

APPLICATION OF GENERATIVE ARTIFICIAL INTELLIGENCE IN GEOGRAPHIC MAP ANALYSIS AND VISUALIZATION: A LITERATURE REVIEW

Nguyen Kim Son*, Luong Thi Minh Hue, Ho Thi Tuyen, Vu Thi Nguyet

TNU - University of Information and Communication Technology

ARTICLE INFO

Received: 20/3/2025
Revised: 26/6/2025
Published: 28/6/2025

KEYWORDS

Generative artificial intelligence
GIS
LLM
Geographic analysis
Map visualization

ABSTRACT

With the rapid advancement of artificial intelligence, generative artificial intelligence has emerged as a pivotal tool in cartography and geospatial analysis. Generative artificial intelligence enhances map accuracy, automates complex spatial analyses, and supports predictive modeling of geographic phenomena. This study provides a comprehensive overview of the field, identifies key trends, and proposes future research directions. Following the PRISMA methodology, research publications from 2014 to 2024 were analyzed. The findings indicate that generative artificial intelligence is primarily applied in automated map generation, spatial pattern recognition, and geographic change prediction. In addition to its potential benefits, this study also examines challenges such as data quality, model interpretability, and ethical considerations in applying generative artificial intelligence to geospatial analysis. Addressing these issues is crucial for optimizing generative artificial intelligence's effectiveness while ensuring accuracy and transparency in real-world applications.

TỔNG QUAN VỀ ỨNG DỤNG CỦA TRÍ TUỆ NHÂN TẠO TẠO SINH TRONG PHÂN TÍCH VÀ TRỰC QUAN HÓA BẢN ĐỒ ĐỊA LÝ

Nguyễn Kim Sơn*, Lương Thị Minh Huệ, Hồ Thị Tuyền, Vũ Thị Nguyệt

Trường Đại học Công nghệ thông tin và Truyền thông – ĐH Thái Nguyên

THÔNG TIN BÀI BÁO

Ngày nhận bài: 20/3/2025
Ngày hoàn thiện: 26/6/2025
Ngày đăng: 28/6/2025

TỪ KHÓA

Trí tuệ nhân tạo tạo sinh
Hệ thống thông tin địa lý
Mô hình ngôn ngữ lớn
Phân tích địa lý
Trực quan hóa bản đồ

TÓM TẮT

Với sự phát triển nhanh chóng của trí tuệ nhân tạo, trí tuệ nhân tạo tạo sinh đang trở thành một công cụ quan trọng trong bản đồ học và phân tích không gian địa lý. Trí tuệ nhân tạo tạo sinh giúp nâng cao độ chính xác của bản đồ, tự động hóa các phân tích không gian phức tạp và hỗ trợ mô hình hóa dự đoán các hiện tượng địa lý. Nghiên cứu này cung cấp một cách tổng quan về lĩnh vực, xác định các xu hướng chính và đề xuất hướng nghiên cứu trong tương lai. Dựa trên phương pháp PRISMA, các nghiên cứu từ năm 2014 đến 2024 đã được phân tích. Kết quả cho thấy trí tuệ nhân tạo tạo sinh chủ yếu được ứng dụng trong tạo bản đồ tự động, nhận diện mẫu không gian và dự đoán sự thay đổi địa lý. Bên cạnh những tiềm năng, nghiên cứu cũng xem xét các thách thức như chất lượng dữ liệu, khả năng giải thích mô hình và vấn đề đạo đức khi ứng dụng trí tuệ nhân tạo tạo sinh vào phân tích không gian. Việc giải quyết những thách thức này sẽ giúp tối ưu hóa hiệu quả của trí tuệ nhân tạo tạo sinh, đồng thời đảm bảo tính chính xác và minh bạch trong các ứng dụng thực tế. Nghiên cứu này cung cấp cơ sở khoa học quan trọng cho các nghiên cứu tiếp theo và định hướng phát triển ứng dụng trong bối cảnh công nghệ đang thay đổi nhanh chóng.

DOI: <https://doi.org/10.34238/tnu-jst.12351>

* Corresponding author. Email: nkson@ictu.edu.vn

1. Giới thiệu

Trong thập kỷ qua, sự phát triển nhanh chóng của công nghệ đã dẫn đến sự ra đời của trí tuệ nhân tạo tạo sinh (GenAI) và các mô hình ngôn ngữ lớn (LLM). Những tiến bộ này đã mở ra nhiều cơ hội mới trong nhiều lĩnh vực, bao gồm cả hệ thống thông tin địa lý (GIS) và khoa học địa lý (GeoScience) [1] - [3]. GenAI, với khả năng tạo ra nội dung mới và hiểu ngôn ngữ tự nhiên, đang dần thay đổi cách chúng ta tương tác và phân tích dữ liệu địa lý [4], [5]. Việc ứng dụng GenAI trong phân tích và trực quan hóa bản đồ hứa hẹn mang lại những đổi mới đột phá cho lĩnh vực GIS. Công nghệ này có thể cải thiện đáng kể khả năng xử lý, phân tích và diễn giải dữ liệu địa lý phức tạp, đồng thời giúp GIS trở nên thân thiện và dễ tiếp cận hơn đối với người dùng không chuyên [6], [7]. Điều này có thể mang lại những góc nhìn mới, đồng thời nâng cao hiệu quả ra quyết định trong các lĩnh vực như quy hoạch đô thị, quản lý tài nguyên và ứng phó thiên tai.

Mặc dù các công nghệ trí tuệ nhân tạo (AI) truyền thống đã được áp dụng trong GIS từ lâu [4], [8], GenAI mang đến một cách tiếp cận đột phá với khả năng thích ứng linh hoạt cho nhiều nhiệm vụ khác nhau, bao gồm tự động tạo dữ liệu, mô hình hóa không gian, nâng cao phân tích địa lý, cũng như hiểu và sinh ngôn ngữ tự nhiên. Nhờ đó, GenAI mở ra những cơ hội mới trong tương tác với dữ liệu địa lý, giúp cải thiện cả tính chính xác và khả năng truy cập thông tin [4], [9], [10]. Tuy nhiên, việc tích hợp GenAI vào GIS vẫn đặt ra nhiều thách thức, bao gồm độ chính xác và độ tin cậy của các mô hình trong bối cảnh địa lý, khả năng mở rộng các giải pháp GenAI, cũng như các vấn đề đạo đức liên quan đến việc sử dụng AI trong phân tích không gian [6], [11], [12]. Một số nghiên cứu đã đề xuất các phương pháp nhằm khắc phục những hạn chế này. Chẳng hạn, Zhu và cộng sự [2] đã phát triển một mô hình ngôn ngữ lớn có ràng buộc kiến thức về dữ liệu địa lý tương tác với GIS, qua đó nâng cao nhận thức của cộng đồng về rủi ro thiên tai. Nghiên cứu của họ cho thấy rằng LLM có thể tạo ra thông tin chính xác về dữ liệu bằng cách tận dụng các thực thể và mối quan hệ trong đồ thị tri thức, qua đó góp phần nâng cao độ tin cậy của phân tích không gian địa lý dựa trên AI. Li và Ning [4] đã giới thiệu khái niệm GIS tự trị, một hệ thống GIS tích hợp AI, khai thác khả năng hiểu ngôn ngữ tự nhiên, lập luận và lập trình của LLM để giải quyết các bài toán không gian. Nghiên cứu này đề xuất một hướng tiếp cận tiềm năng nhằm mở rộng khả năng ứng dụng GenAI trong GIS thông qua năm cấp độ tự trị: tự tạo, tự tổ chức, tự xác minh, tự thực thi và tự phát triển. Bên cạnh đó, Zhang và cộng sự [13] đã giới thiệu GeoGPT - một trợ lý hỗ trợ xử lý các tác vụ địa không gian, trong khi Janowicz và cộng sự [14] phát triển BB-GeoGPT - một mô hình ngôn ngữ lớn chuyên biệt cho phân tích địa không gian. Những nghiên cứu này cùng nhau minh chứng tiềm năng của việc tích hợp GenAI và LLM với GIS nhằm cải thiện độ chính xác, độ tin cậy và khả năng mở rộng trong phân tích và trực quan hóa dữ liệu không gian. Tuy nhiên, hầu hết các nghiên cứu hiện nay vẫn tập trung vào khía cạnh kỹ thuật mà chưa đánh giá đầy đủ tác động thực tế của GenAI trong ứng dụng bản đồ và phân tích không gian địa lý.

Mặc dù GenAI đã đạt được nhiều tiến bộ trong ứng dụng vào GIS, vẫn còn những khoảng trống quan trọng trong việc hiểu và tối ưu hóa quá trình tích hợp công nghệ này vào các quy trình phân tích và trực quan hóa bản đồ. Đặc biệt, những thách thức và hạn chế liên quan đến tính chính xác, độ tin cậy và khả năng mở rộng của GenAI trong lĩnh vực này vẫn chưa được nghiên cứu đầy đủ. Hơn nữa, chưa có một tổng quan hệ thống về các hướng nghiên cứu tương lai, những rào cản kỹ thuật, cũng như các phương pháp tiếp cận hiện tại. Do đó, nghiên cứu này được thực hiện nhằm thu hẹp khoảng trống trên bằng cách trả lời các câu hỏi nghiên cứu cốt lõi:

Câu hỏi 1: Các hướng nghiên cứu trong tương lai về ứng dụng GenAI trong phân tích và trực quan hóa bản đồ là gì?

Câu hỏi 2: Những khó khăn và thách thức nào khi ứng dụng GenAI vào phân tích và trực quan hóa bản đồ?

Câu hỏi 3: Những phương pháp nghiên cứu nào đã được sử dụng trong các nghiên cứu hiện tại về ứng dụng GenAI trong phân tích và trực quan hóa bản đồ?

Bằng cách giải quyết các câu hỏi nghiên cứu trên, nghiên cứu này nhằm cung cấp một cái nhìn

tổng quan về tiềm năng, thách thức và các phương pháp tiếp cận trong việc ứng dụng GenAI vào phân tích và trực quan hóa bản đồ. Đồng thời, nghiên cứu xác định các hướng phát triển quan trọng nhằm định hướng cho các nghiên cứu trong tương lai và hỗ trợ phát triển các ứng dụng GenAI hiệu quả trong GIS.

Phần còn lại của bài báo được tổ chức như sau: Phần 2 trình bày phương pháp nghiên cứu, bao gồm quy trình thu thập, xử lý dữ liệu và tiêu chí đánh giá. Phần 3 tập trung vào phân tích các kết quả chính và thảo luận về ý nghĩa, tác động của các kết quả thu được. Cuối cùng, phần 4 tổng kết nghiên cứu và đưa ra các kết luận quan trọng.

2. Phương pháp nghiên cứu

Bài tổng quan này tuân theo hướng dẫn PRISMA [13] nhằm đảm bảo một phương pháp tiếp cận toàn diện và có hệ thống trong quá trình lựa chọn và phân tích bài báo. Chúng tôi đã thực hiện tìm kiếm có hệ thống trên các cơ sở dữ liệu khoa học để thu thập các nghiên cứu liên quan đến việc ứng dụng GenAI và LLM trong phân tích và trực quan hóa bản đồ. Các bài báo sau đó được phân tích để xác định các xu hướng nghiên cứu chính, những thách thức phổ biến, cũng như các cơ hội phát triển trong tương lai.

2.1. Chiến lược tìm kiếm

Chúng tôi đã thực hiện tìm kiếm có hệ thống các bài báo khoa học được xuất bản trên hai cơ sở dữ liệu điện tử hàng đầu: Web of Science và Scopus. Quá trình tìm kiếm sử dụng các từ khóa kết hợp liên quan đến GenAI và GIS, bao gồm: “Generative Artificial Intelligence Geographic Information System”, “Generative Artificial Intelligence GIS”, “GAI Geographic Information System”, “GAI GIS”, “Geographic Information System Large Language Model”, “GIS Large Language Model”, “Geographic Information System LLM”, và “GIS LLM”. Kết quả tìm kiếm và phân loại tài liệu được trình bày trong Bảng 1.

Bảng 1. Số lượng tìm kiếm từ khóa được trả về bởi mỗi công cụ tìm kiếm

Từ khóa tìm kiếm	Web of Science	Scopus	Tổng số bài báo
Generative Artificial Intelligence GIS	7	4548	4555
GAI Geographic Information System	11	2142	2153
GAI GIS	21	2034	2055
Geographic Information System Large Language Model	103	27125	27228
GIS Large Language Model	189	19088	19177
Geographic Information System LLM	187	133	320
GIS LLM	10	107	117

Trong quá trình tổng hợp tài liệu, cần lưu ý rằng việc sử dụng các từ khóa khác nhau trên cùng một cơ sở dữ liệu hoặc cùng một từ khóa trên các cơ sở dữ liệu khác nhau có thể dẫn đến sự trùng lặp đáng kể trong kết quả tìm kiếm cũng như sự xuất hiện của các tài liệu không liên quan. Do đó, quy trình sàng lọc và lựa chọn tài liệu thủ công, còn được gọi là screening hoặc selection process, đóng vai trò quan trọng trong việc đảm bảo tính chính xác và mức độ phù hợp của các nghiên cứu được đưa vào phân tích.

Quy trình này giúp loại bỏ các bài báo trùng lặp, không liên quan và chỉ giữ lại những nghiên cứu đáp ứng các tiêu chí lựa chọn đã xác định trước. Nhờ đó, tổng quan tài liệu có thể phản ánh một cách khách quan và đáng tin cậy nhất về tình trạng nghiên cứu hiện tại, đồng thời cung cấp cơ sở vững chắc cho các phân tích và thảo luận tiếp theo.

2.2. Tiêu chí đưa vào và loại trừ

Trong phần này, chúng tôi tiến hành sàng lọc các bài báo nhằm loại bỏ những tài liệu trùng lặp hoặc không phù hợp với chủ đề nghiên cứu. Chẳng hạn, một số bài báo có nội dung tập trung vào GenAI, LLM hoặc GIS nhưng không đề cập đến các từ khóa liên quan trong phần tóm tắt sẽ không

được đưa vào phân tích. Các nghiên cứu được đưa vào nếu đáp ứng đầy đủ các tiêu chí sau:

- Bài báo được xuất bản trong khoảng thời gian từ năm 2014 đến 2024.
- Viết bằng tiếng Anh.
- Tập trung vào ứng dụng của GenAI và LLM trong phân tích hoặc trực quan hóa bản đồ địa lý.
- Xuất bản trên các tạp chí khoa học có uy tín.

Chúng tôi loại trừ các nghiên cứu nếu:

- Xuất bản trước năm 2014.
- Không được viết bằng tiếng Anh.
- Tập trung vào GenAI hoặc LLM trong hệ thống thông tin địa lý.
- Không liên quan đến phân tích hoặc trực quan hóa bản đồ.
- Là bài tổng quan lý thuyết hoặc bài viết ý kiến mà không có nghiên cứu thực nghiệm.
- Luận văn thạc sĩ hoặc tiến sĩ.
- Bài báo trùng lặp từ các cơ sở dữ liệu khác nhau.

Để nâng cao độ tin cậy trong quá trình lựa chọn bài báo, nhóm nghiên cứu đã áp dụng kỹ thuật phân loại thẻ đóng (Closed card sorting method), với hai danh mục được xác định trước: liên quan và không liên quan. Mỗi thẻ bao gồm tiêu đề bài báo và tóm tắt nội dung. Quá trình sàng lọc được thực hiện bởi bốn nhà nghiên cứu chuyên sâu, mỗi người độc lập xem xét danh sách kết quả từ hai cơ sở dữ liệu. Sau đó, nhóm tổ chức một phiên thảo luận chuyên môn để so sánh, đánh giá chéo và thống nhất kết quả cuối cùng. Cách tiếp cận này, kết hợp giữa phân loại thẻ đóng và đánh giá chéo, giúp nâng cao tính khách quan, độ tin cậy, đồng thời cải thiện hiệu quả trong việc lựa chọn tài liệu có liên quan cho nghiên cứu.

2.3. Phân tích dữ liệu và đánh giá kết quả

Số lượng bài báo cuối cùng sẽ được phân tích thủ công bởi nhóm tác giả. Quá trình phân tích này sẽ tập trung vào ba khía cạnh chính, tương ứng với ba câu hỏi nghiên cứu đã đặt ra:

Đối với câu hỏi 1, chúng tôi trích xuất và tổng hợp các hướng nghiên cứu tương lai về ứng dụng GenAI trong phân tích và trực quan hóa bản đồ. Các hướng nghiên cứu này sẽ được phân loại theo các chủ đề chính nhằm xác định xu hướng phát triển và cơ hội nghiên cứu tiềm năng.

Để trả lời câu hỏi 2, chúng tôi xác định và tổng hợp các thách thức trong ứng dụng GenAI vào phân tích và trực quan hóa bản đồ. Những thách thức này sẽ được phân nhóm theo các khía cạnh kỹ thuật, đạo đức, pháp lý và xã hội, giúp xây dựng một cái nhìn tổng quan về các rào cản hiện tại trong lĩnh vực này.

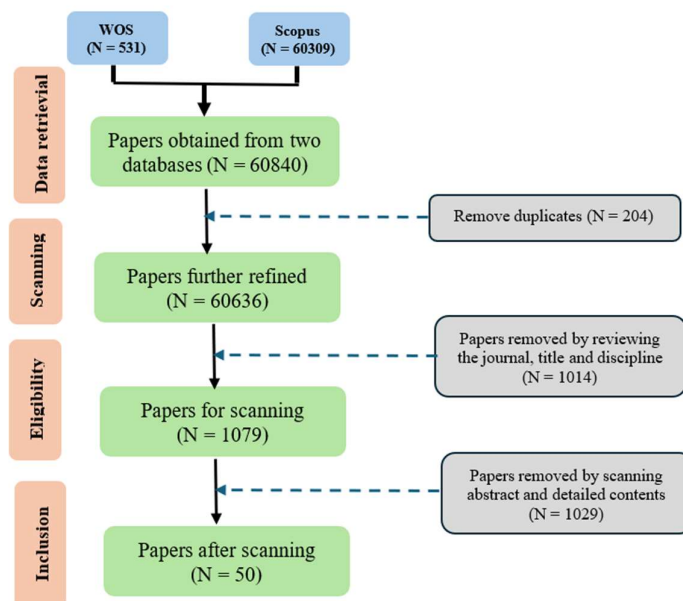
Đối với câu hỏi 3, chúng tôi trích xuất và phân tích các phương pháp nghiên cứu được sử dụng trong các nghiên cứu hiện có về GenAI trong GIS. Các phương pháp này sẽ được phân loại thành định lượng, định tính và hỗn hợp, đồng thời xem xét các kỹ thuật cụ thể như học máy, xử lý ngôn ngữ tự nhiên và mô hình hóa không gian.

Phần mềm MaxQDA sẽ được sử dụng để hỗ trợ quá trình phân tích và tổng hợp kết quả. Chúng tôi sẽ tạo các mã và danh mục trong MaxQDA tương ứng với ba câu hỏi nghiên cứu, giúp tổ chức và phân tích dữ liệu một cách có hệ thống.

Ngoài ra, nghiên cứu cũng tập trung đánh giá các bài báo giới thiệu mô hình GenAI mới trong phân tích và trực quan hóa bản đồ, đặc biệt là các nghiên cứu thực nghiệm so sánh hiệu suất của các mô hình khác nhau. Chúng tôi sẽ tổng hợp kết quả đánh giá hiệu suất trên các bộ dữ liệu chuẩn thường được sử dụng trong lĩnh vực này.

Từ mỗi bài báo được chọn, chúng tôi trích xuất thông tin về hướng nghiên cứu, kỹ thuật GenAI được áp dụng, ứng dụng địa lý, phát hiện chính, cũng như các lợi ích và thách thức đã được đề cập. Dữ liệu được tổng hợp nhằm xác định các chủ đề chung, xu hướng nổi bật và những khoảng trống trong nghiên cứu hiện tại.

Quá trình thu thập và sàng lọc tài liệu, cùng với phương pháp PRISMA để đánh giá tài liệu có hệ thống, được mô tả trong Hình 1. Sau quá trình sàng lọc kỹ lưỡng, 50 bài báo đã được lựa chọn để đưa vào nghiên cứu.



Hình 1. Quy trình PRISMA thu thập bài báo

3. Kết quả và bàn luận

3.1. Các hướng nghiên cứu trong tương lai về ứng dụng GenAI trong phân tích và trực quan hóa bản đồ là gì?

Bảng 2 cho thấy các nghiên cứu hiện tại đã xác định nhiều hướng đi tiềm năng cho việc ứng dụng GenAI vào phân tích và trực quan hóa bản đồ. Một trong những hướng nghiên cứu nổi bật là tự động hóa quá trình tạo bản đồ thông qua các mạng đối kháng tạo sinh (Generative Adversarial Network - GAN), giúp tạo ra các bản đồ có độ phân giải cao và chi tiết chính xác hơn.

Cheng và cộng sự [12] đã phát triển mô hình UrbanGenoGAN, kết hợp giữa GAN, thuật toán di truyền (Genetic Algorithm - GA) và GIS để tạo ra các bản đồ tổng hợp có độ chính xác cao, hỗ trợ quy hoạch không gian đô thị. Đồng thời, nghiên cứu của Smith [15] chỉ ra rằng GAN có thể nâng cao độ phân giải và chi tiết của các bản đồ hiện có, giúp cải thiện quá trình tạo bản đồ tự động.

Ngoài ra, các mô hình GenAI còn được ứng dụng trong việc nhận dạng mẫu không gian trong dữ liệu địa lý phức tạp. Ví dụ, Li và Ning [4] đã phát triển mô hình GIS tự trị, cho phép nhận diện mô hình sử dụng đất và dự đoán sự thay đổi địa lý, từ đó hỗ trợ ra quyết định trong quản lý tài nguyên và quy hoạch đô thị. Bên cạnh đó, GeoGPT, một trợ lý ảo dựa trên trí tuệ nhân tạo, đã chứng minh tiềm năng trong xử lý các tác vụ địa không gian phức tạp thông qua khả năng hiểu và tương tác với dữ liệu ngôn ngữ tự nhiên [13]. Những tiên bộ này cho thấy khả năng mở rộng và cải thiện đáng kể hiệu suất của GIS nhờ vào sự tích hợp của GenAI.

Bảng 2. Bảng số liệu các hướng nghiên cứu tương lai

Ứng dụng	Mô tả	Công nghệ GenAI	Tài liệu tham khảo
Nhận dạng mẫu không gian	Phát hiện khuôn mẫu địa lý từ dữ liệu GIS	CNN, GAN, LLM	[3]
Tạo bản đồ tổng hợp	Cải thiện độ chính xác và chi tiết của bản đồ	GAN, Diffusion Models	[12]
Dự đoán thay đổi địa lý	Dự báo sự thay đổi của các đối tượng địa lý theo thời gian	RNNs, Transformer	[4]
Hỗ trợ phân tích GIS	Tự động phân tích và tạo báo cáo GIS	NLP, LLM	[13]

Các hướng nghiên cứu trong tương lai cần tập trung vào việc phát triển các mô hình GenAI tiên tiến hơn, có khả năng xử lý hiệu quả dữ liệu địa lý phức tạp, đặc biệt trong các lĩnh vực quản lý đô thị, quản lý tài nguyên và ứng phó với thiên tai.

Li [16] đã áp dụng học sâu để phân loại các loại hình sử dụng đất đô thị, từ đó dự đoán sự thay đổi lớp phủ đất trong tương lai, cho thấy tiềm năng lớn của GenAI trong quản lý không gian đô thị. Bên cạnh đó, Wang [3] đã thực hiện một đánh giá có hệ thống về tiềm năng của GenAI, GPT và LLM trong khoa học địa không gian, đặc biệt là trong mô phỏng và dự đoán các hiện tượng địa lý. Những nghiên cứu này nhấn mạnh khả năng cải thiện phân tích và trực quan hóa dữ liệu địa lý của GenAI, đồng thời mở ra nhiều cơ hội mới trong GIS và khoa học thông tin địa lý. Việc tiếp tục nghiên cứu và tối ưu hóa các mô hình GenAI sẽ giúp nâng cao độ chính xác, hiệu quả và khả năng mở rộng trong các ứng dụng địa không gian.

3.2. Những khó khăn và thách thức nào trong việc ứng dụng GenAI vào phân tích và trực quan hóa bản đồ?

Các nghiên cứu hiện tại đã làm nổi bật nhiều thách thức quan trọng trong việc ứng dụng GenAI vào phân tích và trực quan hóa bản đồ. Một trong những trở ngại lớn nhất là độ chính xác và độ tin cậy của các mô hình GenAI khi xử lý dữ liệu địa lý. Chất lượng dữ liệu đầu vào và khả năng của mô hình trong việc nhận dạng các đặc điểm địa lý quan trọng có thể ảnh hưởng đáng kể đến kết quả phân tích và tổng hợp bản đồ [4], [9]. Cheng và cộng sự [17] đã chỉ ra rằng mặc dù mạng đối kháng tạo sinh (GAN) có thể tạo ra các bản đồ tổng hợp chi tiết, nhưng vẫn gặp khó khăn trong việc đảm bảo độ chính xác cao và đáp ứng các yêu cầu kỹ thuật trong nhiều ứng dụng địa lý.

Một thách thức quan trọng khác là khả năng mở rộng của các mô hình GenAI trong bối cảnh phân tích không gian. Li và Ning [4] đã thảo luận về việc tích hợp các LLM vào hệ thống GIS tự trị, nhưng cũng cảnh báo rằng việc mở rộng ứng dụng này để xử lý các tập dữ liệu địa lý lớn và phức tạp vẫn còn là một vấn đề nan giải [18]. Việc xử lý dữ liệu theo thời gian thực và quản lý các lớp thông tin địa lý đa chiều đòi hỏi các mô hình GenAI phải có kiến trúc tối ưu hơn và cần nhiều nghiên cứu để giải quyết bài toán này.

Ngoài ra, các thách thức về đạo đức và quyền riêng tư cũng là một vấn đề đáng quan tâm. Dữ liệu địa lý thường chứa thông tin nhạy cảm về cá nhân và cộng đồng, do đó việc bảo vệ quyền riêng tư và đảm bảo an toàn dữ liệu là rất quan trọng. Wang và cộng sự [11] đã nhấn mạnh rằng việc áp dụng GenAI vào xử lý dữ liệu không gian đòi hỏi các quy tắc đạo đức nghiêm ngặt cùng với các biện pháp bảo vệ dữ liệu cá nhân. GeoGPT cũng gặp những vấn đề tương tự khi thực hiện các tác vụ địa không gian mà không vi phạm quyền riêng tư của người dùng và cộng đồng [19].

Những thách thức này đặt ra yêu cầu cấp bách đối với các nhà nghiên cứu và nhà phát triển trong việc xây dựng các giải pháp GenAI cho GIS. Việc nâng cao độ chính xác, tối ưu hóa khả năng mở rộng và đảm bảo tính đạo đức trong ứng dụng GenAI sẽ là những bước quan trọng để thúc đẩy sự phát triển bền vững của công nghệ này trong phân tích và trực quan hóa bản đồ địa lý.

3.3. Những phương pháp nghiên cứu nào đã được sử dụng trong các nghiên cứu hiện tại về ứng dụng GenAI trong phân tích và trực quan hóa bản đồ?

Nhiều phương pháp khác nhau đã được áp dụng nhằm khám phá tiềm năng của GenAI trong phân tích và trực quan hóa bản đồ. Một trong những phương pháp phổ biến nhất là mạng đối kháng tạo sinh (GAN), được sử dụng để tạo bản đồ tổng hợp có độ phân giải cao và cải thiện chất lượng bản đồ hiện có. Chẳng hạn, Cheng và cộng sự [12] đã phát triển UrbanGenoGAN, kết hợp GAN, thuật toán di truyền (GA) và hệ thống GIS để hỗ trợ quy hoạch không gian đô thị. Bên cạnh đó, Wang [15] đã áp dụng GAN để sửa lỗi băng tần trên ảnh vệ tinh, giúp nâng cao chất lượng dữ liệu không gian phục vụ cho quá trình phân tích.

Ngoài ra, học sâu (deep learning) cũng được sử dụng rộng rãi trong việc phân loại sử dụng đất và dự đoán biến đổi không gian. Ví dụ, Li và Ning [4] đã tích hợp các mô hình học sâu vào GIS tự trị để xử lý dữ liệu địa không gian và hỗ trợ ra quyết định dựa trên các yếu tố địa lý phức tạp.

Wang và cộng sự [2] cũng đánh giá tiềm năng của các mô hình này trong phát hiện mẫu không gian phức tạp, giúp cải thiện khả năng dự đoán và phân tích biến đổi sử dụng đất.

Bên cạnh GAN và học sâu, LLM cũng đóng vai trò quan trọng trong phân tích không gian. Chẳng hạn, GeoGPT, một trợ lý AI trong lĩnh vực GIS, đã chứng minh tiềm năng trong hiểu và xử lý các tác vụ địa không gian, nâng cao hiệu quả phân tích và trực quan hóa dữ liệu không gian [13]. Bảng 3 tổng hợp các phương pháp nghiên cứu chính trong ứng dụng GenAI vào GIS.

Bảng 3. Tổng hợp các phương pháp nghiên cứu

Phương pháp nghiên cứu	Tài liệu tham khảo
Deep Convolutional Neural Networks (DCNN)	[15], [20] - [22]
Generative Adversarial Networks (GAN)	[12], [15], [20]
Natural Language Processing (NLP)	[3], [6], [9], [23] - [25]
Large Language Models (LLM)	[3], [6], [8], [9], [13], [14], [17], [23], [26] - [29]

Tuy nhiên, việc tích hợp GenAI vào nghiên cứu và thực tiễn địa lý cũng đặt ra những thách thức quan trọng liên quan đến tính toàn vẹn của dữ liệu, tính minh bạch của mô hình và đạo đức trong việc sử dụng thông tin không gian do AI tạo ra. Như đã được đề cập trong các nghiên cứu [16], [19], [28], [29], cần phát triển các khung kiểm định chặt chẽ nhằm xác thực độ tin cậy của các kết quả do GenAI tạo ra trong bối cảnh địa lý, đồng thời đảm bảo công nghệ này được triển khai một cách có trách nhiệm và minh bạch.

Để tối ưu hóa ứng dụng của GenAI trong phân tích và trực quan hóa bản đồ, các nghiên cứu sắp tới nên tập trung vào các lĩnh vực sau:

1. Phát triển các kiến trúc GenAI chuyên biệt cho dữ liệu địa lý: Việc thiết kế các mô hình GenAI tập trung vào dữ liệu không gian sẽ giúp nâng cao hiệu suất và độ chính xác khi xử lý bản đồ, từ việc phân loại địa hình, trích xuất thông tin từ ảnh vệ tinh đến mô hình hóa không gian đô thị. Các nghiên cứu trong tương lai có thể tập trung vào việc tối ưu hóa các kiến trúc mô hình, chẳng hạn như sử dụng transformer tùy chỉnh cho dữ liệu địa lý hoặc kết hợp GenAI với hệ thống GIS tiên tiến.

2. Cải thiện khả năng giải thích và minh bạch của mô hình GenAI: Khi các mô hình GenAI ngày càng được áp dụng trong quyết định không gian, cần có cơ chế giúp con người hiểu rõ cách chúng đưa ra dự đoán. Xu hướng trong tương lai sẽ tập trung vào phát triển các phương pháp giải thích mô hình, ví dụ như sử dụng kỹ thuật giải thích dựa trên chú thích hoặc bản đồ nhiệt để hiển thị mức độ tin cậy của dự đoán.

3. Tích hợp GenAI với các công nghệ tiên tiến: Sự kết hợp giữa GenAI và Internet vạn vật (IoT) sẽ cho phép thu thập và phân tích dữ liệu không gian theo thời gian thực, hữu ích trong giám sát môi trường, đô thị thông minh và dự báo thiên tai. Bên cạnh đó, blockchain có thể được sử dụng để đảm bảo tính bảo mật và bất biến của dữ liệu bản đồ, giúp nâng cao độ tin cậy trong các hệ thống GIS hiện đại.

4. Ứng dụng GenAI vào Digital Twins và mô hình hóa không gian: GenAI có tiềm năng lớn trong việc tạo ra các mô hình số hóa mô phỏng thế giới thực, hỗ trợ quy hoạch đô thị, thiết kế hạ tầng và quản lý tài nguyên thiên nhiên. Trong tương lai, các nghiên cứu sẽ hướng tới việc cải thiện độ chính xác của các bản sao số, cho phép tương tác thời gian thực với dữ liệu không gian.

5. Tăng cường tính đạo đức và quy chuẩn chính sách trong ứng dụng GenAI: Việc sử dụng GenAI trong phân tích và trực quan hóa dữ liệu địa lý đặt ra nhiều vấn đề về đạo đức và quyền riêng tư. Các xu hướng nghiên cứu sẽ tập trung vào phát triển các chính sách bảo vệ dữ liệu, đảm bảo tính công bằng trong mô hình AI và giảm thiểu sai lệch khi xử lý thông tin không gian.

6. Tối ưu hóa hiệu suất và khả năng mở rộng của GenAI: Khi khối lượng dữ liệu địa lý ngày càng tăng, việc tối ưu hóa hiệu suất của mô hình GenAI trở nên cần thiết. Xu hướng trong tương lai có thể bao gồm việc áp dụng kỹ thuật học tăng cường, huấn luyện mô hình phân tán hoặc tận dụng các phần cứng chuyên biệt như GPU và TPU để cải thiện tốc độ xử lý.

Bên cạnh đó, các bộ dữ liệu đóng vai trò quan trọng trong phân tích và trực quan hóa dữ liệu địa lý bằng GenAI, bao gồm dữ liệu ảnh vệ tinh (Landsat [11], Sentinel [3], Google Earth Engine [10]), dữ

liệu địa hình và không gian (Shuttle Radar Topography Mission - SRTM [2], OpenStreetMap [9]), dữ liệu khí hậu và môi trường (Moderate Resolution Imaging Spectroradiometer - MODIS [4], ECMWF Reanalysis v5 - ERA5 [6]), dữ liệu dân số và xã hội (WorldPop [7], Global Human Settlement Layer - GHSL [10]) và các nguồn dữ liệu chuyên ngành khác (USGS Earth Explorer [8], FAO GeoNetwork [5]). Việc lựa chọn và kết hợp các bộ dữ liệu phù hợp giúp nâng cao độ chính xác và hiệu quả của các mô hình GenAI trong phân tích không gian.

4. Kết luận

Bài tổng quan này cung cấp một cái nhìn toàn diện về vai trò của GenAI và LLM trong phân tích và trực quan hóa bản đồ địa lý. Những công nghệ này đang dần trở thành công cụ quan trọng trong việc xử lý, phân tích và diễn giải dữ liệu không gian, góp phần nâng cao độ chính xác và khả năng tự động hóa trong GIS. Mặc dù GenAI mang lại nhiều tiềm năng, việc áp dụng rộng rãi công nghệ này vẫn gặp những thách thức đáng kể, đặc biệt là trong các bài toán yêu cầu độ chính xác cao, tính minh bạch và khả năng giải thích của mô hình. Do đó, các nghiên cứu trong tương lai cần tập trung vào phát triển phương pháp đánh giá hiệu quả của GenAI, đồng thời tăng cường khả năng diễn giải và kiểm định tính tin cậy của mô hình trong bối cảnh địa không gian. Khi GenAI tiếp tục phát triển, sự tích hợp của nó vào khoa học thông tin địa lý có thể mở ra những phương pháp phân tích không gian chính xác, hiệu quả và chuyên sâu hơn. Tuy nhiên, để khai thác toàn diện tiềm năng của GenAI trong GIS, cần có sự nghiên cứu liên tục, hợp tác liên ngành, cũng như xem xét kỹ lưỡng các tác động đạo đức của công nghệ này. Bài tổng quan này đóng góp vào kho tri thức đang mở rộng về ứng dụng AI trong địa lý, đồng thời cung cấp nền tảng cho các nghiên cứu và ứng dụng thực tiễn trong tương lai. Việc tiếp tục nghiên cứu và ứng dụng GenAI trong khoa học địa không gian không chỉ giúp thúc đẩy sự phát triển của lĩnh vực này mà còn mở ra nhiều cơ hội mới cho nghiên cứu, quản lý và giáo dục địa lý.

Lời cảm ơn

Nghiên cứu này được hỗ trợ bởi đề tài T2025-07-06, thực hiện tại Trường Đại học Công nghệ Thông tin và Truyền thông, Đại học Thái Nguyên, Việt nam.

TÀI LIỆU THAM KHẢO/ REFERENCES

- [1] S. Silva, L. Alçada-Almeida, and L. C. Dias, "Development of a web-based multi-criteria spatial decision support system for the assessment of environmental sustainability of dairy farms," *Computers and Electronics in Agriculture*, vol. 108, pp. 46-57, 2014.
- [2] J. Zhu *et al.*, "A flood knowledge-constrained large language model interactable with GIS: enhancing public risk perception of floods," *International Journal of Geographical Information Science*, vol. 38, no. 4, pp. 603-625, 2024.
- [3] S. Wang *et al.*, "GPT, large language models (LLMs) and generative artificial intelligence (GAI) models in geospatial science: a systematic review," *International Journal of Digital Earth*, vol. 17, no. 1, 2024, Art. no. 2353122.
- [4] Z. Li and H. Ning, "Autonomous GIS: the next-generation AI-powered GIS," *International Journal of Digital Earth*, vol. 16, no. 2, pp. 4668-4686, 2023.
- [5] S. Seraj and M. R. Delavar, "Developing a GIS-based rough fuzzy set granulation model to handle spatial uncertainty for hydrocarbon structure classification, case study: Fars domain, Iran," *Geo-spatial Information Science*, vol. 25, no. 3, pp. 399-412, 2022.
- [6] C.-H. Chang and G. Kidman, "The rise of generative artificial intelligence (AI) language models-challenges and opportunities for geographical and environmental education," *International Research in Geographical and Environmental Education*, vol. 32, no. 2, pp. 85-89, 2023.
- [7] Y. Gao, A. Herrmann, and C. Chen, "Leveraging GIS and ChatGPT for social goods and higher education," in *Proceedings of the ALISE Annual Conference, (Works-in-Progress Poster)*, 2023, doi: 10.21900/j.alise.2023.1392.
- [8] L. Juhász and B. Guan, "Utilizing Large Language Models in geographic contexts: Experiences from the FIU GIS Center," presented at *the South Florida GIS Expo*, West Palm Beach, FL, Aug. 24, 2023. [Online]. Available: <https://digitalcommons.fiu.edu/gis/103/>. [Accessed Mar. 17, 2025].

- [9] Y. Zhang *et al.*, "BB-GeoGPT: A framework for learning a large language model for geographic information science," *Information Processing & Management*, vol. 61, no. 5, 2024, Art. no. 103808.
- [10] J. Khan *et al.*, "Development and performance evaluation of new AirGIS—a GIS based air pollution and human exposure modelling system," *Atmospheric Environment*, vol. 198, pp. 102-121, 2019.
- [11] A. G. Yeh, "From urban modelling, GIS, the digital, intelligent, and the smart city to the digital twin city with AI," *Environment and Planning B: Urban Analytics and City Science*, vol. 51, no. 5, pp. 1085-1088, 2024.
- [12] W. Cheng *et al.*, "UrbanGenoGAN: pioneering urban spatial planning using the synergistic integration of GAN, GA, and GIS," *Frontiers in Environmental Science*, vol. 11, 2023, Art. no. 1287858.
- [13] Y. Zhang, C. Wei, Z. He, and W. Yu, "GeoGPT: An assistant for understanding and processing geospatial tasks," *International Journal of Applied Earth Observation and Geoinformation*, vol. 131, 2024, Art. no. 103976.
- [14] Z. Liu, K. Janowicz, K. Currier, and M. Shi, "Measuring geographic diversity of foundation models with a natural language-based geo-guessing experiment on GPT-4," *AGILE: GIScience Series*, vol. 5, 2024, Art. no. 38.
- [15] Z. L. Paola, L. S. Jesús, A. H. Christian, and R. U. Sonia, "Correction of banding errors in satellite images with generative adversarial networks (gan)," *IEEE Access*, vol. 11, pp. 51960-51970, 2023.
- [16] N. Khan, Z. Khan, A. Koubaa, M. K. Khan, and R. B. Salleh, "Global insights and the impact of generative AI-ChatGPT on multidisciplinary: a systematic review and bibliometric analysis," *Connection Science*, vol. 36, no. 1, 2024, Art. no. 2353630.
- [17] I. G. Hatvani, M. Mudelsee, and Z. Kern, "Essential questions in earth and geosciences according to large language models," *Open Geosciences*, vol. 16, no. 1, 2024, Art. no. 20220677.
- [18] S. Thomas and D. Tishechkin, "Leveraging Generative AI to Simplify Spatial Analysis and Visualization of Digital Twins," in *Fourth EAGE Digitalization Conference & Exhibition*, European Association of Geoscientists & Engineers, 2024, vol. 2024, no. 1, pp. 1-5.
- [19] S. Pang, E. Nol, and K. Heng, "ChatGPT-4o for English language teaching and learning: Features, applications, and future prospects," *Cambodian Journal of Educational Research*, vol. 4, no. 1, pp. 35–56–35–56, 2024.
- [20] Y. Kang, S. Gao, and R. E. Roth, "Transferring multiscale map styles using generative adversarial networks," *International Journal of Cartography*, vol. 5, no. 2-3, pp. 115-141, 2019.
- [21] Q. Qiu, Z. Xie, L. Wu, and L. Tao, "GNER: A generative model for geological named entity recognition without labeled data using deep learning," *Earth and Space science*, vol. 6, no. 6, pp. 931-946, 2019.
- [22] Z. Li, B. Chen, S. Wu, M. Su, J. M. Chen, and B. Xu, "Deep learning for urban land use category classification: A review and experimental assessment," *Remote Sensing of Environment*, vol. 311, 2024, Art. no. 114290.
- [23] S. C. Tan, W. Chen, and B. L. Chua, "Leveraging generative artificial intelligence based on large language models for collaborative learning," *Learning: Research and Practice*, vol. 9, no. 2, pp. 125-134, 2023.
- [24] R. Ramachandran, M. Ramasubramanian, P. Koirala, I. Gurung, and M. Maskey, "Language model for earth science: Exploring potential downstream applications as well as current challenges," in *IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium*, 2022, pp. 4015-4018.
- [25] P. Gupta, B. Ding, C. Guan, and D. Ding, "Generative AI: A systematic review using topic modelling techniques," *Data and Information Management*, vol. 8, no. 2, 2024, Art. no. 100066.
- [26] P. Mooney, W. Cui, B. Guan, and L. Juhász, "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*, 2023, pp. 85-94.
- [27] A. Hadid, T. Chakraborty, and D. Busby, "When geoscience meets generative AI and large language models: Foundations, trends, and future challenges," *Expert Systems*, vol. 41, no. 10, 2024, Art. no. e13654.
- [28] A. Crooks and Q. Chen, *Exploring the new frontier of information extraction through large language models in urban analytics*, vol. 51, SAGE Publications Sage UK: London, England, 2024, pp. 565-569.
- [29] N. Fulman, A. Memduhoğlu, and A. Zipf, "Distortions in judged spatial relations in large language models," *The Professional Geographer*, vol. 76, no. 6, pp. 703-711, 2024.