Transactions in GIS

WILEY

Transactions in GIS

RESEARCH ARTICLE

GISphere Knowledge Graph for Geography Education: Recommending Graduate Geographic Information System/Science Programs

Zhining Gu[1,2] | Wenwen Li[1] | Bing Zhou[2,3] | Yikang Wang[2,4] | Yanbing Chen[2,5] | Shan Ye[2,6] | Kejin Wang[2,7] | Hongkai Gu[2,8] | Yuhao Kang[2,9,10]

[1]School of Geographical Sciences & Urban Planning, Arizona State University, Tempe, Arizona, USA | [2]GISphere Corporation, South Carolina, SC, USA | [3]Department of Geography, Texas A&M University, College Station, Texas, USA | [4]Centre of Advanced Spatial Analysis, University College London, London, UK | [5]Department of Geography, University of Wisconsin-Madison, Madison, Wisconsin, USA | [6]National Institute of Clean-and-Low-Carbon Energy, Changping, Beijing, China | [7]Department of Environmental Sciences, Louisiana State University, Baton Rouge, Louisiana, USA | [8]Department of Geography, Ghent University, Ghent, Belgium | [9]GISense Lab, Department of Geography and the Environment, The University of Texas at Austin, Austin, Texas, USA | [10]Department of Geography, University of South Carolina, Columbia, South Carolina, USA

Correspondence: Yuhao Kang (yuhao.kang@austin.utexas.edu)

Received: 9 July 2024 | **Revised:** 7 November 2024 | **Accepted:** 17 November 2024

Funding: This work was supported by Cartography and Geographic Information Society (CaGIS) Rising Award.

Keywords: education | GIS | information retrieval | knowledge graph | large language model | program recommendation

ABSTRACT

The growing global interest in Geographic Information System/Science (GIS) programs has led to an increased demand for higher education in this field. However, students often struggle to identify suitable programs and faculty due to the overwhelming options and the lack of personalized guidance. This paper presents GISphere-KG, an AI-powered platform based on the GISphere project. It combines knowledge graph (KG) and large language models (LLMs) to enhance the search and recommendation of GIS-related graduate programs. GISphere-KG offers four key features: (1) interactive conversation that provides natural language responses to applicants' inquiries; (2) efficient information retrieval through semantic relationships built within the KG; (3) discovery of professors whose research interests align with those of the applicants, offering more choices within specific research fields; and (4) personalized program recommendations tailored to applicants' academic and career developments. Our platform aims to provide a user-friendly tool that assists prospective students in achieving their career goals and enriching the geography community by attracting more talent and promoting global geography education.

1 | Introduction

The increasing importance of Geographic Information System/Science (GIS), cartography, and geospatial analytics has made a prominent expansion to various domains such as urban and environmental planning, remote sensing, disaster management, and sustainable development (Goodchild and Longley 2021; Rosenkrantz et al. 2021; Wu et al. 2023). These successful practices have attracted more students to pursue graduate programs in GIS to enhance their professional skills in geography and other relevant fields (Bodycott 2009; Bodycott and Lai 2012; Bednarz, Heffron, and Huynh 2013; Bohman 2014; Cebolla-Boado, Hu, and Soysal 2018; de Wit, Minaeva, and Wang 2022). These graduate programs include academic programs (e.g., master's and doctoral programs) and non-academic ones (e.g., certificate programs, online programs). Given this context, numerous higher education programs have been designed for prospective students to achieve their career goals. For easy access to these programs, several educational resources, beyond program websites, have been developed to assist applicants in program navigation. For instance, the *AAG Guide to Geography Programs and Opportunities*[1] provides valuable resources for students to explore various programs in geography. Additionally, GISphere,[2] a non-profit platform led by graduate students and young scholars from GIS-related fields, offers opportunities for discovering and retrieving GIS graduate programs. Its website summarizes comprehensive global information on over 600 GIS programs across more than 97 countries and regions as of December 2023 (Wang et al. 2023).

Despite the availability of these valuable educational resources, applicants still face several challenges. First, abundant global opportunities are too overwhelming for applicants to select suitable problems accurately fitting their interests across various websites (Wang et al. 2023). Second, the lack of detailed information and personalized guidance hinders them from discovering programs that satisfy their specific interests and preferences (Naz, Siddiqua, and Asim 2019). For instance, when considering future programs, applicants may take factors such as location and research interests into account (Wang et al. 2023; Lei and Chuang 2010). Applicants also need guidance on how to explore and utilize the available data more effectively so that they can dive deep into professors' research interests and understand the range and prospects of research fields (A Tactical Guide to Graduate Student Recruitment 2024; Paris, Birnbaum, and Dix 2020). All of these emphasize the necessity for a more personalized and insightful decision-making process while exploring different graduate programs.

Consequently, we identify a significant gap between educational resources and applicants in information retrieval. How to assist applicants in navigating programs and finding the most relevant opportunities based on their academic and career goals? In this paper, building upon the GISphere project, we introduce GISphere-KG, an AI-powered platform supporting the search and recommendation for GIS-related graduate programs. It incorporates two major technologies: knowledge graph (KG) and large language models (LLMs). We ask the following research questions: (1) How can the platform facilitate the search process, making it easier and more intuitive for applicants to access program resources that align with their research interests? (2) How can the platform discover professors with similar research interests? (3) How can the platform offer personalized program recommendations for applicants by considering their preferences? Using our GISphere-KG, applicants can efficiently identify programs that align with their specific needs such as country, university, and research interests. It not only saves the time and energy of applicants but also benefits the entire geography community in attracting the next generation of geographers.

Such a powerful platform was implemented by incorporating KG and LLMs. As a semantic web technology, KG organizes structured and unstructured data into a graph (W. Li et al. 2022). In the KG, each concept is considered as a node, and the semantic relationship between two concepts is represented as an edge connecting the two semantic nodes, which enables knowledge exchange and data integration through standardization (Li et al. 2023). This cutting-edge technology has shown extraordinary performance in various knowledge discovery tasks, such as link prediction (Wang, Qiu, and Wang 2021), question and answering (QA) (Yani and Krisnadhi 2021), and information retrieval (Reinanda, Meij, and de Rijke 2020). In the field of GIS, numerous practices has also leveraged KG, such as world KG for the semantic representation of geographic entities (Dsouza et al. 2021); a knowledge graph visualization platform, GeoGraphVis, for disaster relief (Li et al. 2023); an electric vehicle (EV) knowledge graph for smart transportation system (Qi et al. 2023); and a cartographic geovisualization platform with the support of semantic web technologies (Huang and Harrie 2019). However, KG techniques have not received sufficient attention in GIS and Geography education, and there has been no practical applications that use KG yet. Integrating KG offers new opportunities to dynamically link research interests among different professors, allowing for the discovery of similar research direction that matches students with suitable opportunities. More importantly, the strong capability of KG on heterogeneous data linkages enables GIS educational program retrieval and promotion across diverse datasets comprehensively.

In addition to KG, the rapid advancement of LLMs, such as generative pre-trained transformer (GPT) models (Brown et al. 2020) and bidirectional encoder representations from transformers (BERT) (Devlin et al. 2018), has led to outstanding performance on a wide range of natural language processing (NLP) tasks. These include but are not limited to information retrieval (Zhu et al. 2023), QA (Ishwari et al. 2019), and text generation (J. Li et al. 2022). ChatGPT, one of the most powerful interactive platforms supported by various LLMs, developed by OpenAI (2022, 2023), enables a wide range of NLP tasks through a question-answering process (OpenAI 2022). It takes users' questions in a natural language format and responds with easy-to-understand answers in the same way. This can simplify the question-answering process, offering students a convenient and efficient way to access important information and program recommendations. Although ChatGPT, powered by LLMs, has displayed extraordinary performance in natural language understanding, it is not omnipotent. Researchers have observed that LLMs can sometimes produce inaccuracies or misleading information, a phenomenon referred to as "hallucination" (Tonmoy et al. 2024). Thus, errors in response accuracy and context interpretation can occur because of the limitations inherent in these LLMs.

Our proposed GISphere-KG platform has the following characteristics:

1. *Interactive conversation*: Utilizing LLMs, the platform enables a conversational interface where applicants can input their questions and receive responses in natural language.
2. *Efficient information retrieval*: Through the application of KG, GISphere-KG connects diverse attributes through semantic relationships, facilitating quick and efficient retrieval of information about GIS programs. It should be acknowledged that our KG is built based on diverse GIS/Cartography-related information collected from external sources hosted on the GISphere platform. New data sources related to GIS/Cartography programs will be incorporated into updates of the KG on the GISphere platform.
3. *Discovery of professors with similar research interests*: The platform uses KG to connect professors through semantic

similarity based on their research interests. It allows applicants to find professors who match their own interests, thereby offering a comprehensive overview of the research field.

4. *Program recommendation*: By leveraging LLMs, GISphere-KG can analyze applicants' interests and recommend GIS-related programs that align with their academic and career goals based on semantic similarity calculations.

By integrating these cutting-edge technologies, GISphere-KG aims to provide a user-friendly tool based on an up-to-date global database of GIS programs. We showcase how advanced technologies such as KG and LLMs can be utilized to enrich GIS and geography education and help recruit future generations of geographers. The GISphere-KG facilitates personalized matches between applicants and the programs or professors that can best support their future careers. It should be acknowledged that, although there are various programs, including both academic (research-based) and non-academic (non-research-based) ones, we primary focus on academic master's and doctoral programs due to the current dataset coverage. In these programs, graduate students often need to finish thesis and dissertation as graduate requirements beyond training practical skills. Differently, "non-academic" programs aim to cultivate students pursing their industry careers instead of academia. In this category, students from non-academic graduate programs are probably not required to finish dissertation for graduation. Furthermore, our work did not include the information regarding research experience or funding package. In the future, we plan to expand our work to include non-academic programs and collect more data sources covering diverse information. Despite the fact that we solely focus on GIS, our GISphere-KG showcases the potential of advanced technology to enhance education within this field and might be utilized to benefit broader domains such as geography and social science.

2 | Related Work

2.1 | Rising Importance of GIS Education

GIS has evolved into an essential component across various domains (Longley 2000; Baker and Witham Bednarz 2003; Rickles, Ellul, and Haklay 2017; Jakab, Ševcík, and Grežo 2017). Specifically, GIS aids city design and management in urban planning (Batty 2018), facilitates environmental monitoring in remote sensing (Weng 2010), supports biodiversity conservation in ecology (Boyd and Foody 2011), and enhances cultural research in digital humanities (Bodenhamer, Corrigan, and Harris 2010). The rising demand for specialized knowledge and skills in these areas has led to a significant increase in educational programs specializing in GIS (Revell and Benfield 2022). These programs include master's and doctoral programs, professional master's courses, and certificate programs (DeMers, Kerski, and Sroka 2021), as well as non-research degrees and online programs. They aim to align students' aspirations with the dynamic and multifaceted needs of industry (Mathews and Wikle 2019; Tian 2017), offering a diverse range of specializations. They focus on areas like GIS methods, spatial analysis, cartography and geovisualization, and domain-specific applications in urban planning and environmental management (Wang et al. 2023).

Despite the significant growth of GIS in both industry and academia, prospective students still face challenges in efficiently searching for and receiving personalized recommendations for programs that align with their interests and career goals. Massive amounts of fragmented information are spreading across various websites, often lacking proper links, in multiple languages, and much of it is not publicly available on the internet. This makes it difficult for prospective students worldwide to find comprehensive and up-to-date program details. Although the GISphere database offers valuable information on GIS programs and professors from more than 400 universities globally (GISphere Contributors 2023), the student experience could be significantly enhanced by providing personalized program recommendations that closely align with individual research interests and career aspirations. Additionally, the ability to search for relevant information across universities would be beneficial. This gap highlights the necessity for a systematic approach to gathering and sharing program information, which is crucial for students navigating the complexities of GIS educational paths and for the broader geographical community aiming to attract new talent (Tate and Jarvis 2017).

2.2 | Knowledge Graphs in Geographical Applications

The concept of "Knowledge Graph" was originated back in 1980s for semantic networks (Knowledge Graph 2024). In 2012, Google introduced the Knowledge Graph built on DBpedia and Freebase to facilitate the information retrieval (Knowledge Graph 2024). A knowledge graph represents data in a human and machine-readable way, enabling the scalability and timely updates of new cross-domain knowledge (Janowicz et al. 2020, 2022; Qi et al. 2023). It has been increasingly applied to various geographical applications (Liu et al. 2022; Li et al. 2023; Qi et al. 2023; Huang and Harrie 2019). For instance, KnowWhereGraph system was developed to build a highly interconnected, cross-disciplinary knowledge graph with geographical information enriched to support environmental intelligence (Janowicz et al. 2022). GeoGraphVis platform, empowered by a knowledge graph (Li et al. 2023; Wang, Li, and Gu 2023), integrates cross-domain data, such as disaster impact, medical resources, and regional population, facilitating disaster relief and humanitarian aid. Qi et al. (2023) constructed an EV KG encapsulating various EV-relevant knowledge, such as supply equipment and electricity transmission network, to support decision-making and potential interoperability with additional domain knowledge. However, the application of knowledge graphs in GIS education promotion has not been explored yet. As GIS programs are dynamic and rapidly evolving, knowledge graphs can offer the flexibility and scalability to incorporate and summarize new and emerging geographical knowledge.

Unlike tabular data, where each individual data point is independent, a knowledge graph links heterogeneous data through explicit semantic relationships, allowing sophisticated querying and efficient knowledge discovery. Liu et al. (2022) developed a faceted search platform to support information retrieval from 12 billion statements across more than 30 domains. Li et al.'s (2023) GeoGraphVis interface, supported by a knowledge graph, enables disaster-related queries across various domains, such as medical resources and disaster impact. These efforts facilitate the retrieval of heterogeneous knowledge already present in the database.

Applying knowledge graph to GIS education can enhance the efficiency of searching for the desired information under complicated requirements. An educational knowledge graph, constructed by Troussas et al. (2023), incorporates a cosine similarity method to measure preferences for different educational resources, delivering personalized education material recommendations. Motivated by this, we hope to integrate KG to provide valuable insights for GIS education. In addition to querying existing information, students are often interested in analyzing which programs have similar research fields or which professors work on topics they are passionate about. Consequently, knowledge graphs can be a promising solution for information discovery and program recommendation in GIS education.

2.3 | GIS Promotion Platform Empowered by LLMs

The rapid progress of LLMs can be characterized by a series of GPT models (Floridi and Chiriatti 2020; Brown et al. 2020), which form a solid foundation for the LLMs approaching human-level comprehension. These various LLMs achieve outstanding performance on many NLP tasks, including information retrieval (Qiu et al. 2023), QA (Feng, Ding, and Xiao 2023), and text generation (Wang, Thompson, and Iyyer 2021; Tang, Chuang, and Xia 2023). OpenAI (2022, 2023) has developed cutting-edge LLMs, enabling various NLP tasks through a question-answering process on the ChatGPT platform (OpenAI 2022). This innovative integration allows users to enter their questions on the ChatGPT platform and receive immediate answers, tailored to their specific inquiries and intentions, without requiring prior knowledge of the underlying technologies.

Through the application programming interfaces (APIs) offered by OpenAI, researchers can leverage its capabilities to enhance their domain-specific applications. This access allows for the integration of advanced language processing features into a wide range of specialized fields, including our domain of geography. In domain-specific applications, LLMs can reinforce processes through prompt engineering, which involves carefully designing the inputs that align with the system's capabilities to elicit the optimal responses (Dai et al. 2022; Hu et al. 2023). These advanced tools enable spatial–temporal data collection and analysis, geodatabase queries, and map generations (Yin et al. 2019; Zhang et al. 2023; Feng, Ding, and Xiao 2023).

As the capability of representing space and time began to be recognized (Gurnee and Tegmark 2023), scholars have attempted to unleash the potential of state-of-the-art ChatGPT in fundamental geographical education and research, investigating its effectiveness and limitations. For example, ChatGPT has been used to generate teaching materials (Nguyen et al. 2023) and evaluate exams in GIS education (Mooney et al. 2023). These studies have demonstrated the potential of LLMs to enrich the education field.

3 | Methodology

The GISphere-KG platform is designed to enhance applicants' experience in seeking Cartography/GIS programs. Building upon a worldwide dataset of graduate programs, GISphere-KG enables flexible searches to help applicants find programs that best fit their needs. A key technique of the platform refers to a semantic similarity measure, which can retrieve multiple professors working in similar research fields based on their research interests. Additionally, the platform utilizes semantic similarity to compare applicants' research interests with those in the database, facilitating the recommendation of appropriate programs.

Figure 1 illustrates the workflow of our platform, including: (1) data collection and preprocessing, (2) knowledge graph construction, (3) semantic similarity calculation, and (4) graph search with LLMs. The data collection and preprocessing step details dataset preparation. We then designed an ontology and built the knowledge graph to represent the relationships and entities within the data. In the phase of semantic similarity calculation, we used LLMs to embed research interests into numerical vectors and measured semantic similarities between them. This supports the discovery of professors with similar research interests and program recommendations. The final step streamlines the question-answering process for information retrieval with both explicit and implicit graph searches, discovering professors with similar research interests, and providing personalized program recommendations.

3.1 | Data Collection and Preprocessing

From 2019 to 2023, the GISphere project has collected information from over 600 GIS-related programs worldwide in 97 countries or regions as of December 2023. This extensive dataset covers faculty information from about 2000 professors across more than 400 academic institutions, specializing in fields such as Geographic Information Science (GIS), Global Navigation Satellite System (GNSS), Remote Sensing (RS), physical geography, human geography, and urban planning. To obtain this rich dataset, volunteers utilized web crawling techniques and manual searches. They also regularly updated the database with the latest information on academic-oriented MS/PhD opportunities. The raw datasets were preprocessed and cleaned by extracting specific details and removing emojis, punctuation, and repetitive texts. An annual validation and revision process is in place to ensure the data's accuracy and relevance. Consequently, this comprehensive dataset serves as the data foundation for our KG.

Figure 2 displays the format of the collected data, categorized into four entities: *Country*, *City*, *University*, and *Professor*. Each entity contains multiple attributes. The *Country* entity includes the attribute "continent". The *City* entity includes the attributes "country", "latitude", and "longitude". The *University* entity includes the attribute "city" where the university is located, a "description" of the university, the "department" associated

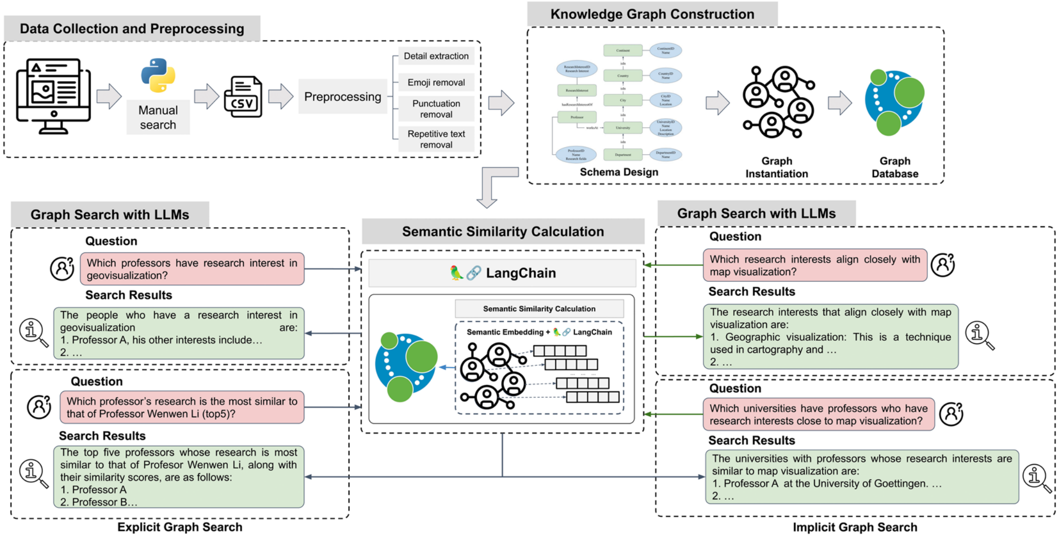


FIGURE 1 | The workflow of the GISphere-KG: (1) data collection and preprocessing, (2) knowledge graph construction, (3) semantic similarity calculation, and (4) graph search with LLMs with both explicit and implicit graph search.

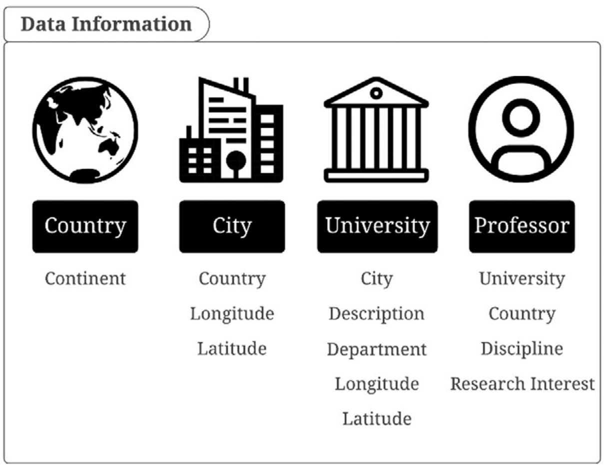


FIGURE 2 | Data structures of the collected data including the four entities country, city, university, professor, and their attributes.

with GIS-related programs, and the university's "latitude" and "longitude". The *Professor* entity includes the "university" and "country" where the professor works, the attribute "discipline", indicating the professor's broad field, such as "physical geography", "human geography", "urban planning", "GIS", "RS", and "GNSS", and the attribute "research interest", which refers to specific areas of focus. Research interests of professors are retrieved from professors' personal/department websites or Google Scholar. This attribute is essential as it supports semantic similarity measures, enabling the identification of professors with similar research interests and could support recommending appropriate programs for applicants.

3.2 | Knowledge Graph Construction

Based on the collected and cleaned dataset, we designed an ontology for GISphere-KG to define the semantic structure among concepts, properties, and their relationships. Figure 3 shows this schema, displaying entities and their corresponding relationships. Compared with LLMs that may generate inaccurate information due to the "hallucination" phenomena.

The GISphere-KG ontology contains seven classes: *Continent*, *Country*, *City*, *University*, *Department*, *Professor*, and *Research Interest*. Each class has multiple attributes, as shown in Figure 3. *Continent*, *Country*, and *City* represent the administrative regions of each program and include the "name" attribute. The *University* includes "name", "description", "latitude", and "longitude". The geolocation of the *University* entity is stored in WKT format to support spatial inquiries. The *Department* class has a "name" attribute related to the GIS programs. The *Professor* class includes "name" and "expertise" attributes. The "expertise" attribute contains six broad fields, namely, physical geography, human geography, urban planning, GIS, RS, and GNSS. The *Research Interest* class is the most important entity in our KG. It defines various research interests of professors, which can be used to analyze similar expertise among professors and provide appropriate program recommendations for applicants.

According to the constructed seven classes, four types of relationships are defined: *isIn*, *worksAt*, *hasResearchInterestOf*, and *isSimilarTo*. The "*isIn*" relationship connects administrative regions or places, such as <*Country*, *isIn*, *Continent*> to denote whether a *Country* is part of a *Continent*. The "*worksAt*" relationship specifies the *University* where a *Professor* is employed, represented by <*Professor*, *worksAt*, *University*>. The "*hasResearchInterestOf*" relationship indicates the research interests of a *Professor*, represented as <*Professor*, *hasResearchInterestOf*, *Research Interest*>. Specifically, the relationship between *Professor* and *Research Interest* is one-to-many, meaning that a *Professor* can have multiple research interests. "*isSimilarTo*" represents the similarity between professors based on their research interests. More details will be introduced in Section 3.3.

Based on the designed ontology, we converted the dataset into a graph format compatible with Neo4j's data loader requirements. After compiling all instances for nodes and relationships, we

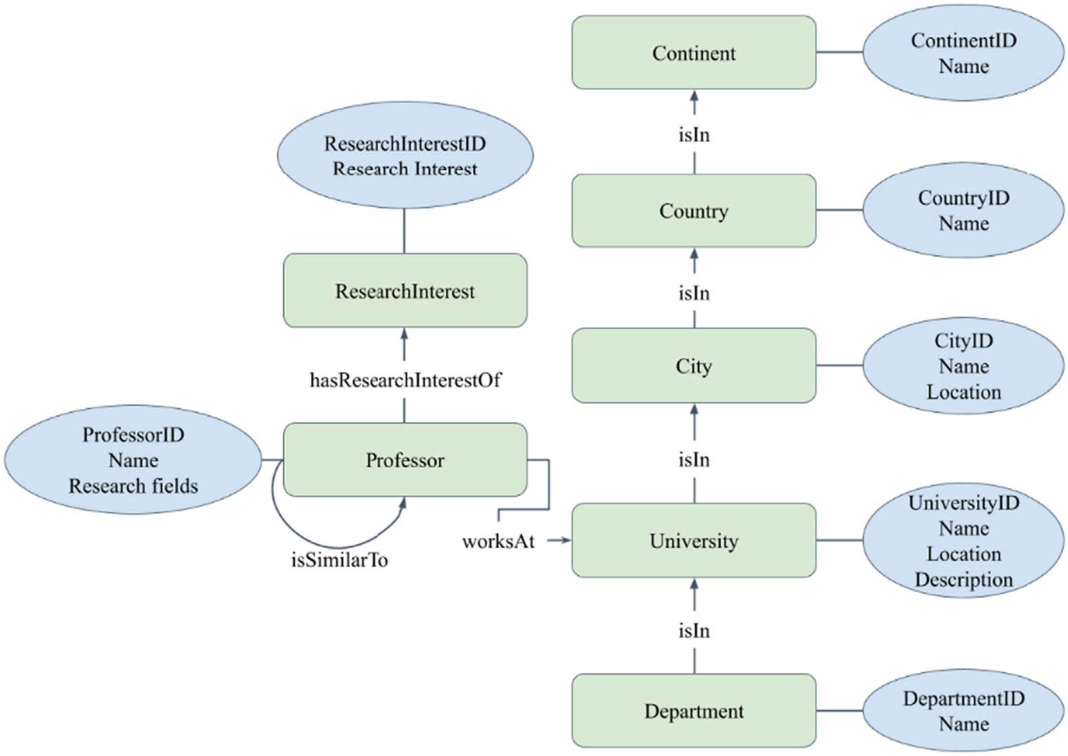


FIGURE 3 | Designed ontology for GISphere-KG.

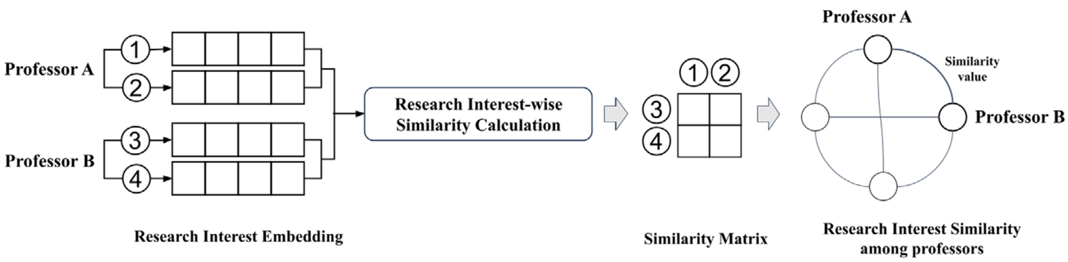


FIGURE 4 | Research similarity calculation between two professors.

used Neo4j's built-in bulk import tool to load them for graph instantiation.

3.3 | Semantic Similarity Calculation

When applicants search for professors with certain research interests, they often want to explore other professors with similar interests or receive relevant program recommendations based on their desired research areas. Measuring the semantic similarity of research interests is crucial for retrieving professors with similar research interests and facilitating program recommendations.

To accomplish this, we leveraged a state-of-the-art embedding model, the *text-embedding-ada-002* (Gao 2023), to convert each research interest into a high-dimensional numerical vector that captures its hidden semantics. This model[3], accessible from OpenAI documentation guide,[3] efficiently handles long contexts (Gao 2023). After obtaining the semantic vectors, we applied cosine similarity to calculate the semantic similarity between each pair of research interests.

To represent professors' research similarities, we built a new semantic relationship, *isSimilarTo*, assigning similarity values to it. These relationships were stored in a graph database, allowing the system to directly identify professors with similar research interests. Figure 4 illustrates the process of calculating the semantic similarity between professors.

Using the *text-embedding-ada-002* model, each research interest is embedded into a vector representing its semantic features. The semantic similarity between each pair of research interests is calculated using cosine similarity, as represented in Equation (1). The value of semantic similarity ranges from 0 to 1, with higher values indicating higher relevance between the two research interests.

$$s(\boldsymbol{a}, \boldsymbol{b}) = \cos(\theta) = \frac{\boldsymbol{a} \bullet \boldsymbol{b}}{\|\boldsymbol{a}\|\|\boldsymbol{b}\|} \tag{1}$$

The semantic similarity of each pair of research interests was calculated, building a similarity matrix, as shown in Figure 4. The similarity matrix of the two professors reflects their research-interest-wise semantic similarity, as shown in Equation (2).

$$
\boldsymbol{s}^{\boldsymbol{AB}} = \begin{bmatrix} s_{11} & s_{12} & \dots & s_{1n} \\ s_{21} & s_{22} & \dots & s_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ s_{m1} & s_{m1} & \dots & s_{mn} \end{bmatrix} \tag{2}
$$

where $\boldsymbol{s}^{\boldsymbol{AB}} \in R^{m \times n}$ is a similarity matrix of research interests for professors $\boldsymbol{A}$ and $\boldsymbol{B}$. It has the size of $m \times n$, corresponding to the number of research interests of two professors. Specifically, m is the number of research interests for professor $\boldsymbol{A}$, while n is the number of research interests that another one has. The element s_{ij} in the matrix indicates the similarity value between the ith research interest from one professor and jth one from another. As a result, for the research similarity between two professors, we aggregated the pairwise scores from the above similarity matrix into a single representative score by calculating the average value of the matrix, as demonstrated in Equation (3).

$$
similarity(professor\ A, professor\ B) = mean\left(\boldsymbol{s}^{\boldsymbol{AB}}\right) \tag{3}
$$

After calculating the semantic similarity values between professors, we defined a new relationship called "*isSimilarTo*" within our constructed KG. This relationship includes an attribute "similarity value" that saves the similarity values between professors. It allows applicants to easily perform natural language queries to discover professors whose research interests closely align with their own. Consequently, applicants can retrieve multiple professors with similar interests and identify suitable programs efficiently.

3.4 | Graph Search With LLMs

After constructing the KG, we developed two types of graph searches so that users could retrieves information of professors from the graph database and receive recommendations on programs. In particular, two search types are performed. First, "*Explicit Graph Search*" retrieves information directly from the graph database and discovers professors with similar research interests. Second, "*Implicit Graph Search*" recommends programs and retrieves information based on semantic similarity inference. To accomplish this, Neo4j graph database is utilized to store the established knowledge graph and semantic similarity relationships. This setup allows for efficient information retrieval and program recommendation through a question-answering process using Cypher statements.

Cypher, the graph query language for Neo4j, enables querying the database with specific syntax formats. However, predefined Cypher statements can be inflexible, limiting the ability to handle complex queries and failing to accommodate unexpected and diverse questions provided by applicants with varying preferences. To address this limitation and fulfill diverse querying demands, we integrated LLMs through the LangChain framework using OpenAI's APIs. LangChain, as an off-the-shelf open-source framework, provides convenient access to different LLMs, thus facilitating their integration into NLP-supported applications. Notably, by integrating LangChain with the Neo4j database, we enabled the generation of Cypher statements from natural language inputs. This strategy offers a more adaptable and user-friendly solution. As a result, applicants can conveniently communicate with the platform as they would with a human, without encountering technical barriers. As we have pointed out problems associated with LLMs that can lead to inaccurate responses from ChatGPT, our GISphere-KG platform, which is also supported by LLMs, may produce undesirable answers as well. When we provide natural language requests to the platform, the process of Cypher statement generation requires the LLM to be versatile enough to understand the intention and context of the inputs. This limitation necessitates that we provide the model with as much explicit guidance as possible to ensure more accurate outputs for users. Additionally, any updates to the GISphere-KG based on new data sources undergo thorough review and validation to maintain the system's accuracy and reliability.

Given the diverse needs of students during their application process, we developed two information retrieval functions based on the most relevant and frequent questions they may have: "*Explicit Graph Search*" and "*Implicit Graph Search*". The explicit graph search function allows users to search for specific information stored in the KG without inference. Examples include queries such as "Where is a university located?" or "What are a professor's research interests?" Additionally, using the explicit graph search, applicants can discover other professors with similar interests based on the pre-calculated "*isSimilarTo*" relationship. The implicit graph search aims to recommend the most suitable programs by inferring semantic similarity between the applicants' input research interests and those stored in the database. The two functions ensure that applicants can efficiently retrieve information and receive personalized program recommendations, enhancing their overall experience using the platform.

3.4.1 | Explicit Graph Search

This function is designed to retrieve existing information from the graph. Natural language questions provided by applicants are converted into Cypher statements using LLMs, which are further used to query the Neo4j graph database for direct knowledge without inference. Applicants can retrieve information for several predefined entities (*Continent*, *Country*, *City*, *University*, *Department*, *Professor*, and *Research Interest*) along with their corresponding attributes and relationships between these entities (*isIn*, *worksAt*, *hasResearchInterestOf*). For instance, Question 1.1 illustrates how to query the graph for specific knowledge retrieval, as shown in Figure 1. In addition to fetching direct information from the KG, the platform utilizes the "*isSimilarTo*" relationship to connect pairs of professors, incorporating a "similarity value" attribute that measures the similarity of their research interests. This function enables users to discover other professors whose expertise closely aligns with their interests, thereby facilitating potential academic collaborations and mentorships. It should be noted the similarity of professors' research interests has been precalculated and stored in the KG to save computational resources, thus this function belongs to the explicit graph search. This exploration also serves as a valuable tool for applicants to understand the broader trends and emerging research areas within their field of interest. Question 1.2 is an example querying question for exploring professors' research similarities.

Question 1.1. Which professors have research interests in geovisualization?

Question 1.2. Which professor's research is most similar to that of Professor Wenwen Li (top5)?

3.4.2 | Implicit Graph Search

The implicit graph search complements the explicit graph search by incorporating research-interest-based recommendations and graph retrieval based on semantic similarity with inference. It not only recommends programs but also provides other relevant information associated with research interests in the database, inferred through semantic similarity. When an applicant submits a query in natural language, it includes their preferred research interest, which is converted into a vector. The system employs LLMs to perform semantic similarity calculations between the applicant's input and the embeddings in the database, identifying the most similar research interests and additional entities attached to those research interests.

By integrating advanced natural language understanding, the GISphere-KG system can respond accurately to direct inquiries and uncover implicit insights, thereby enhancing the overall utility and effectiveness of the program recommendation process. Below are examples of implicit graph searches. Question 2.1 focuses on identifying research interests based on semantic similarity calculation, while Question 2.2 takes this a step further by retrieving specific information (e.g., university and professors) related to such semantic similarities to enhance program recommendations.

Question 2.1. Which research interests align closely with map visualization?

Question 2.2. Which universities have professors who have research interests close to map visualization?

4 | Experiments

In this section, we showcase several outputs of the GISphere-KG platform. Our GISphere-KG operates as a text-based interface, and we display the ontology structure of the constructed knowledge graph using Neo4j's built-in visualization function. We present several examples to highlight the ontology visualization of the KG and its responses to a variety of queries. Our code of the system has been openly available on GitHub: https://github.com/GIS-Info/GISphereKG-ChatBot.git.

4.1 | Knowledge Graph Ontology Visualization

Figure 5 shows examples of the ontology structure of the constructed knowledge graph using the Neo4j visualization platform. It shows ontology structures among instances from various entities and relationships. Circles with different colors represent different entities, such as *City*, *Country*, *Continent, Department, Professor, Research Interest*, and *University*. Arrows indicate the semantic relationships between pairs of entities, including *hasResearchInterestOf*, *isIn*, *isSimilarTo*, and *WorksAt*. Figure 5a shows a subgraph that contains <*City*, *isIn*, *Country*> and <*Country*, *isIn*, *Continent*>. In this subgraph, orange circles represent *City* entities, light green circles represent *City*, and light-yellow circles are *Continent* nodes. In this figure, different types of nodes are connected by an edge of "*isIn*". Figure 5b shows a few examples of <*Professor*, *hasResearchInterestOf*, *Research Interest*> and <*Professor*, *WorksAt*, *University*>. Specifically, the *Professor* node is in blue, the *Research Interest* nodes are in orange, and the node with pink color is the *University* node. These visualizations clearly illustrate the structure of the knowledge graph, demonstrating how various entities are interconnected within the GISphere-KG platform.

4.2 | Question-Answering Process

Figures 6 and 7 illustrate the interface of the GISphere-KG, which allows applicants to inquire about GIS-related programs by simply typing their questions. This user-friendly, intuitive question-answering process requires no prior knowledge of the platform. Building upon the GISphere dataset, the GISphere-KG offers a comprehensive and accessible platform for exploring GIS programs, significantly enhancing the user experience for prospective students and researchers. To demonstrate its querying capabilities, we use four example questions mentioned in the previous section. Both explicit graph search and implicit graph search functions are performed. In addition, we provide a scenario to illustrate how the GISphere-KG platform supports personalized user experiences.

4.2.1 | Explicit Graph Search

Figure 6 shows the results of the explicit graph search of GISphere-KG. Figure 6a displays the search results for the question "Which professors have research interests in geovisualization?". The system first summarizes the query, searches for the keyword in the database, and lists ten different professors who match the criteria. Moreover, each professor's research interests are briefly introduced, and external links to their personal/institutional websites are provided with further details. Figure 6b shows the results for the query, "*Which professor's research is most similar to that of Professor Wenwen Li (top5)*?". Since research similarities among professors have been pre-calculated and stored in the graph database, LLMs can directly interpret the question, generate the appropriate Cypher statement, and retrieve the top five professors who have the most similar research interests. The results are listed in descending order of similarity, with each professor's name and corresponding similarity value provided.

4.2.2 | Implicit Graph Search

Figure 7 displays the search results of finding research interests similar to those input by applicants. Recognizing that research interests can be expressed in various ways, the system identifies similar interests within the graph database. This helps applicants to discover research fields that match their interests. As shown in Figure 7a, the platform responds to the question "What research interests align closely with map visualization?". The

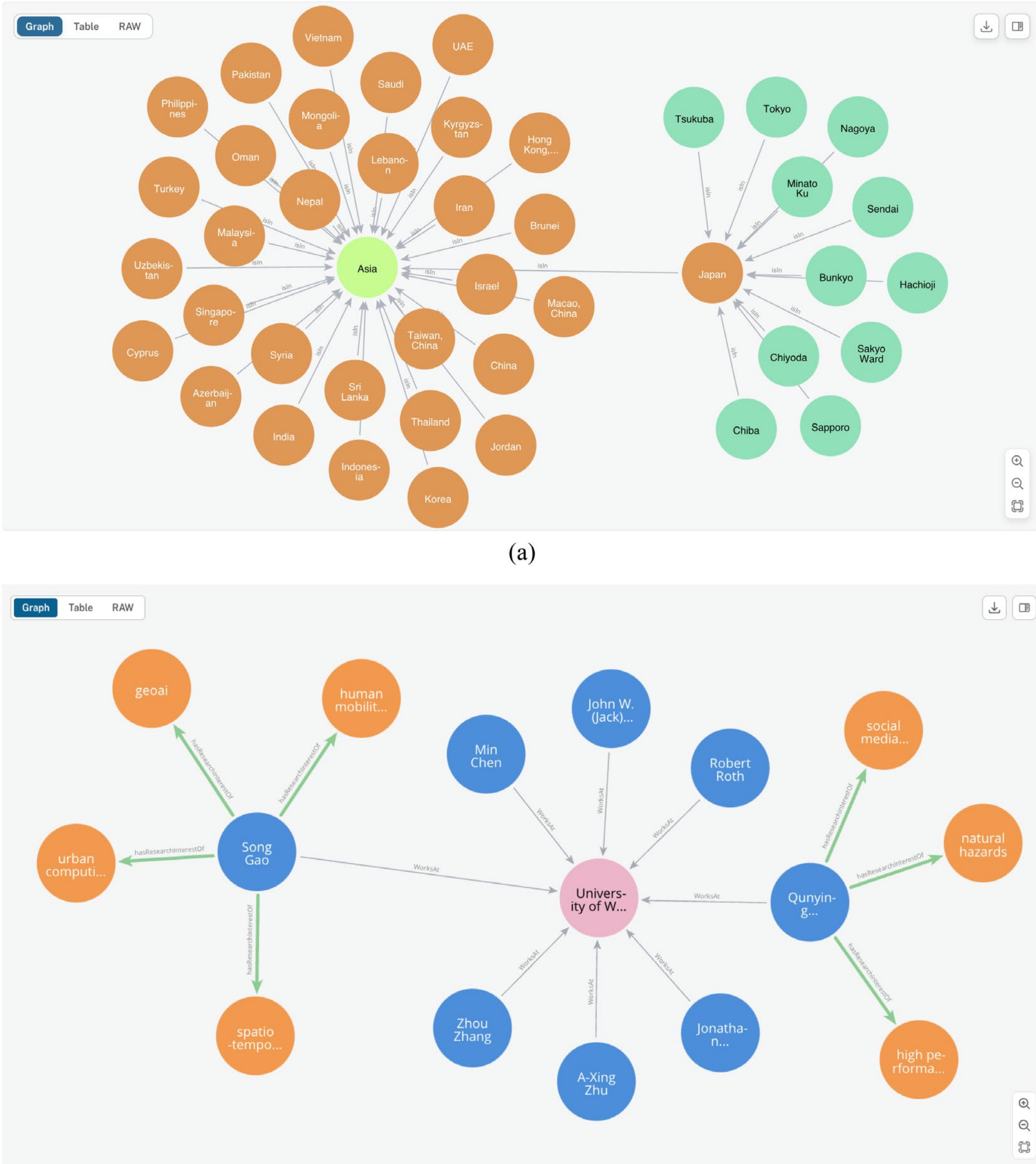


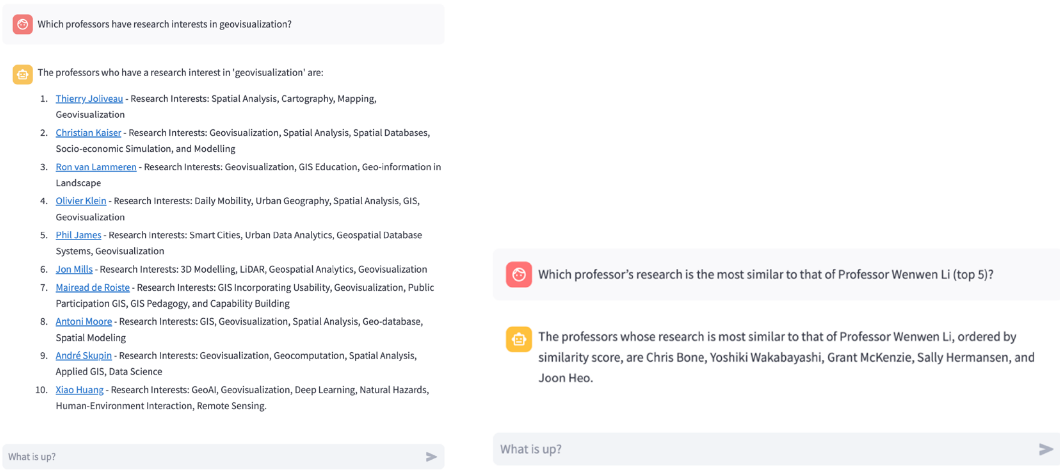
FIGURE 5 | The ontology of the GISphere-KG using the Neo4j built-in visualization function. (a) shows the subgraph for <*City*, *isIn*, *Country*> and <*Country*, *isIn*, *Continent*>; (b) shows the subgraph for <*Professor*, *hasResearchInterestOf*, *Research Interest*> and <*Professor*, *WorksAt*, *University*>.

GISphere-KG converts the question into a vector, retrieves relevant embeddings from the database, and calculates the semantic similarity between the input and research interests stored in the KG. It then identifies and presents the most similar research interests in the KG in a human-readable format. Research interests similar to "map visualization" include "geographic visualization", "geographic visualization/visual analytics", "cartographic visualization", and "cartography and geovisualization".

This platform also retrieves information related to research interests inferred by semantic similarity, as shown in Figure 7b. This process involves several backend processing steps: (1) analyzing the query with NLP models to identify relevant research interests. (2) searching for "Research Interest" nodes in the graph database based on semantic similarity, and (3) generating Cypher statements to query information associated with these research interests. As a result, the system retrieves information about professors, their research interests, and universities. This information is then retrieved and presented in natural language on the frontend web page, making it convenient for applicants to understand and explore.

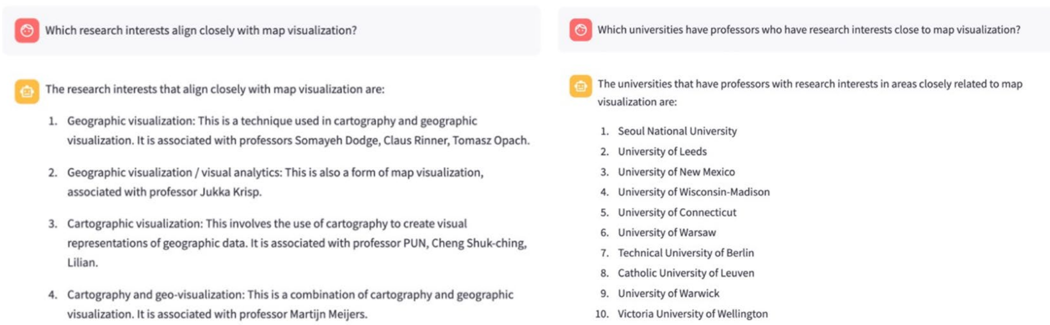
4.2.3 | User Retrieval Scenario

In addition, we have provided another one uses cases to further display that our platform can understand complicated questions to generate desirable answers for applicants. Figure 8 shows the response for the question "*John Doe is from the United States and is interested in hazards research. He is looking for a university to pursue his graduate studies and is looking to develop skills in remote sensing and spatial analysis.*" The platform understands the intention of the applicant, which is to find a university, and captures the research interests from it,



(a) (b)

FIGURE 6 | Examples of the explicit graph search results. (a) "Which professors have research interests in geovisualization?"; (b) "Which professor's research is most similar to that of Professor Wenwen Li (top5)?".



(a) **(b)**

FIGURE 7 | Implicit graph search for program recommendation. (a) Search for similar research interests; (b) Search for information associated with research interests.

including hazards and remote sensing. Except for those explicitly mentioned in questions, the platform also finds out similar ones related to applicant's interest, including "geohazards" and "hazard". As a result, it intelligently integrates key information as a clear request, which is to find a university where he can work on hazards, remote sensing, and spatial analysis. The response lists professors linked with websites affiliated with their universities as well as research interests. In this way, the applicant can not only know the university to apply for but also corresponding professors the applicant might want to work with. Despite that our system effectively retrieves relevant information, some limitations should be acknowledged in understanding complicated questions. We observed that the system did not consistently interpret the input questions successfully. During testing, there were instances where the system accurately captured all relevant research interests, while at other times, only a few were identified. This inconsistency may be due to the fact that current LLMs are primarily trained on general contexts and lack sufficient domain-specific knowledge. To address this, future work could focus on fine-tuning the LLMs with more geographic terminologies and enhancing the understanding of the system by incorporating prompts with more geographic knowledge. These improvements would enable the platform to better understand complex user queries, ultimately enhancing the overall user experience.

5 | Discussion and Future Work

This research demonstrates a valuable example of how cutting-edge technologies such as Knowledge Graphs and Large Language Models can be combined to enrich GIS and geography education. By leveraging these emerging technologies, GISphere-KG enables students to efficiently explore and discover relevant educational programs, delve into GIS research topics, gain insights into research trends, and discover professors who match their research interests. These capabilities, previously unattainable through conventional educational methods, represent a significant advancement. GISphere-KG offers a personalized recommendation system that customizes program suggestions based on user interests and academic backgrounds, thereby fostering a more directed and meaningful educational journey. Additionally, the integration of LLMs with knowledge graphs allows users to query complex research questions and access insights that were previously achieved with extensive manual coding. We advocate for the further integration of LLMs and other advanced AI tools in educational frameworks to equip future geographers with a robust, intuitive, and comprehensive learning environment.

Notably, the GISphere-KG only focuses on facilitating the search for GIS-related programs and professors in this paper, while its

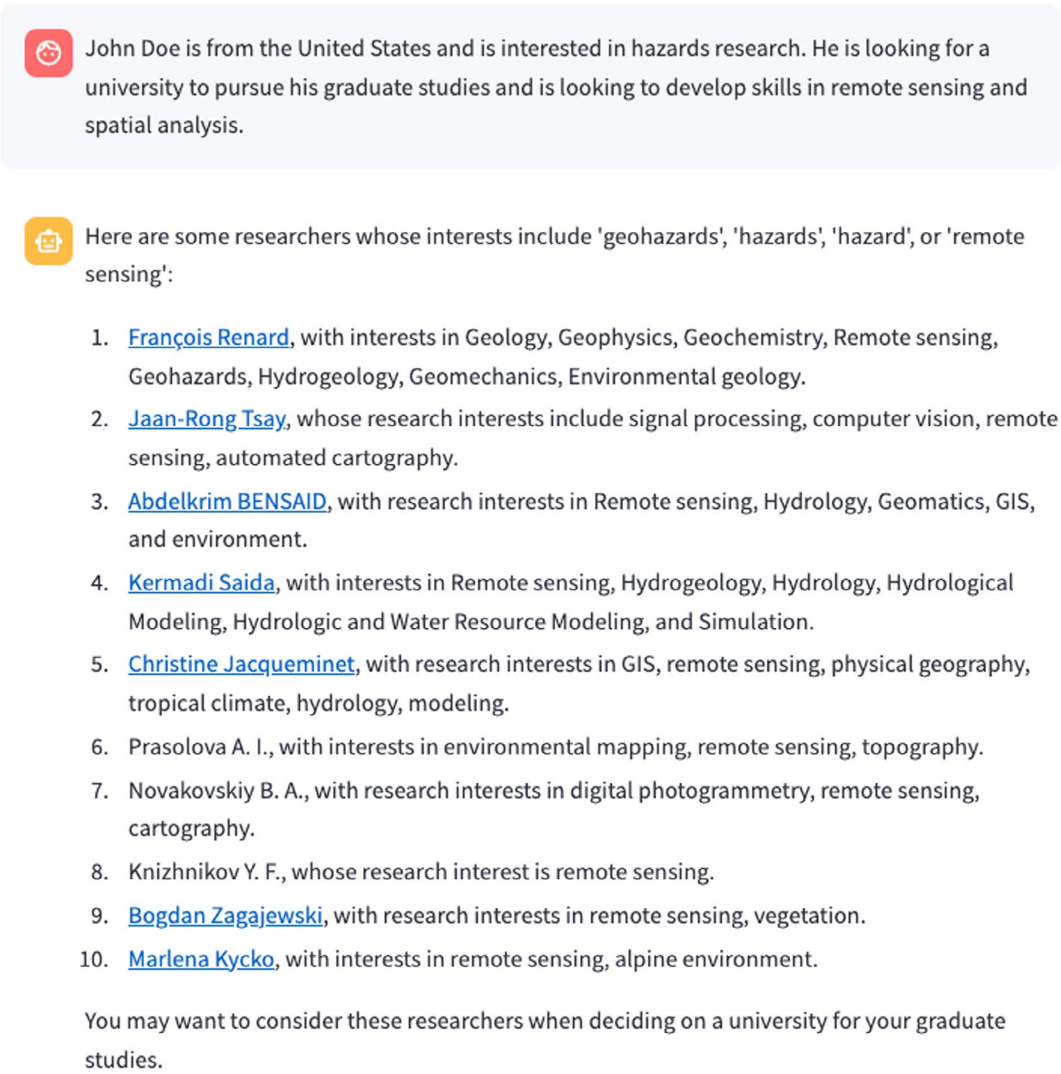


FIGURE 8 | Implicit graph search for university recommendation.

underlying technology holds the potential to support a wide range of applications, significantly broadening the platform's utility and user base. Beyond program search, GISphere-KG can assist students in understanding GIS structures and research dynamics, learning about specific sub-research areas, exploring their research interests, and preparing for career developments. The platform allows its users to explore research trends and researchers' interests, making it a valuable resource for researchers seeking collaboration, peer review, or panel participation opportunities. Additionally, combined with up-to-date recruitment information from the GISphere database, the platform could offer insights for early career researchers in job hunting and professionals contemplating career shifts. Its knowledge in GIS provides the opportunity for individuals pursuing a general understanding of GIS. Future enhancements could include the development of tools for visualizing GIS research trends, creating interactive forums for academic discussion, and implementing AI-driven advising systems for personalized guidance in education and career planning. These initiatives would not only extend the platform's capabilities but also foster a more engaged and informed GIS community.

While GISphere-KG offers a valuable tool for users, we acknowledge several limitations that can be addressed in future work. First, we have noticed that the data quality has a significant impact on the platform's responses. For instance, professors' research interests can evolve with the emergence of new technologies and research topics, making their initial research interests potentially outdated. Additionally, research interests claimed by researchers may not accurately reflect their work. Publications, on the other hand, may provide a more accurate reflection of their current research work at the time. To address this, we plan to expand our data sources to incorporate a broader range of academic information, such as publications, citations, teaching records, and project experiences. This expansion will enable a more comprehensive understanding of professors' research interests, refining the similarity measure and providing more personalized recommendations and discovery of similar research interests.

Second, despite the extensive coverage of universities and professors in the GISphere database, some data and information may still be missing, especially in several Global South countries. Also, our datasets primarily focus on academic programs, such as academic master's and doctoral programs. As a result, regular updates of the database are crucial as research and educational programs continuously evolve, and we plan to expand our efforts to include non-academic programs and collect data from additional sources. For example, online programs are becoming increasingly important and popular, and are playing a crucial role in GIS education promotion. To accomplish this, we plan to advocate public participatory GIS (PPGIS), encouraging users to actively engage with the platform by providing feedback, contributing new programs, updating information, and sharing reviews. This will expand and update the platform's knowledge base and foster a collaborative learning environment.

Finally, although the GISphere-KG platform is generally effective at retrieving faculty and program information based on user queries, it still encounters challenges when processing and interpreting complex input questions. This primarily arises from the "hallucination" issue of current LLMs that generate inaccurate or misleading information. Additionally, the existing LLMs are trained on general knowledge, and specific domain knowledge might be integrated in future systems. In the future, with the advancements in LLMs, we will also aim to refine the model's capability to deal with more complicated questions by integrating latest data sources and time information of program and expanding the range of academic and non-academic programs.

6 | Conclusions

In this paper, we introduce GISphere-KG, a knowledge graph and LLM-empowered platform for information retrieval and recommendations of GIS-related programs. It provides information about the continent, country, city, university, professors, and research interests. By integrating LLMs through the LangChain framework, applicants can enter their questions and get accurate and intelligent responses in natural language without prior knowledge about querying statements or other technical skills. Based on the GISphere-KG, applicants can search for a wide range of information stored in the graph database and explore research similarities among professors. They can also get recommendations on programs and related information based on their specified research interests, leveraging semantic similarity calculations.

In the future, we will keep updating program resources worldwide so that applicants can get to know the latest information for their admission. We also plan to enrich the GISphere-KG with more functions, such as KG visualization with the input of textual questions, a multi-hop question-answering function that enables more complicated information search, and temporal information integration which allows applicants to query programs and get recommendations within specific periods. In this way, our platform will understand more complicated questions intelligently.

We believe that our platform will significantly advance Cartography/GIS education, and more broadly, geography education. The comprehensive search functions enable applicants to dive deep into a rich repository of academic and professional information, fostering a deep understanding in making informed decisions for their educational journey. Furthermore, it facilitates research and collaboration, and could help the recruitment of the next generation of geographers.

Acknowledgments

The authors would like to thank all volunteers and contributors of the GISphere project and the community. In particular, Jingyuan Zhang and Pengyu Chen. This work is supported by the Cartography and Geographic Information Society (CaGIS) Rising Award, and the non-profit organization GISphere Corporation. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the funders.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data and code that support the findings of this study are openly available in GitHub at https://github.com/GIS-Info/GISphereKG-ChatBot.

Endnotes

[1] https://www.aag.org/guide-to-geography-programs-opportunities/.

[2] https://gisphere.info/.

[3] https://platform.openai.com/docs/guides/embeddings/use-cases.

References

A Tactical Guide to Graduate Student Recruitment. 2024. https://www.hanoverresearch.com/insights-blog/higher-education/a-tactical-guide-to-graduate-student-recruitment/#:~:text=Stay%20up%20to%20date%20with,likely%20to%20impact%20our%20institution?

Baker, T. R., and S. Witham Bednarz. 2003. "Lessons Learned From Reviewing Research in GIS Education." *Journal of Geography* 102, no. 6: 231–233.

Batty, M. 2018. *Inventing Future Cities*. Cambridge, MA: MIT Press.

Bednarz, S., S. Heffron, and N. Huynh. 2013. *A Road Map for 21st Century Geography Education*. Washington, DC: Association of American Geographers.

Bodenhamer, D. J., J. Corrigan, and T. M. Harris. 2010. *The Spatial Humanities: GIS and the Future of Humanities Scholarship*. Bloomington: Indiana University Press.

Bodycott, P. 2009. "Choosing a Higher Education Study Abroad Destination: What Mainland Chinese Parents and Students Rate as Important." *Journal of Research in International Education* 8, no. 3: 349–373.

Bodycott, P., and A. Lai. 2012. "The Influence and Implications of Chinese Culture in the Decision to Undertake Cross-Border Higher Education." *Journal of Studies in International Education* 16, no. 3: 252–270.

Bohman, E. 2014. "Attracting the World: Institutional Initiatives' Effects on International Students' Decision to Enroll." *Community College Journal of Research & Practice* 38, no. 8: 710–720.

Boyd, D. S., and G. M. Foody. 2011. "An Overview of Recent Remote Sensing and GIS Based Research in Ecological Informatics." *Ecological Informatics* 6, no. 1: 25–36.

Brown, T., B. Mann, N. Ryder, et al. 2020. "Language Models Are Few-Shot Learners." *Advances in Neural Information Processing Systems* 33: 1877–1901.

Cebolla-Boado, H., Y. Hu, and Y. N. Soysal. 2018. "Why Study Abroad? Sorting of Chinese Students Across British Universities." *British Journal of Sociology of Education* 39, no. 3: 365–380.

Dai, Z., V. Y. Zhao, J. Ma, et al. 2022. "Promptagator: Few-Shot Dense Retrieval from 8 Examples." *arXiv Preprint arXiv:2209.11755*.

de Wit, H., E. Minaeva, and L. Wang, eds. 2022. *International Student Recruitment and Mobility in Non-Anglophone Countries: Theories, Themes, and Patterns*. London: Routledge.

DeMers, M. N., J. J. Kerski, and C. J. Sroka. 2021. "The Teachers Teaching Teachers GIS Institute: Assessing the Effectiveness of a GIS Professional Development Institute." *Annals of the American Association of Geographers* 111, no. 4: 1160–1182. https://doi.org/10.1080/24694452.2020.1799745.

Devlin, J., M. W. Chang, K. Lee, and K. Toutanova. 2018. "Bert: Pre-Training of Deep Bidirectional Transformers for Language Understanding." *arXiv Preprint arXiv:1810.04805*.

Dsouza, A., N. Tempelmeier, R. Yu, S. Gottschalk, and E. Demidova. 2021. "Worldkg: A World-Scale Geographic Knowledge Graph." In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*, 4475–4484.

Feng, Y., L. Ding, and G. Xiao. 2023. "GeoQAMap—Geographic Question Answering With Maps Leveraging LLM and Open Knowledge Base (Short Paper)." In *DROPS-IDN/v2/Document*. Schloss-Dagstuhl–Leibniz Zentrum für Informatik. https://doi.org/10.4230/LIPIcs.GIScience.2023.28.

Floridi, L., and M. Chiriatti. 2020. "GPT-3: Its Nature, Scope, Limits, and Consequences." *Minds and Machines* 30: 681–694.

Gao, A. K. 2023. "Vec2Vec: A Compact Neural Network Approach for Transforming Text Embeddings With High Fidelity." *arXiv Preprint arXiv:2306.12689*.

GISphere. 2023. *GISphere Guide*. https://gisphere.github.io/.

Goodchild, M. F., and P. A. Longley. 2021. "Geographic Information Science." In *Handbook of Regional Science*, 1597–1614. Berlin: Springer.

Gurnee, W., and M. Tegmark. 2023. "Language Models Represent Space and Time." *arXiv*. https://doi.org/10.48550/arXiv.2310.02207.

Hu, Y., G. Mai, C. Cundy, et al. 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, no. 11: 2289–2318. https://doi.org/10.1080/13658816.2023.2266495.

Huang, W., and L. Harrie. 2019. "Towards Knowledge-Based Geovisualisation Using Semantic Web Technologies: A Knowledge Representation Approach Coupling Ontologies and Rules." *International Journal of Digital Earth* 13, no. 9: 976–997.

Ishwari, K. S. D., A. K. R. R. Aneeze, S. Sudheesan, H. J. D. A. Karunaratne, A. Nugaliyadde, and Y. Mallawarrachchi. 2019. "Advances in Natural Language Question Answering: A Review." *arXiv Preprint arXiv:1904.05276*.

Jakab, I., M. Ševcík, and H. Grežo. 2017. "Model of Higher GIS Education." *Electronic Journal of e-Learning* 15, no. 3: 220–234.

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, no. 4: 625–636. https://doi.org/10.1080/13658816.2019.1684500.

Janowicz, K., P. Hitzler, W. Li, et al. 2022. "Know, Know Where, KnowWhereGraph: A Densely Connected, Cross-Domain Knowledge Graph and Geo-Enrichment Service Stack for Applications in Environmental Intelligence." *AI Magazine* 43, no. 1: 30–39.

Knowledge Graph. 2024. https://en.wikipedia.org/wiki/Knowledge_graph.

Lei, S. A., and N. K. Chuang. 2010. "Demographic Factors Influencing Selection of an Ideal Graduate Institution: A Literature Review With Recommendations for Implementation." *College Student Journal* 44, no. 1: 84.

Li, J., T. Tang, W. X. Zhao, J. Y. Nie, and J. R. Wen. 2022. "Pretrained Language Models for Text Generation: A Survey." *arXiv Preprint arXiv:2201.05273*.

Li, W., S. Wang, X. Chen, et al. 2023. "Geographvis: A Knowledge Graph and Geovisualization Empowered Cyberinfrastructure to Support Disaster Response and Humanitarian Aid." *ISPRS International Journal of Geo-Information* 12, no. 3: 112.

Li, W., S. Wang, S. Wu, Z. Gu, and Y. Tian. 2022. "Performance Benchmark on Semantic Web Repositories for Spatially Explicit Knowledge Graph Applications." *Computers, Environment and Urban Systems* 98: 101884.

Liu, Z., Z. Gu, T. Thelen, et al. 2022. "Knowledge Explorer: Exploring the 12-Billion-Statement KnowWhereGraph Using Faceted Search (Demo Paper)." In *Proceedings of the 30th International Conference on Advances in Geographic Information Systems*, 1–4.

Longley, P. A. 2000. "The Academic Success of GIS in Geography: Problems and Prospects." *Journal of Geographical Systems* 2, no. 1: 37–42. https://doi.org/10.1007/s101090050027.

Mathews, A. J., and T. A. Wikle. 2019. "GIS&T Pedagogies and Instructional Challenges in Higher Education: A Survey of Educators." *Transactions in GIS* 23, no. 5: 892–907.

Mooney, P., W. Cui, B. Guan, and L. Juhász. 2023. "Towards Understanding the Geospatial Skills of ChatGPT: Taking a Geographic Information Systems (GIS) Exam." In *Proceedings of the 6th ACM SIGSPATIAL International Workshop on AI for Geographic Knowledge Discovery*, 85–94. Hamburg, Germany: ACM. https://doi.org/10.1145/3615886.3627745.

Naz, Z., N. Siddiqua, and M. Asim. 2019. "Problems in Higher Education: Graduate Program, Selection of Program, Predictor of Success." *Pakistan Journal of Educational Research* 2, no. 1: 59.

Nguyen, X. H., H. A. Nguyen, L. Cao, and H. Truong. 2023. *Unleashing the Potential and Recognizing the Limitations of ChatGPT in Vietnamese Geography Education*. https://doi.org/10.35542/osf.io/r4dg6.

OpenAI. 2022. *Introducing ChatGPT*. https://openai.com/blog/chatgpt.

OpenAI. 2023. *GPT-4 Technical Report*. https://arxiv.org/abs/2303.08774.

Paris, J. H., M. Birnbaum, and N. Dix. 2020. "Selecting Students: Graduate Admissions Criteria and Evaluation Methodologies." In *A Comprehensive Guide to Graduate Enrollment Management*, 72–91. London: Routledge.

Qi, Y., G. Mai, R. Zhu, and M. Zhang. 2023. "EVKG: An Interlinked and Interoperable Electric Vehicle Knowledge Graph for Smart Transportation System." *Transactions in GIS* 27: 949–974.

Qiu, Q., Z. Huang, X. Dexin, et al. 2023. "Integrating NLP and Ontology Matching Into a Unified System for Automated Information Extraction From Geological Hazard Reports." *Journal of Earth Science* 34, no. 5: 1433–1446. https://doi.org/10.1007/s12583-022-1716-z.

Reinanda, R., E. Meij, and M. de Rijke. 2020. "Knowledge Graphs: An Information Retrieval Perspective. Foundations and Trends®." *Information Retrieval* 14, no. 4: 289–444.

Revell, M., and M. Benfield. 2022. *The State of Geography: Data and Trends in Higher Education*. https://www.aag.org/the-state-of-geography-data-and-trends-in-higher-education/.

Rickles, P., C. Ellul, and M. Haklay. 2017. "A Suggested Framework and Guidelines for Learning GIS in Interdisciplinary Research." *Geo: Geography and Environment* 4, no. 2: e00046. https://doi.org/10.1002/geo2.46.

Rosenkrantz, L., N. Schuurman, N. Bell, and O. Amram. 2021. "The Need for GIScience in Mapping COVID-19." *Health & Place* 67: 102389.

Tang, R., Y.-N. Chuang, and H. Xia. 2023. "The Science of Detecting LLM-Generated Texts." *arXiv*. https://doi.org/10.48550/arXiv.2303.07205.

Tate, N. J., and C. H. Jarvis. 2017. "Changing the Face of GIS Education With Communities of Practice." *Journal of Geography in Higher Education* 41, no. 3: 327–340. https://doi.org/10.1080/03098265.2017.1315534.

Tian, J. 2017. "Mentoring Undergraduates in Cartography and Geographic Information Science: An Apprenticeship Model." *Transactions in GIS* 21, no. 6: 1148–1164. https://doi.org/10.1111/tgis.12272.

Tonmoy, S. M., S. M. Zaman, V. Jain, et al. 2024. "A Comprehensive Survey of Hallucination Mitigation Techniques in Large Language Models." *arXiv Preprint arXiv:2401.01313*.

Troussas, C., A. Krouska, P. Tselenti, D. K. Kardaras, and S. Barbounaki. 2023. "Enhancing Personalized Educational Content Recommendation Through Cosine Similarity-Based Knowledge Graphs and Contextual Signals." *Information* 14, no. 9: 505.

Wang, M., L. Qiu, and X. Wang. 2021. "A Survey on Knowledge Graph Embeddings for Link Prediction." *Symmetry* 13, no. 3: 485.

Wang, S., W. Li, and Z. Gu. 2023. "GeoGraphViz: Geographically Constrained 3D Force-Directed Graph for Knowledge Graph Visualization." *Transactions in GIS* 27: 931–948.

Wang, S., L. Thompson, and M. Iyyer. 2021. "Phrase-Bert: Improved Phrase Embeddings From Bert With an Application to Corpus Exploration." *arXiv Preprint arXiv:2109.06304*.

Wang, Y., Y. Kang, H. Liu, et al. 2023. "Choosing GIS Graduate Programs From Afar: Chinese Students' Perspectives." *Transactions in GIS* 27, no. 2: 450–475.

Weng, Q. 2010. *Remote Sensing and GIS Integration*. New York, NY: McGraw Hill.

Wu, X., W. Dong, L. Wu, and Y. Liu. 2023. "Research Themes of Geographical Information Science During 1991–2020: A Retrospective Bibliometric Analysis." *International Journal of Geographical Information Science* 37, no. 2: 243–275.

Yani, M., and A. A. Krisnadhi. 2021. "Challenges, Techniques, and Trends of Simple Knowledge Graph Question Answering: A Survey." *Information* 12, no. 7: 271.

Yin, Z., C. Zhang, D. W. Goldberg, and S. Prasad. 2019. "An NLP-Based Question Answering Framework for Spatio-Temporal Analysis and Visualization." In *Proceedings of the 2019 2nd International Conference on Geoinformatics and Data Analysis*, 61–65.

Zhang, Y., C. Wei, S. Wu, Z. He, and W. Yu. 2023. "Geogpt: Understanding and Processing Geospatial Tasks Through an Autonomous Gpt." *arXiv Preprint arXiv:2307.07930*.

Zhu, Y., H. Yuan, S. Wang, et al. 2023. "Large Language Models for Information Retrieval: A Survey." *arXiv Preprint arXiv:2308.07107*.