and outcomes, examining how location as well as the social, natural or built environment influence disease patterns, and health-care accessibility. Similarly, environmental health applications investigate how external and human induced factors like climate change, pollution and urbanisation impact the environment and well-being. This consideration of a multitude of context factors is key in the context of the World Health Organization’s (WHO) One Health initiative that “addresses the interconnected risks and vulnerabilities of human, animal and ecosystem health” [2].

Commonly employed methodologies, such as time series analysis and the use of discrete geospatial units like administrative boundaries, have often disregarded the complexities of spatial interactions and relationships. This research gap is attributable to a range of challenges that previously hindered progress in the field:

Data Limitations: Historically, the required geospatial data were either unavailable or accessible only in limited forms, thus limiting the significance and scalability of approaches for digital health research [3].

Modelling Complexity: The development and application of spatial interaction models are typically complex and resource-intensive. While one-dimensional models, e.g., for time series analysis and prediction, have been widely researched and adopted, health research still struggles with scalable and generalist spatial models, partly due to complex spatiotemporal covariance matrices [4]. Geographic data often exhibit spatial dependence or autocorrelation, which violates the assumption of independence common in many traditional statistical models, requiring specialised techniques in spatial statistics or explicit geospatial machine learning approaches. Models must account for this dependence to avoid biases and ensure accurate results.

Disease-Specific Focus: Existing models often lacked generalisability, being narrowly tailored to specific diseases without accommodating broader applications. This limits both their reliability and robustness as well as their potential for practical applicability [5].

Image-Centric Approaches: Research on geospatial health analysis models oftentimes focused on image data, thereby excluding other valuable geospatial modalities.

Paradigm changes in geospatial health surveillance

Despite these shortcomings, recent advancements in technology, widespread data availability and new analysis methods have transformed the landscape of geospatial health research, creating new opportunities to address these historical limitations. Methodologically, this also drove the emergence of geospatial multimodal approaches that integrate various data types for comprehensive insights. Particularly AI algorithms have matured to address many of the issues inherent in traditional purely hypothesis-driven modelling approaches. Foundation models, in particular, provide a powerful framework for leveraging largescale geospatial datasets and overcoming prior computational and modelling barriers.

These developments have led to a shift in geospatial health research, for instance, in the following areas:

- **Disease Prediction and Surveillance:** AI models have significantly improved the accuracy of predicting disease outbreaks, such as dengue fever and COVID-19, by analysing environmental and social determinants [6].
- **Environmental Health Assessment:** AI can effectively monitor air and water quality, linking these factors to health outcomes. To better understand environmental public health risks, a multi-layer geospatial analysis was conducted by combining environmental data (air and water quality), demographic information (population density, age distribution), and epidemiological records (disease incidence) [7]—Fig. 1. Simi-

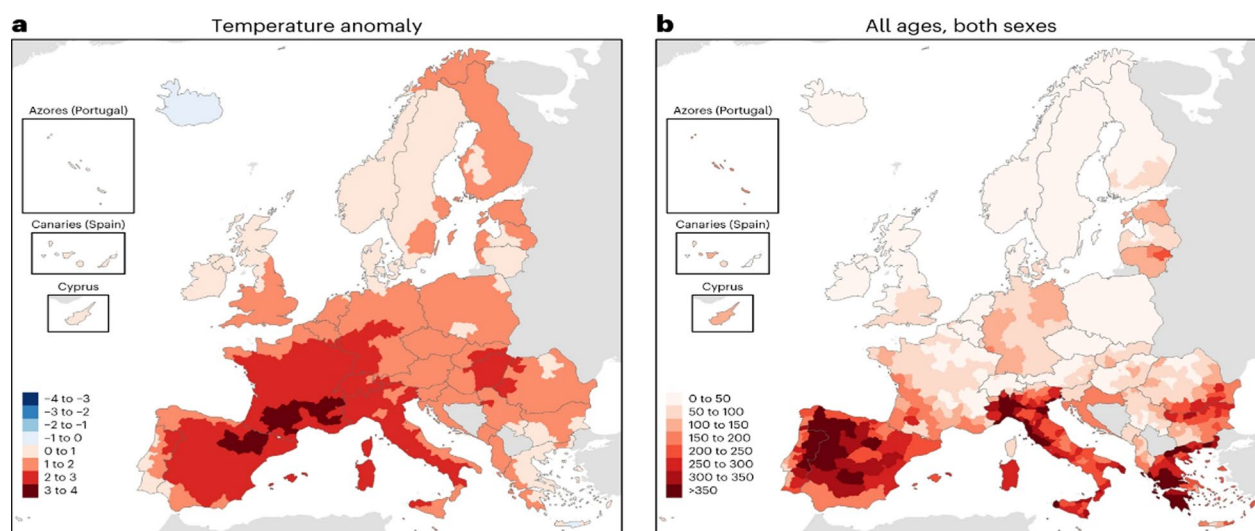


Fig. 1 **a** Regional temperature anomaly (°C) averaged over the summer of 2022. **b** Regional heat-related mortality rate (summer deaths per million) aggregated over the summer of 2022 for the whole population (image adapted under a Creative Commons Attribution 4.0 International License from [7])

larly, the Malaria Atlas Project [8] aims at converging geospatial data with AI to provide targeted insights into malaria prevalence and control strategies, and the Map of Life initiative [9] aims to help identify and close key information gaps and highlight species of greatest concern.

- **Public Health Risk Prediction:** GenAI methods, utilizing machine learning and natural language processing, enhance risk prediction by integrating diverse data sources, detecting non-linear relationships, and identifying latent patterns [10].
- **Pandemic Modelling:** A notable initiative was the correlation of pollution parameters with mortality [11]. By integrating Large Language Models (LLMs), these models provided localised risk assessments and tailored public health guidance, such as targeted lockdowns and vaccination drives, demonstrating the potential for proactive pandemic response [12].
- **Resource Allocation:** AI-driven models help in optimising the distribution of healthcare resources, ensuring better preparedness and response to health crises. For example, researchers utilised mobility data from smartphones and environmental factors such as temperature and humidity to predict the spread of the virus [13] to support policy-making and health resource allocation.
- **Identifying Health Disparities:** AI is capable of identifying and addressing health disparities by analysing geospatial data to uncover patterns related to socioeconomic factors. Geospatial AI models help identify healthcare areas lacking adequate medical facilities

by analysing population density, transportation networks, and socioeconomic data [14, 15]. Foundation models, particularly LLMs, augment these efforts by interpreting demographic surveys and historical health outcomes [16] to suggest equitable resource allocation strategies, such as establishing mobile clinics or optimising ambulance routes.

Foundation models in health research

The field of AI has recently undergone a substantial paradigm shift towards FMs, i.e., large AI models pre-trained on vast Internet-scale datasets [17]. Rather than training task-specific individual models from scratch, FMs are fine-tuned using few-shot or zero-shot learning strategies on top of pre-training. Fine-tuned FMs have presented remarkable performance across a diverse range of tasks including text classification [18], question answering [19], image classification [20] or segmentation [21].

These rapid advancements have naturally transferred to the domain Geospatial Artificial Intelligence (GeoAI), though the development of a geospatial FM remains challenging due to the multimodal nature of most geospatial data, which encompasses geospatial information such as geometries along with text, images, graph or vector data [22]. From an application perspective, geospatial FMs hold significant potential for health geography. Mai et al. [22] demonstrated the capabilities of the pre-trained Generative Pre-trained Transformer (GPT) language model [23] and its variant InstructGPT [24] to predict

dementia-related death counts on a state- and county-level in the United States of America (USA), even outperforming a traditional AutoRegressive Integrated Moving Average (ARIMA) model.

Regardless, FMs are still underused in the domain of geospatial health. Their adoption is particularly complicated due to the multimodal complexity of diseases [25]. The spread of viral infections, for instance, is highly heterogeneous in space and the process of contagion is location-dependent as the virus moves through different regions [26]. Health is additionally influenced by various socio-ecological covariates, making disease dynamics a highly complex system [27].

Next-generation natural language conversational interfaces and agentic systems (AI Agents) for querying geospatial catalogues: the geoexposomics demonstrator example

In a recent ISPRS (International Society for Photogrammetry and Remote Sensing) funded project under ISPRS Scientific Initiatives 2023, researchers from five countries proof concepted and disseminated their initial development concepts about a much-needed metadata catalogue

of Earth observation data sources/products and types that are relevant to human health research in exposomics [28]. The proof-of-concept (PoC) searchable catalogue takes the form of a dedicated Geoexposomics Web portal that is provided as a free service to interested parties worldwide at [29]. A key part of this effort involved examining a number of complementary user interface and experience options to make the PoC catalogue searchable and more accessible to its end users, and to improve the discoverability of its content. Notable among these experimental user experiences is GeoX-GPT, a ChatGPT-like natural language conversational interface for querying the metadata catalogue powered by an LLM and Retrieval-Augmented Generation (RAG), whereby the PoC catalogue became an external knowledgebase to GPT, in order to provide more relevant and better cross-linked search results in response to user's prompts. Geo-X GPT has the potential of better handling user questions and interactions with the catalogue by allowing users to query the catalogue in an unrestricted, natural way using their own words and terms and to receive more expert-human-like relevant answers from the system (Fig. 2) [28, 29].

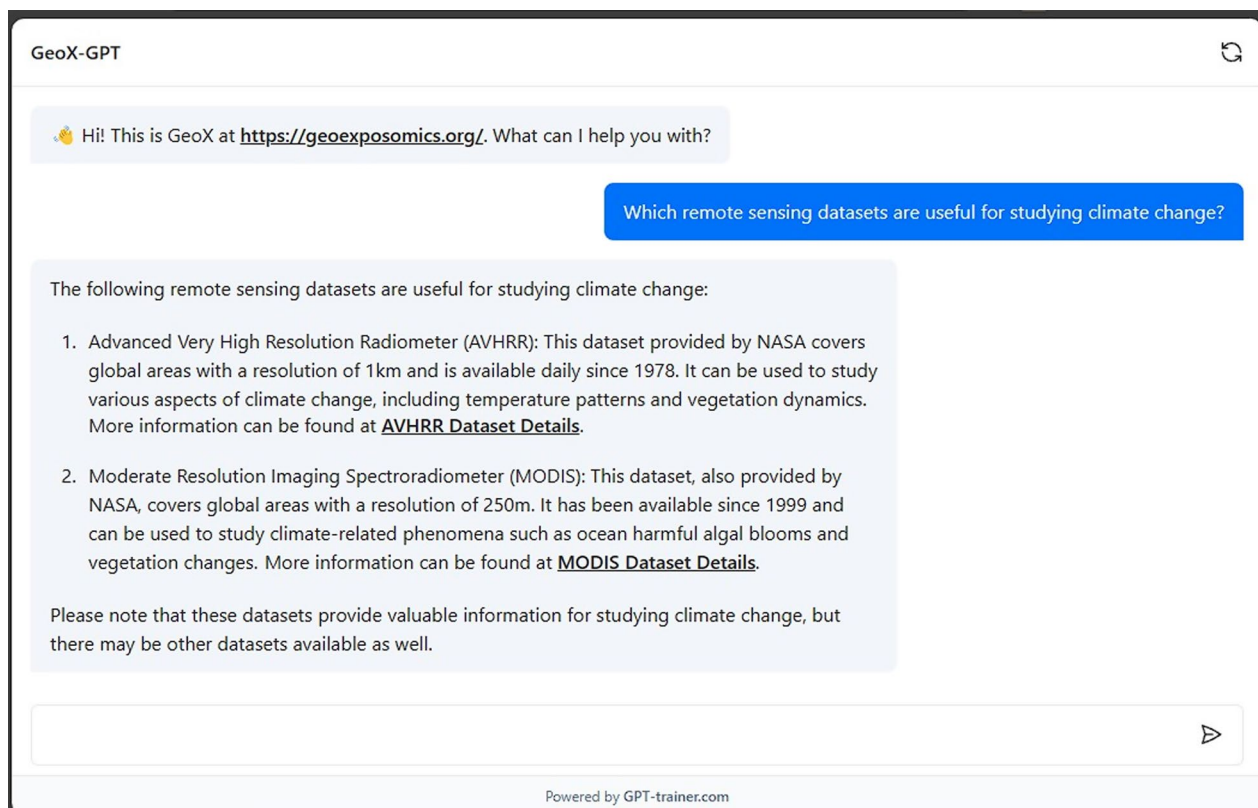


Fig. 2 Screenshot of GeoX-GPT. This experimental RAG implementation was developed in GPT-trainer. It features a natural-language-based, AI-driven conversational user interface to enhance discovery and explorability by allowing the search user interface to more flexibly adapt to users' needs in more natural ways instead of asking users to adapt to and master a rather rigid (preprogrammed) and less forgiving conventional interface

However, using remote sensing datasets in own research can often prove challenging to novice researchers. For example, researchers should be aware of, and address, the fact that various remote sensing datasets can have highly variable accuracy. Users who are less experienced in exposure measurement would definitely benefit from this kind of practical tips. One possible future direction for Geo-X GPT would be to expand it into a fully-fledged agentic system (AI agent) that contextually incorporates and automates this tip and others (the 'know-how' of expert users) to assist less experienced users (as well as experienced researchers) in selecting, preparing, and running the most appropriate datasets, approaches and methods for a given investigation or research question. AI agents and conventional workflows are not the same; agentic systems can offer a better solution when operating in a wide-ranging, dynamic research and investigation environment covering different health conditions and datasets with many variations, variables, interdependencies, uncertainties, data formats, quality and completeness issues, and other data selection and analysis factors to consider. For more information about agentic systems, please refer to [30–32].

Foundation models in geospatial health: state of the art

The history of geospatial foundation models reflects the convergence of advances in geospatial sciences, machine learning, and big data analytics. Early GIS platforms like ESRI's ArcGIS, GRASS GIS, and others laid the groundwork for geospatial data processing, focused on static mapping and spatial data visualisation. The integration of FMs in geospatial health has seen various degrees of integration. Language FMs, though frequently used in the medical and public health domain, have seen little consideration for geographic space. Conversely, vision FMs and their geospatial variants are used more widely in the field of remote sensing. However, a research geospatial FM that are capable of handling geo-data beyond images is still underrepresented, although they can increase the effectiveness and precision of geographical modelling, improving public health outcomes [33, 34].

Language foundation models

Large pre-trained language FMs, usually referred to as LLMs in the literature, are a subset of AI models designed to process and generate human-like text. They have fundamentally changed the field of Natural Language Performance (NLP) in less than a decade. Models such as Bidirectional Encoder Representations from Transformers (BERT) [35], OpenAI's GPT [36] and Meta's Llama [37] are trained on immense amounts of textual data scraped from the web [38] using self-supervised learning

tasks, thus involving no human labelling. Fine-tuned variants of these LLMs have demonstrated exceptional performance for numerous language understanding tasks such as content summarisation [39], translation [40], zero-shot classification [41], question answering [19] or even logical reasoning [42].

The internal structure of these language FMs usually relies on the transformer architecture [43] and frequently involves billions of parameters, making them cost-intensive to run. However, they are significantly more sample-efficient compared to smaller, non-pre-trained models, making fine-tuning with little data substantially more effective [44]. Chatbot-like interfaces to generative language FMs such as ChatGPT also allow for instruction-based fine-tuning using natural language [24] and prompting techniques [45]. More recent models with reasoning capabilities, e.g., by OpenAI (o1) and DeepSeek (R1), have further improved the performance of LLMs. Users can now run some of these LLMs on their own (suitable) laptops and PCs entirely offline to keep all their interactions with the LLM private and secure, e.g., DeepSeek R1 runs smoothly on NVIDIA GeForce RTX 50 Series AI PCs [46], but also on some earlier RTX 30 and 40 Series machines.

In this context, Clinical Language Models (CLaMs), language models fine-tuned for and with electronic medical data, have gained significant popularity [47]. CLaMs have been successfully used to extract drug names or medical information from text [48], medical question answering [49] or medical dialogue summarisation [50]. While general-purpose FMs such as GPT are becoming increasingly powerful CLaMs still tend to vastly outperform these models for medical tasks [51, 52]. Beyond the pure medical scope, language FMs have also been utilised to process large amounts of textual information like social media posts or news articles in the context of epidemic outbreaks, particularly during the COVID-19 crisis in 2020 [53].

However, the capabilities of language FMs in a health context has not been extended by geospatial capabilities. Previous research in Geoinformatics tends to focus on the generalised applications of language FMs in Geographic Information System (GIS), from information and location extraction [54] to fully autonomous GIS and mapping workflows [21, 55]. Consequently, neither do CLaMs explicitly consider geographic context, nor have language FMs been specifically utilised for investigations in geospatial health like context-based diagnosis or disease count prediction.

Geospatial and vision foundation models

The concept of FMs was originally pioneered by vision models like VGGNet [56], ResNet [57] and AlexNet [58],

and was later advanced by the Vision Transformer (ViT) architecture [59], which adapted transformer-based approaches from natural language processing to image processing. Foundation models in geospatial contexts are agnostically trained on diverse and massive unlabeled geospatial datasets and can generalise across tasks [60, 61]. In this way, these models can be fine-tuned on relatively smaller, task-specific annotated datasets for specific tasks, such as predicting disease outbreaks, assessing environmental risks, and optimizing resource allocation, leading to more efficient and effective solutions.

Standard vision FMs are frequently employed in medical imaging for tasks like image segmentation [62] and classification [63], though more complex tasks like question answering are usually powered by multimodal FMs instead of being purely vision-based [64]. However, vision FMs have been widely adopted in the geospatial domain, resulting in the first generation of geospatial FMs which are pre-trained on large amounts of geospatial image data from satellites. Additional challenges introduced by using satellite imagery include the spectral differences between satellites and variations in spatial resolutions, e.g. 10 m for Sentinel-2 and 30 m for Landsat-9 [65]. NASA/IBM's Prithvi model, for instance addresses the problems of spectral differences through harmonisation and was pre-trained on more than 1 TB of multispectral satellite imagery from Sentinel-2 [66]. Pre-trained Prithvi outperformed the SotA on various earth observation tasks like for flood mapping on 10 m resolution or wildfire scar segmentation. Other notable geospatial FMs include SatMAE [60], Scale-MAE [67] and DINO-MC [68]. These geospatial FMs can effectively be used for public health tasks like water quality detection [69], pollution monitoring [70] or mosquito breeding site detection [71].

However, this first generation of geospatial FMs can be described more accurately as geospatial *vision* as they rely exclusively on image data. This not only limits their applicability to other spatial data such as geo-referenced text, numbers or trajectories but also defies the original goal of developing generalisable geospatial FM. As Mai et al. [22] state, a true geospatial FM must be capable of handling different data sources and types beyond satellite images [22]. The broad applicability of geospatial FMs in a health context is therefore currently limited by the nature of these models. Epidemiological modelling or contextualised medical diagnosis is not possible using current geospatial FMs.

Multimodal foundation models

Based on the progress made in language modelling, FM development has recently shifted towards so-called LMMs. These are large-scale FMs that integrate multiple modalities like language, vision and audio. One of

the first breakthroughs of this kind was achieved by the Vision-Language Model (VLM) Contrastive Language-Image Pre-training (CLIP) [23], which uses self-supervised contrastive learning to learn joint embeddings for images and text. Numerous follow-up works like Bootstrapping Language-Image Pretraining (BLIP) [72, 73], BEiT [74] and Large Language and Vision Assistant (LLaVa) [75] have been presented, improving upon the original CLIP approach. The commonly used GPT-variant is also an inherently multimodal FM, supporting both vision and language by default [36]. GPT now additionally supports audio and video, and so does Google's family of Gemini models [76]. Both models allow for almost natural human interaction using spoken language with capabilities to generate and read visual content alongside. The simultaneous handling of audio, vision and language has also been discussed in data2vec [77] which presents a way to learn a joint latent embedding space for all three modalities.

Multimodal FMs are increasingly used in the medical domain with the main application areas being visual question answering powered by a fine-tuned model like Gemini [78]. With image-understanding capabilities, multimodal FMs are also well-suited generating medical reports from both text and image data [79]. Liu et al. [80] further suggest the use of multimodal FMs as personal virtual assistants in medicine, enabling remote healthcare, multilingual medical communication and the efficient integration of academic literature, clinical guidelines and case studies [80].

However, geographic space has not yet been considered in the development of multimodal medical FMs. Conversely, health geography has also not received much attention as an application area for multimodal FMs in Geoinformatics. So far, multimodal FMs are primarily viewed as reasoners with extended capabilities compared to language FMs in the geospatial sciences. For instance, Wang et al. [81] proposed the use of vision-language models to generate captions for street view imagery or explainable flood depth estimation where the model outputs a reason for the estimated flood depth such as 'The flood inundates the adult male's knee, so I guess the depth is about 0.5 m' [81]. We therefore identify a substantial research gap concerned with bringing together the geospatial and medical capabilities of FMs.

Current research challenges with foundation models in geospatial health

While FMs have demonstrated remarkable performance for both medical and geospatial tasks, their usage in health geography is limited by significant challenges in data curation, generalisability, prediction quality and credibility.

Geospatial multimodality

The increasing availability of geospatial health data in recent years has facilitated advancements in applications like COVID-19 monitoring [82] or Dengue disease control [83]. Despite these advances, a model for handling multiple types of geospatial health data remains undeveloped. FMs pre-trained on large amounts of such data could overcome the limitations of disease-specific models. Such an effort, however, would require an FM that is capable of (1) handling multiple modalities such as numbers, text, images or trajectories simultaneously while (2) being able to consider geographic space during training and prediction. As current geospatial FMs only integrate image data, the development of a multimodal geospatial FM is essential for leveraging their capabilities in health geography.

Spatially explicit learning

In many ways, geospatial data defies the assumption made in traditional Machine Learning (ML) settings such as the hypothesis that the data points are independent and identically distributed (i.i.d.). It furthermore presents challenges like spatial dependence, spatial heterogeneity and the Modifiable Areal Unit Problem (MAUP). Therefore, spatially explicit learning techniques have gained significant traction and have been shown to outperform traditional non-spatial methods [84–88]. In recent years, ML methods have been widely adopted in health geography with applications in public health surveillance [89, 90], environmental analysis [91], infection risk factor identification [92] or vector control [83]. However, the utilised techniques are rarely spatially explicit, highlighting the potential for improvement through the explicit consideration of geographic space. Consequently, geospatial FMs for health geography should not only be viewed from a data perspective but also integrate geographic space from an architectural viewpoint.

Data availability, generation and curation

In order to achieve generalist capabilities, FMs must be trained on massive, diverse and high-quality health-related geospatial datasets [93]. The creation of such datasets requires access to credible data on disease spread, patient data, pollution measurements and other variables from around the globe. It must be diverse enough to cover all regions and populations of the world while following a common understanding of diseases and infections. The construction of a large-scale dataset therefore encompasses ethical, economic and social challenges. Currently, access to health data within research groups is often limited to small sample sizes from small geographic areas [94]. Data curation therefore poses a

major research challenge for leveraging FMs effectively in a health context.

In part, insufficient training data can be augmented using data generation techniques to improve model performance [95]. For instance, [96] have demonstrated that the classification of tweets regarding COVID-19 intervention measures can be improved using distillation [97]. Additionally, Kim et al. [98] presented an approach for generating like geo-social urban movement trajectories [98]. Extending these methods to geospatial health data while incorporating multiple modalities can mitigate some of the lack in training data. However, generating credible spatio-temporal synthetic datasets is far from trivial, requires careful assessment for bias and has not been tackled yet. In this context, FMs (1) are able to generate higher quality synthetic data compared to previous methods [99] and (2) can benefit from this additional data during pre-training or fine-tuning. However, multimodal data generation and trustworthiness are still a key research challenge.

Credibility and explainability

An increase in the reliance on FMs in health geography requires the models to be highly credible. That is, users must be confident in the model's output. This encompasses both accuracy and consistency. A model must be able to withstand bias and noise in the input data while producing consistent results across nuanced datasets [100]. While the credibility of modern FMs has improved over traditional deep learning methods [101], health data and its associations with the environment are frequently fuzzy [89]. Credibility therefore remains a significant challenge. While eXplainable Artificial Intelligence (XAI) can help make models more credible by providing a reason for the output, interpretability alone does not necessarily encompass credibility if the model's output is wrong [102]. Simultaneously, large-scale models with billions of parameters are increasingly challenging to understand, even if the architecture is completely transparent [101].

Geographic bias

FMs tend to amplify existing biases in the training data [103, 104]. In the context of health geography, this particularly concerns geographic bias which has gained increasing interest in GeoAI research. Manvi et al. [105] found that LLMs like GPT and Llama are clearly biased against locations with lower socioeconomic conditions, especially most of Africa, on a range of subjective topics such as attractiveness, morality and intelligence [105]. Liu et al. [106] further showed that neural-network-based geo-parsing models were highly biased in quality towards data-rich regions in Europe and the USA [106]. Liu et al.

[20] also found significant inter-regional disparities in the geo-guessing performance of several GPT variants for UNESCO World Heritage Sites [20]. Large-scale FMs also frequently carry inherent prejudices against certain groups of people such as the black population in the USA and tend to work best for high-resource languages such as English [107]. As debiasing massive FMs in retrospect is very difficult [22], unbiased training data is of utmost importance. Biases are particularly dangerous in a health context as they can easily be learned during fine-tuning [103], risking significant harm if the model is widely adopted.

This bias is oftentimes caused by hidden spurious correlations and can result in inaccurate and inequitable outcomes, such as neglecting marginalised communities. A notable example is the disparity in air quality monitoring data, where urban areas often have more sensors compared to rural or remote/isolated regions. This disparity can skew AI predictions and decision support interventions. Developers must ensure that AI/LLMs and geospatial models are inclusive and representative by integrating adequate and diverse datasets and performing regular audits. Identifying the reasons behind a biased system is not straightforward, since in many occasions they are associated with hidden spurious correlations which are not easy to spot, and specific tools like DOMINO, FACTS, ViG-Bias and Bias-to-Text are being used for its mitigation [108].

Privacy and security

Geospatial FMs also pose significant privacy risks due to potentially sensitive information that can be learned and disclosed by the model. For instance, a LLM could memorise home addresses present in the training data and disclose them when asked. In a health context, a multimodal FM could even learn the health status of people depicted in images. Someone could then upload a picture of a person and ask for the health of individual persons, potentially disclosing information that could threaten people's lives [109]. In case the model has access to additional geospatial databases, sensitive information could also be leaked from those. Malicious user interaction and adversarial attacks pose supplementary challenge in this context. Such attacks can cause the model to produce wrong results through well-designed noises that are frequently invisible to humans [101]. For instance, Perez and Ribeiro [110] state that GPT could easily be guided towards ignoring all previous prompts including system prompts by injecting specific attack prompts [110]. Similarly, Schlarmann and Hein [111] found that multimodal vision-language models can be attacked to lead users to malicious websites or provide fake information using

imperceptible alterations on images [111]. Such security issues would be intolerable in health applications where interventions critically impact patient outcomes and well-being.

Furthermore, during the development of contact tracing apps during the COVID-19 pandemic, questions arose about how much personal location data should be shared to balance public safety with individual privacy. Striking a balance between utility and confidentiality is critical, requiring robust encryption, anonymisation, and ethical oversight. Haltaufderheide and Ranisch conducted a systematic review on the ethics of ChatGPT and other LLMs in medicine, identifying four broad LLM application categories (covering health professionals and researchers, patient support, clinical applications, and public health) and a number of recurring ethical issues related to epistemic values (reliability, transparency, hallucinations), therapeutic relationships, and privacy, among others [112].

Finally, the integration of geospatial data with health information raises privacy concerns. For example, during the COVID-19 pandemic, location data from contact tracing apps raised debates on privacy versus public health benefits. Striking a balance between utility and confidentiality is critical, requiring robust encryption, anonymisation, and ethical oversight to protect sensitive data while enabling actionable insights.

Ethical and functional considerations

One of the major drawbacks of FMs, particularly LLMs, are so called “hallucinations”, which stem from the inherent stochasticity of these models, whereby the models generate plausible sounding but factually incorrect or nonsensical information. This is a major issue for the utilisation of LLMs especially in healthcare and public health, but also to many other industries that cannot tolerate a 10% or even 5% error rate/misinformation rate. In medical applications for instance, if left undetected, these hallucinations can pose significant clinical risks to patients, and lead to misdiagnoses and inappropriate treatments.

Several tools are used to mitigate the hallucinations problem. For example, Hypercube is a tool for the automated detection and mitigation of LLM hallucinations, which integrates medical knowledge bases, symbolic reasoning, and NLP, allowing for an initial automated detection step before human expert review [113]. RAG technology can reduce a model's hallucinations by grounding the generated content in retrieved, verified data, but it is not a complete solution to the hallucination problem. There is ongoing research on this topic [114] which is of critical importance when it comes to health applications.

Other notable limitations in relation to LLM stochasticity include 'prompt brittleness' (slight modifications in prompts leading to significantly different outputs) and LLM unpredictability or reproducibility issues (their ability to generate different responses when prompted repeatedly with exactly the same prompt) [115]. These issues can affect the reliability and consistency of LLM-generated answers over multiple runs of the same or slightly paraphrased user queries.

Another recent finding concerns the inability of especially VLMs to understand negation. In multimodal tasks, VLMs like CLIP play a crucial role in areas such as image retrieval, image captioning, and medical diagnosis. The goal of these models is to align visual data with language data for more efficient information processing. However, current VLMs still face significant challenges in understanding negation, which is an issue particularly with tasks and questions that are defined by comprehensive inclusion and exclusion criteria. To address these issues, researchers from MIT, Google DeepMind, and the University of Oxford proposed the NegBench framework to evaluate and improve VLMs' understanding of negation [116].

Developing and deploying AI-powered geospatial solutions can be technically challenging, requiring high-performance computing. Ironically, training large models for environmental applications has its own environmental cost by requiring significant computational resources, thus significantly contributing to carbon emissions. Researchers must prioritise energy-efficient algorithms and reduce computational demands as much as possible.

Future research avenues for foundation models in geospatial health

Geospatial context: data/information and methods

The integration of geospatial information into health surveillance and prediction constitutes a critical frontier in medical informatics and public health. This approach leverages the spatio-temporal nature of health spread through analysing digital geodata to improve the accuracy and timeliness of real-time health surveillance.

A key component of this integration is the development and application of spatio-temporal plausibility measures. These measures evaluate the likelihood of disease occurrence based on the co-occurrence of temporal and spatial factors. For example, understanding the incubation periods and transmission dynamics of specific diseases allows for more precise predictions when combined with regional case trends. Further research should explore algorithms capable of dynamically weighting these factors based on disease-specific parameters, local epidemiological data, and environmental influences.

From a diagnostic viewpoint, chatbot-based systems such as Symptoma [117] could utilise geospatial data to enhance diagnostic likelihood calculations. For instance, if epidemiological records indicate that an unusually high number of measles cases have been recently reported in a specific region such as a city or a metropolitan area, symptoms reported in the same region could significantly increase the likelihood of a subsequent diagnosis of measles. This highlights the need for real-time incorporation of regional case counts into diagnostic decision-making processes.

The use of a wide variety of large-scale user-generated data represents another promising avenue. Sources such as Internet search trends, geo-social media activity, and data from smart thermometers can act as early indicators of disease outbreaks. Particularly geo-social media have been shown to significantly enhance the prediction of disease spreads [82, 90]. Furthermore, the rapid rise of large-scale wearable sensing technologies due to the rise of the "quantified self" movement has established a large-scale health monitoring system through the availability of affordable physiological sensors measuring heart rate, heart rate variability, electrodermal activity or skin and body temperature.

These data streams, while inherently noisy, could be refined through advanced natural language processing (NLP) and machine learning techniques to extract meaningful signals. For instance, increases in searches for "rash and fever" within a geographic region could correlate with the emergence of increasing disease case counts, providing an early warning system for public health authorities.

Moreover, patients' self-recorded images, as demonstrated in recent studies [118], also hold significant potential for diagnosis. Smartphone-captured images, analysed through deep learning-based image recognition algorithms, could identify dermatological or visual symptoms indicative of specific diseases. Future research should focus on enhancing these algorithms to accommodate diverse populations and varying image qualities, ensuring diagnostic reliability.

Incorporating clinical data from physicians' visits adds another layer of depth. This includes not only the documentation of symptoms and diagnoses but also metadata such as timestamps and geographic locations, enabling synchronisation with broader epidemiological trends. Linking these data points with travel activity can further refine diagnostic accuracy. Collective mobility patterns, such as increases in commuting or travel during holidays, can inform population-level disease risk models. Meanwhile, individual travel

Table 1 Key recommendations for successful generative AI adoption in geospatial health applications

<ul style="list-style-type: none"> • Fostering interdisciplinary/stakeholder collaboration and user engagement: Addressing the challenges and realizing the full potential of generative AI in geospatial and environmental health requires interdisciplinary collaboration among various researchers, policymakers, practitioners, and the public. Collaboration between AI developers, geospatial scientists, and public health experts has already shown success in projects like the Malaria Atlas Project. This initiative brought together expertise across disciplines to map disease patterns and inform targeted public health interventions, demonstrating the innovation that can be reached through interdisciplinary efforts • Mitigating LLM hallucinations: Ensuring faithfulness of results by mitigating or eliminating LLM hallucinations is of critical importance for many applications, especially those that could pose a threat to human health if inaccurate or incomplete information is generated and used. For generative AI tools to be safely and effectively integrated into mainstream health and healthcare, substantial technological advancements and mitigation strategies towards hallucination elimination are necessary • Ethical and unbiased generative AI deployment: Establishing a framework for ethical and unbiased generative AI deployment is essential, such as adopting principles of transparency, accountability, and inclusivity. Engaging local communities in AI initiatives ensures that solutions are contextually relevant and inclusive • Open data and open-source tools: Promoting open data sharing and developing open-source models and tools can accelerate the development and deployment of generative AI-powered, geospatial scalable and cost-effective solutions. Such tools would allow researchers and policymakers in low-resource settings to harness the power of advanced geospatial technologies, enabling data-driven decisions to improve public and environmental health outcomes • Promoting AI literacy, education and training: Investing in education and training material and courses to equip the next generation of researchers and practitioners with the skills and knowledge necessary to develop and implement effective generative AI-powered geospatial health solutions is equally important. As AI continues to advance, it is of paramount importance for all relevant researchers and practitioners to be offered regular access to suitable training opportunities to allow them to stay informed and continuously adapt their methods to emerging technologies and approaches
--

history provides critical exposure data, especially for diseases with known hotspots or outbreak regions.

To fully realise the potential of these approaches, robust methodologies for data integration and analysis are essential. Advances in data interoperability standards, privacy-preserving data sharing frameworks, and federated learning techniques could enable the secure and ethical utilisation of sensitive geospatial and health data. Moreover, interdisciplinary collaboration between epidemiologists, data scientists, clinicians, and public health officials is crucial to develop systems that are both scientifically rigorous and practically implementable.

Future research should also address the challenges associated with real-time data processing and visualisation. Developing interactive dashboards capable of presenting layered geospatial and temporal data could enhance decision-making for healthcare professionals and policymakers. These tools should be designed to accommodate the dynamic spatio-temporal nature of disease spread, incorporating predictive modelling to forecast potential outbreaks and resource needs.

In summary, integrating geospatial information into health surveillance and diagnosis demands comprehensive research across multiple domains. By combining real-time regional data, spatio-temporal analytics, user-generated content, clinical records, and mobility patterns, future systems can offer unprecedented precision in health analysis and monitoring as well as disease prediction. These efforts promise to transform public health responses, improving both individual patient outcomes and broader epidemic management strategies.

Geospatially enhanced foundation models

Geospatially explicit FM and question answering

Geospatially explicit FM offer significant advantages over non-spatial models by considering spatial structures, interactions, and correlations into their analytical approach. Unlike traditional models, which often treat spatial data as independent observations, geospatially explicit FM consider geographic dependencies, spatial autocorrelation, and the influence of topological relationships. They capture flows within and between regions, supporting a deeper understanding of spatial phenomena, such as migration patterns, transportation networks, and environmental dynamics. By embedding spatial structures into their learning process, these models enable more accurate and context-aware predictions, making them particularly valuable for geospatial health surveillance and risk assessment. Their ability to model spatial relationships enhances not only predictive performance but also the interpretability of spatial patterns, facilitating more informed decision-making in domains where geography plays a crucial role.

Beyond their predictive capabilities, geospatially explicit FM enable advanced geospatial question answering through dialogue-based interaction. These models allow users to pose complex spatial queries and receive insightful responses that incorporate feature selection, model architecture design, and coding explanations. Unlike conventional geographic information systems (GIS), which purely rely on domain expertise to extract meaningful insights, geospatially explicit FM enhance accessibility by enabling natural language queries. This capability extends to

interactive model refinement, where users can iteratively improve results by specifying relevant features or requesting tailored explanations. Thus, geospatial question answering serves as an intuitive interface between human expertise and machine intelligence, making it a powerful method for decision support in spatial health analysis. This integration of generative AI with geospatial analytics marks a significant step toward democratising spatial data science and expanding its utility across disciplines.

Training data enhancement and XAI

Generative AI for spatial data presents transformative opportunities by enhancing both the size and quality of training datasets, ultimately improving predictive performance. Many geospatial applications suffer from data sparsity, noise, and biases, limiting the effectiveness of traditional machine learning approaches. By synthesising realistic spatial data, generative AI can address these challenges by augmenting existing datasets, filling gaps in underrepresented areas, and refining label quality. This is particularly valuable for many problems in health surveillance, where high-quality labelled datasets are typically scarce and expensive to acquire. Furthermore, generative AI can create counterfactual scenarios to assess policy impacts, simulate climate change effects, or explore urban growth under different conditions. These capabilities not only enhance model robustness, but also enable more comprehensive and fair analysis, paving the way for more reliable geospatial AI in scientific research and policy-making.

Despite these advancements, a critical challenge in geospatial AI remains the explainability, transparency, and interpretability of a model's learning process and its outputs. The complex, high-dimensional nature of spatial data—combined with the opacity of deep learning models—makes it difficult to understand how predictions are generated and which spatial factors influence its outcomes. This lack of interpretability poses risks in high-stakes applications, such as targeted response to an epidemic or health policy-making, where decision-makers require trustworthy AI-driven insights. Addressing this issue requires the development of XAI methods tailored to geospatial contexts, including spatial attention maps, interpretable embeddings, and post-hoc explanation techniques. Furthermore, transparency in training data, model architecture, and decision-making processes is essential to ensure fairness and accountability. By advancing explainability, transparency and interpretability in geospatial AI, researchers can foster trust, improve user adoption, and ensure that spatial models are aligned with ethical and scientific standards.

Multimodal geospatial foundation models

To effectively leverage FMs in geospatial health, future research needs to focus on developing multimodal, geospatially enhanced models that can handle diverse types of data including text, numbers, images and time while incorporating geographic space. Spatially explicit learning could address the challenges of spatial dependencies and heterogeneity in health data. Fine-tuned FMs could then be used for diverse health applications, from dialogue-based question answering to incidence prediction. Pre-training on high-quality, diverse datasets is crucial for ensuring credibility and minimal bias. In this context, synthetic data generation offers a potential solution to data scarcity, though multimodal data generation and trustworthiness remain a challenge. FMs can be effectively used for data generation in this context but also benefit from it. A suitable setting for combining the two could be self-supervised learning. Another avenue for future studies concerns human-in-the-loop learning techniques like active learning.

Sigma Geography is a first step in this direction. Released in 2024 by a team of researchers from China's Institute of Geographic Sciences and Natural Resources Research (IGSNRR) and other organizations, all under the Chinese Academy of Sciences, and touted as "the world's first multimodal geographic science model" with a better understanding of the language patterns, domain-specific terminology and professional knowledge in the field of geography compared with general purpose LLMs, Sigma Geography can answer professional geographic questions, analyse geographic literature, query geographic data resources, and create thematic maps. It can also pair its generated text responses with geographical landscape images, thematic maps, or schematic charts to provide users with a more visual understanding of the information [119].

Credibility and usability

As the complexity of models grow, future research efforts must also focus on ensuring that outputs are credible and align with user expectations. Simultaneously, reducing geographic or other biases in both the training data and model output should be of utmost importance. Additionally, future research needs to concentrate on debiasing training data and evaluating potential biases learned by the model. Furthermore, techniques for preserving privacy in geospatial FMs are essential to protect sensitive health data from disclosure, especially under adversarial attacks. Security and privacy thus present another avenue for future research.

Even with once these challenges are solved, the effective use and interaction with geospatial FMs remains a key challenge. Recent years have been characterised by

generative AI systems that provide dialogue-based question answering capabilities through natural language-based interaction. The respective output can contain explanations, code, classification labels, prediction values or feature relevance. However, with geospatial and multimodal capabilities, models are not limited to textual interaction, opening up avenues for more nuanced interaction that also includes visual elements like maps and charts, video or even audio. This enables responses that are specifically tailored to the user, allowing for more detailed insights, better understanding of the problem at hand and improved decision-making.

New architectures

To date, transformers have achieved impressive results, but not without many flaws and limitations. Alternative methods, such as JEPA (Joint Embedding Predictive Architecture), are currently being explored to enable AI to attain human-level intelligence [120].

Recommendations and conclusions

The rapid rise of generative AI models presents a unique opportunity to revolutionise geospatial health and environmental health. Notable applications include the use of satellite data to predict malaria outbreaks, leveraging the multimodal nature of large data sources like geosocial media data as a basis for large-scale early warning systems, or AI-powered dashboards to manage epidemics for efficient and targeted decision-making. By integrating spatial analysis with AI-driven insights, these technologies can address complex challenges, including health surveillance or monitoring the spread of a specific disease. Advanced AI agents are already showing great potential and will be increasingly contributing to this vital integration in the coming months and years.

However, their adoption must be guided by faithfulness, ethical principles, inclusiveness, and sustainability to ensure a positive impact on society (Table 1). Addressing the challenges and embracing the opportunities will require continued innovation, collaboration, and a commitment to ethical and responsible development. By harnessing the power of these technologies, we can develop more effective health surveillance systems, improve public health outcomes, mitigate the impacts of newly emerging diseases, and create a more resilient health system.

Author contributions

MNKB conceived the manuscript's scope and direction, and invited BR, PK, DH, and MAB to contribute. BR, PK, DH, MAB, and MNKB all made contributions of equal importance to the paper and participated in its literature review, writing and revision. All authors read and approved the final version of the manuscript. Disclaimer: Views and opinions expressed are those of the author(s) only. Reference in the manuscript to any specific commercial product, process

or service by trade name, trademark, manufacturer or otherwise does not necessarily constitute or imply its endorsement, recommendation or favouring by the authors or the entities they are affiliated to, and shall not be used for commercial advertising or product endorsement purposes.

Funding

BR and DH's research was funded by the Austrian Science Fund (FWF) under the project GeoEpi (Grant Number I-5117). MNKB's GeoExposomics research described in this paper was funded by the International Society for Photogrammetry and Remote Sensing (ISPRS) under ISPRS Scientific Initiatives 2023 (SI2023 Awards).

Availability of data and materials

No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 3 March 2025 Accepted: 4 March 2025

Published online: 02 April 2025

References

- Meng X, Yan X, Zhang K, Liu D, Cui X, Yang Y, et al. The application of large language models in medicine: a scoping review. *iScience*. 2024;27(5):109713.
- World Health Organization. 2025. <https://www.who.int/europe/initiatives/one-health>. Accessed 23 Feb 2025.
- Peters M, Zeeb H. Availability of open data for spatial public health research. *GMS German Med Sci*. 2022;20:Doc01.
- Orozco-Acosta E, Riebler A, Adin A, Ugarte MD. A scalable approach for short-term disease forecasting in high spatial resolution areal data. *Biom J*. 2023;65(8):2300096.
- Nguyen D, Kay F, Tan J, Yan Y, Ng YS, Iyengar P, et al. Deep learning-based COVID-19 pneumonia classification using chest CT images: model generalizability. *Front Artif Intell*. 2021;4:694875.
- El Morr C, Ozdemir D, Asdaah Y, Saab A, El-Lahib Y, Sokhn ES. AI-based epidemic and pandemic early warning systems: a systematic scoping review. *Health Inform J*. 2024;30(3):14604582241275844.
- Ballester J, Quijal-Zamorano M, Méndez Turrubiates RF, Pegenaute F, Herrmann FR, Robine JM, et al. Heat-related mortality in Europe during the summer of 2022. *Nat Med*. 2023;29(7):1857–66.
- Hay SI, Snow RW. The Malaria atlas project: developing global maps of malaria risk. *PLoS Med*. 2006;3(12):e473.
- MOL. Map of Life. 2025. <https://mol.org/>. Accessed 23 Feb 2025.
- Kang M, Ko E, Mersha TB. A roadmap for multi-omics data integration using deep learning. *Brief Bioinform*. 2022;23(1):bbab454. <https://doi.org/10.1093/bib/bbab454>.
- Amoroso N, Cilli R, Maggipinto T, Monaco A, Tangaro S, Bellotti R. Satellite data and machine learning reveal a significant correlation between NO2 and COVID-19 mortality. *Environ Res*. 2022;204:111970.
- Sufi F. An innovative way of Analyzing COVID topics with LLM. *J Econ Technol*. 2024. <https://doi.org/10.1016/j.ject.2024.11.004>.
- Loché Fernández-Ahúja JM, Fernández Martínez JL. Effects of climate variables on the COVID-19 outbreak in Spain. *Int J Hyg Environ Health*. 2021;234:113723.
- Google. How we're using AI to make emergency healthcare more accessible. 2025. <https://blog.google/technology/health/google-ai-healthcare-accessibility/>. Accessed 21 Feb 2025.

15. Sakib N, Hyer K, Dobbs D, Peterson L, Jester DJ, Kong N, et al. A GIS enhanced data analytics approach for predicting nursing home hurricane evacuation response. *Health Inf Sci Syst.* 2022;10(1):28.
16. Al Nazi Z, Peng W. Large language models in healthcare and medical domain: a review. *Informatics.* 2024;11:57. <https://doi.org/10.3390/informatics11030057>.
17. Schneider J, Meske C, Kuss P. Foundation models. *Bus Inf Syst Eng.* 2024;66(2):221–31.
18. Brown T, Mann B, Ryder N, Subbiah M, Kaplan JD, Dhariwal P, et al. Language models are few-shot learners. In: *advances in neural information processing systems*. Newry: Curran Associates, Inc.; 2020. p. 1877–901.
19. Tan Y, Min D, Li Y, Li W, Hu N, Chen Y, et al. Can ChatGPT replace traditional KBQA models? An in-depth analysis of the question answering performance of the GPT LLM family. In: Payne TR, Presutti V, Qi G, Poveda-Villalón M, Stoilos G, Hollink L, et al., editors. *The semantic web – ISWC 2023*. Cham: Springer Nature Switzerland; 2023. p. 348–67.
20. Liu X, Zhou T, Wang C, Wang Y, Wang Y, Cao Q, et al. Toward the unification of generative and discriminative visual foundation model: a survey. *Vis Comput.* 2024. <https://doi.org/10.1007/s00371-024-03608-8>.
21. Li Z, Ning H. Autonomous GIS: the next-generation AI-powered GIS. *Int J Digit Earth.* 2023;16(2):4668–86.
22. Mai G, Huang W, Sun J, Song S, Mishra D, Liu N, et al. On the Opportunities and Challenges of Foundation Models for Geospatial Artificial Intelligence. *arXiv*; 2023.
23. Radford A, Narasimhan K, Salimans T, Sutskever I. Improving Language Understanding by Generative Pre-Training. 2018.
24. Ouyang L, Wu J, Jiang X, Almeida D, Wainwright C, Mishkin P, et al. Training language models to follow instructions with human feedback. *Adv Neural Inf Process Syst.* 2022;35:27730–44.
25. Cordes J, Castro MC. Spatial analysis of COVID-19 clusters and contextual factors in New York City. *Spat Spatiotemporal Epidemiol.* 2020;34:100355.
26. Chinazzi M, Davis JT, Ajelli M, Gioannini C, Litvinova M, Merler S, et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science.* 2020;368(6489):395–400.
27. Arthur RF, Gurley ES, Salje H, Bloomfield LSP, Jones JH. Contact structure, mobility, environmental impact and behaviour: the importance of social forces to infectious disease dynamics and disease ecology. *Philos Trans Royal Soc B Biol Sci.* 2017;372(1719):20160454.
28. Koh K, Kamel Boulos MN, Zheng G, Zhang H, Iyyanki MV, Bwambale B, et al. A proof-of-concept metadata catalogue and online portal of earth observation datasets for health research in exposomics. *J Acad Public Health.* 2025. <https://doi.org/10.70542/rcj-japh-art-1wjzmn4>.
29. GeoExposomics Project Team. Geoexposomics Web Portal. 2025. <https://geoexposomics.org/>. Accessed 21 Feb 2025.
30. Wiesinger J, Marlow P, Vuskovic V. Agents. 2024. <https://www.kaggle.com/whitepaper-agents>. Accessed 21 Feb 2025.
31. Anthropic. Building Effective Agents. 2024. <https://www.anthropic.com/research/building-effective-agents>. Accessed 21 Feb 2025.
32. Crockett K. What Is Agentic AI? 2025. <https://transmitter.ieee.org/what-is-agentic-ai/>. Accessed 21 Feb 2025.
33. Olawade DB, Wada OJ, David-Olawade AC, Kunonga E, Abaire O, Ling J. Using artificial intelligence to improve public health: a narrative review. *Front Public Health.* 2023. <https://doi.org/10.3389/fpubh.2023.1196397>.
34. Kamel Boulos MN, Peng G, VoPham T. An overview of GeoAI applications in health and healthcare. *Int J Health Geogr.* 2019;18(1):7.
35. Devlin J, Chang MW, Lee K, Toutanova K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In: *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)*. Minneapolis, Minnesota: Association for Computational Linguistics. 2019; 4171–86.
36. OpenAI, Achiam J, Adler S, Agarwal S, Ahmad L, Akkaya I, et al. GPT-4 Technical Report. *arXiv*; 2024.
37. Dubey A, Jauhri A, Pandey A, Kadian A, Al-Dahle A, Letman A, et al. The Llama 3 Herd of Models. *arXiv*; 2024.
38. Luccioni A, Viviano J. What's in the Box? An Analysis of Undesirable Content in the Common Crawl Corpus. In: Zong C, Xia F, Li W, Navigli R, editors. *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*. Online: Association for Computational Linguistics. 2021; 182–9.
39. Pu X, Gao M, Wan X. Summarization Is (Almost) Dead. *arXiv*; 2023.
40. Hendy A, Abdelrehim M, Sharaf A, Raunak V, Gabr M, Matsushita H, et al. How Good Are GPT Models at Machine Translation? A Comprehensive Evaluation. *arXiv*; 2023.
41. Wang Z, Pang Y, Lin Y. Large Language Models Are Zero-Shot Text Classifiers. *arXiv*; 2023.
42. Xu J, Fei H, Pan L, Liu Q, Lee ML, Hsu W. Faithful logical reasoning via symbolic chain-of-thought. *arXiv preprint arXiv:240518357*. 2024.
43. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention is all you need. In: *Advances in neural information processing systems*. Newry: Curran Associates, Inc.; 2017.
44. Kaplan J, McCandlish S, Henighan T, Brown TB, Chess B, Child R, et al. Scaling Laws for Neural Language Models. *arXiv*; 2020.
45. Wei J, Wang X, Schuurmans D, Bosma M, Ichter B, Xia F, et al. Chain-of-Thought Prompting Elicits Reasoning in Large Language Models. In: *Proceedings of the 36th International Conference on Neural Information Processing Systems*. Red Hook, NY, USA: Curran Associates Inc. 2024; 24824–37. (NIPS '22).
46. Chockalingam A. Accelerate DeepSeek Reasoning Models With NVIDIA GeForce RTX 50 Series AI PCs. 2025. <https://blogs.nvidia.com/blog/deepseek-r1-rtx-ai-pc/>. Accessed 23 Feb 2025.
47. Wornow M, Xu Y, Thapa R, Patel B, Steinberg E, Fleming S, et al. The shaky foundations of large language models and foundation models for electronic health records. *NPJ Digit Med.* 2023;6(1):1–10.
48. Agrawal M, Hegselmann S, Lang H, Kim Y, Sontag D. Large Language Models Are Few-Shot Clinical Information Extractors. In: Goldberg Y, Kozareva Z, Zhang Y, editors. *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*. Abu Dhabi, United Arab Emirates: Association for Computational Linguistics. 2022; 1998–2022.
49. Singhal K, Azizi S, Tu T, Mahdavi SS, Wei J, Chung HW, et al. Large language models encode clinical knowledge. *Nature.* 2023;620(7972):172–80.
50. Chintagunta B, Katariya N, Amatriain X, Kannan A. Medically Aware GPT-3 as a Data Generator for Medical Dialogue Summarization. In: Shivade C, Gangadharaiyah R, Gella S, Konam S, Yuan S, Zhang Y, et al., editors. *Proceedings of the Second Workshop on Natural Language Processing for Medical Conversations*. Online: Association for Computational Linguistics. 2021; 66–76.
51. Lehman E, Hernandez E, Mahajan D, Wulff J, Smith MJ, Ziegler Z, et al. Do We Still Need Clinical Language Models? In: *Proceedings of the Conference on Health, Inference, and Learning*. PMLR; 2023; 578–97.
52. Moradi M, Blagec K, Haberl F, Samwald M. GPT-3 Models Are Poor Few-Shot Learners in the Biomedical Domain. *arXiv*; 2022.
53. Borazio F, Croce D, Gambosi G, Basili R, Margiotta D, Sciaella A, et al. Semi-Automatic Topic Discovery and Classification for Epidemic Intelligence via Large Language Models. In: Afli H, Bouamor H, Casagran CB, Ghannay S, editors. *Proceedings of the Second Workshop on Natural Language Processing for Political Sciences @ LREC-COLING 2024*. Torino, Italia: ELRA and ICCL. 2024; 68–84.
54. Hu Y, Mai G, Cundy C, Choi K, Lao N, Liu W, et al. Geo-knowledge-guided GPT models improve the extraction of location descriptions from disaster-related social media messages. *Int J Geogr Inf Sci.* 2023;37(11):2289–318.
55. Fernandez A, Dube S. Core Building Blocks: Next Gen Geo Spatial GPT Application. *arXiv*; 2023.
56. Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. 3rd International Conference on Learning Representations (ICLR 2015). 2015.
57. He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016; 770–8.
58. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Commun ACM.* 2017;60(6):84–90.
59. Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, et al. An Image Is Worth 16x16 Words: Transformers for Image Recognition at Scale. *arXiv*; 2021.

60. Cong Y, Khanna S, Meng C, Liu P, Rozi E, He Y, et al. SatMAE: Pre-Training Transformers for Temporal and Multi-Spectral Satellite Imagery. In: Proceedings of the 36th International Conference on Neural Information Processing Systems. Red Hook, NY, USA: Curran Associates Inc. 2024; 197–211. (NIPS '22).
61. Christie G, Fendley N, Wilson J, Mukherjee R. Functional map of the world. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018; 6172–80.
62. Shi P, Qiu J, Abaxi SMD, Wei H, Lo FPW, Yuan W. Generalist vision foundation models for medical imaging: a case study of segment anything model on zero-shot medical segmentation. *Diagnostics*. 2023;13(11):1947.
63. Prabhod KJ, Gadhiraju A. Foundation models in medical imaging: revolutionizing diagnostic accuracy and efficiency. *J Artif Intell Res Appl*. 2024;4(1):471–511.
64. Hartsock I, Rasool G. Vision-language models for medical report generation and visual question answering: a review. *Front Artif Intell*. 2024;7:1430984.
65. Vatsavai RR. Geospatial Foundation Models: Recent Advances and Applications. In: Proceedings of the 12th ACM SIGSPATIAL International Workshop on Analytics for Big Geospatial Data. New York, NY, USA: Association for Computing Machinery. 2024; 30–3. (BigSpatial '24).
66. Jakubik J, Roy S, Phillips CE, Fraccaro P, Godwin D, Zadrozny B, et al. Foundation Models for Generalist Geospatial Artificial Intelligence. arXiv; 2023.
67. Reed CJ, Gupta R, Li S, Brockman S, Funk C, Clipp B, et al. Scale-MAE: A Scale-Aware Masked Autoencoder for Multiscale Geospatial Representation Learning. In: 2023 IEEE/CVF International Conference on Computer Vision (ICCV). 2023; 4065–76.
68. Wanyan X, Seneviratne S, Shen S, Kirley M. Extending Global-local View Alignment for Self-supervised Learning with Remote Sensing Imagery. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024; 2443–53.
69. Balakrishnan S, Preetam Raj PM, Somasekar J, Kumar KV, Amutha S, Sangeetha A. Remote sensing data-based satellite image analysis in water quality detection for public health data modelling. *Remote Sens Earth Syst Sci*. 2024;7(4):532–41.
70. Yu H, Zahidi I. Environmental hazards posed by mine dust, and monitoring method of mine dust pollution using remote sensing technologies: an overview. *Sci Total Environ*. 2023;864:161135.
71. Knoblauch S, Li H, Lautenbach S, Elshiaty Y, de Rocha AAA, Resch B, et al. Semi-supervised water tank detection to support vector control of emerging infectious diseases transmitted by Aedes Aegypti. *Int J Appl Earth Obs Geoinf*. 2023;119:103304.
72. Li J, Li D, Xiong C, Hoi S. BLIP: Bootstrapping Language-Image Pre-training for Unified Vision-Language Understanding and Generation. arXiv; 2022.
73. Li J, Hui B, Qu G, Yang J, Li B, Li B, et al. Can LLM already serve as a database interface? A big bench for large-scale database grounded text-to-SQLs. *Adv Neural Inf Process Syst*. 2023;36:42330–57.
74. Wang W, Bao H, Dong L, Björck J, Peng Z, Liu Q, et al. Image as a Foreign Language: BEiT pretraining for all vision and vision-language tasks. arXiv; 2022.
75. Liu H, Li C, Wu Q, Lee YJ. Visual instruction tuning. arXiv; 2023.
76. Team G, Anil R, Borgeaud S, Alayrac JB, Yu J, Soricut R, et al. Gemini: a family of highly capable multimodal models. arXiv; 2024.
77. Baevski A, Hsu WN, Xu Q, Babu A, Gu J, Auli M. Data2vec: a general framework for self-supervised learning in speech, vision and language. arXiv; 2022.
78. Yang L, Xu S, Sellergren A, Kohlberger T, Zhou Y, Ktena I, et al. Advancing multimodal medical capabilities of gemini. arXiv; 2024.
79. Thawkar O, Shaker A, Mullappilly SS, Cholakkal H, Anwer RM, Khan S, et al. XrayGPT: chest radiographs summarization using medical vision-language models. arXiv; 2023.
80. Liu C, Jin Y, Guan Z, Li T, Qin Y, Qian B, et al. Visual-language foundation models in medicine. *Vis Comput*. 2024. <https://doi.org/10.1007/s00371-024-03579-w>.
81. Wang X, Wang Z, Gao X, Zhang F, Wu Y, Xu Z, et al. Searching for Best Practices in Retrieval-Augmented Generation. arXiv; 2024.
82. Stolerman LM, Clemente L, Poirier C, Parag KV, Majumder A, Masyn S, et al. Using digital traces to build prospective and real-time county-level early warning systems to anticipate COVID-19 outbreaks in the United States. *Sci Adv*. 2023;9(3):eabq0199.
83. Knoblauch S, Su Yin M, Chatrinan K, de Aragão Rocha AA, Haddawy P, Biljecki F, et al. High-resolution mapping of urban Aedes Aegypti immature abundance through breeding site detection based on satellite and street view imagery. *Sci Rep*. 2024;14(1):18227.
84. Janowicz K, Gao S, McKenzie G, Hu Y, Bhaduri B. GeoAI: spatially explicit artificial intelligence techniques for geographic knowledge discovery and beyond. *Int J Geogr Inf Sci*. 2020;34(4):625–36.
85. Yan B, Janowicz K, Mai G, Zhu R. A spatially explicit reinforcement learning model for geographic knowledge graph summarization. *Trans GIS*. 2019;23(3):620–40.
86. Chu G, Potetz B, Wang W, Howard A, Song Y, Brucher F, et al. Geo-Aware Networks for Fine-Grained Recognition. arXiv; 2019.
87. Honzák K, Schmidt S, Resch B, Ruthensteiner P. Contextual enrichment of crowds from mobile phone data through multimodal geo-social media analysis. *ISPRS Int J Geoinf*. 2024;13(10):350.
88. Quiñones S, Goyal A, Ahmed ZU. Geographically weighted machine learning model for untangling spatial heterogeneity of type 2 diabetes mellitus (T2D) prevalence in the USA. *Sci Rep*. 2021;11(1):6955.
89. Kogan NE, Clemente L, Liautaud P, Kaashoek J, Link NB, Nguyen AT, et al. An early warning approach to monitor COVID-19 activity with multiple digital traces in near real time. *Sci Adv*. 2021;7(10):eabd6989.
90. Arifi D, Resch B, Santillana M, Guan WW, Knoblauch S, Lautenbach S, et al. Geosocial media's early warning capabilities across US county-level political clusters: observational study. *JMIR Infodemiology*. 2025;5:e58539.
91. Rundle AG, Bader MDM, Mooney SJ. Machine learning approaches for measuring neighborhood environments in epidemiologic studies. *Curr Epidemiol Rep*. 2022;9(3):175–82.
92. Savin I, Ershova K, Kurdyumova N, Ershova O, Khomenko O, Danilov G, et al. Healthcare-associated ventriculitis and meningitis in a neuro-ICU: incidence and risk factors selected by machine learning approach. *J Crit Care*. 2018;45:95–104.
93. He Y, Huang F, Jiang X, Nie Y, Wang M, Wang J, et al. Foundation model for advancing healthcare: challenges, opportunities, and future directions. arXiv; 2024.
94. Willemink MJ, Koszek WA, Hardell C, Wu J, Fleischmann D, Harvey H, et al. Preparing medical imaging data for machine learning. *Radiology*. 2020;295(1):4–15.
95. Mumuni A, Mumuni F. Data augmentation: a comprehensive survey of modern approaches. *Array*. 2022;16:100258.
96. Miao L, Last M, Litvak M. Twitter Data Augmentation for Monitoring Public Opinion on COVID-19 Intervention Measures. In: Verspoor K, Cohen KB, Conway M, de Bruijn B, Dredze M, Mihalcea R, et al., editors. Proceedings of the 1st Workshop on NLP for COVID-19 (Part 2) at EMNLP 2020. Online: Association for Computational Linguistics; 2020.
97. Radosavovic I, Dollár P, Girshick R, Gkioxari G, He K. Data Distillation: Towards Omni-Supervised Learning. arXiv; 2017.
98. Duives DC, Wang G, Kim J. Forecasting pedestrian movements using recurrent neural networks: an application of crowd monitoring data. *Sensors*. 2019;19(2):382.
99. Zhou Y, Guo C, Wang X, Chang Y, Wu Y. A Survey on Data Augmentation in Large Model Era. arXiv; 2024.
100. Yilmaz L, Liu B. Model credibility revisited: concepts and considerations for appropriate trust. *J Simul*. 2022;16(3):312–25.
101. Li YF, Wang H, Sun M. ChatGPT-like large-scale foundation models for prognostics and health management: a survey and roadmaps. *Reliab Eng Syst Saf*. 2024;243:109850.
102. Wang J, Oh J, Wang H, Wiens J. Learning Credible Models. In: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York, NY, USA: Association for Computing Machinery. 2018; 2417–26. (KDD '18).
103. Bommasani R, Hudson DA, Adeli E, Altman R, Arora S, von Arx S, et al. On the opportunities and risks of foundation models. arXiv; 2022.
104. Zhang S, Roller S, Goyal N, Artetxe M, Chen M, Chen S, et al. OPT: open pre-trained transformer language models. arXiv; 2022.
105. Manvi R, Khanna S, Burke M, Lobell D, Ermon S. Large language models are geographically biased. arXiv; 2024.

106. Liu Z, Janowicz K, Cai L, Zhu R, Mai G, Shi M. Geoparsing: solved or biased? An evaluation of geographic biases in geoparsing. *AGILE GISci Ser.* 2022;3:1–13.
107. Navigli R, Conia S, Ross B. Biases in large language models: origins, inventory, and discussion. *J Data Inf Qual.* 2023;15(2):10:1–10:21.
108. Marani BE, Hanini M, Malayarukil N, Christodoulidis S, Vakalopoulou M, Ferrante E. ViG-Bias: visually grounded bias discovery and mitigation. 2025; 414–29.
109. Rao J, Gao S, Mai G, Janowicz K. Building Privacy-Preserving and Secure Geospatial Artificial Intelligence Foundation Models (Vision Paper). In: *Proceedings of the 31st ACM International Conference on Advances in Geographic Information Systems*. New York, NY, USA: Association for Computing Machinery. 2023; 1–4. (SIGSPATIAL '23).
110. Perez F, Ribeiro I. Ignore previous prompt: attack techniques for language Models. *arXiv*; 2022.
111. Schlarman C, Hein M. On the adversarial robustness of multi-modal foundation models. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2023; 3677–85.
112. Haltaufderheide J, Ranisch R. The ethics of ChatGPT in medicine and healthcare: a systematic review on Large Language Models (LLMs). *NPJ Digit Med.* 2024;7(1):183.
113. Vishwanath PR, Tiwari S, Naik TG, Gupta S, Thai DN, Zhao W, et al. Faithfulness Hallucination Detection in Healthcare AI. In: *Artificial Intelligence and Data Science for Healthcare: Bridging Data-Centric AI and People-Centric Healthcare*. 2024.
114. Zhang P, Shi J, Kamel Boulos MN. Generative AI in medicine and healthcare: moving beyond the 'peak of inflated expectations.' *Future Internet.* 2024;16(12):462.
115. Park SH, Suh CH, Lee JH, Kahn CE Jr, Moy L. Minimum reporting items for clear evaluation of accuracy reports of large language models in healthcare (MI-CLEAR-LLM). *Korean J Radiol.* 2024;25(10):865.
116. Alhamoud K, Alshammari S, Tian Y, Li G, Torr P, Kim Y, et al. Vision-language models do not understand negation. *arXiv preprint arXiv:250109425*. 2025.
117. Symptoma GmbH. Digital Health Assistant & Symptom Checker | Symptoma. 2024. <https://www.symptoma.com/>. Accessed 22 Feb 2025.
118. Liu Y, Jain A, Eng C, Way DH, Lee K, Bui P, et al. A deep learning system for differential diagnosis of skin diseases. *Nat Med.* 2020;26(6):900–8.
119. China's IGSNRR, CAS. World's First Geographic Multimodal Premiered in China. 2024. http://english.igsnr.cas.cn/newsroom/news/202409/t20240923_690360.html. Accessed 21 Feb 2025.
120. Kuka V. What is JEPA?. 2024. <https://www.turingpost.com/p/jepa>. Accessed 21 Feb 2025.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.