

OVERVIEW OF APPLICATION OF GENERATIVE ARTIFICIAL INTELLIGENCE IN SOFTWARE SOURCE CODE GENERATION

Nguyen Van Viet^{1*}, Nguyen Huu Khanh², Nguyen The Vinh¹,
Vu Van Dien¹, Nguyen Kim Son¹, Luong Thi Minh Hue¹

¹TNU of Information and Communication Technology, ²Thai Nguyen University

ARTICLE INFO		ABSTRACT
Received:	13/3/2025	This paper provides an overview of the application of generative artificial intelligence in the process of software source code generation. Large language models such as GPT-4, CodeBERT, Codex, and AlphaCode are helping programmers automate many tasks, including generating code from natural language descriptions, detecting programming errors, optimizing code, and improving software maintainability. The study uses the PRISMA method to analyze scientific literature from Web of Science during 2021-2025, focusing on important topics and research trends of Large language models in software engineering. The results show that the number of articles on this topic increased sharply in 2024, reflecting the growing interest in artificial intelligence in software development. The studies also show that Elsevier and IEEE are the two sources of documents with the largest number of publications in this field. Although generative artificial intelligence offers many benefits, the study also addresses important challenges such as code accuracy, error detection, security and privacy issues. Integrating generative artificial intelligence into the software development process requires appropriate approaches to exploit the full potential of this technology. The paper concludes that research on Large language models in software engineering still has many gaps, opening up opportunities for new directions of development in the future.
Revised:	26/6/2025	
Published:	28/6/2025	

KEYWORDS

Generative artificial intelligence
Software engineering
Transformer
Artificial intelligence
PRISMA

TỔNG QUAN VỀ ỨNG DỤNG TRÍ TUỆ NHÂN TẠO TRONG QUÁ TRÌNH PHÁT SINH MÃ NGUỒN PHẦN MỀM

Nguyễn Văn Việt^{1*}, Nguyễn Hữu Khánh², Nguyễn Thế Vinh¹,
Vũ Văn Diện¹, Nguyễn Kim Sơn¹, Lương Thị Minh Huệ¹

¹Trường Đại học Công nghệ thông tin và Truyền thông – ĐH Thái Nguyên, ²Đại học Thái Nguyên

THÔNG TIN BÀI BÁO		TÓM TẮT
Ngày nhận bài:	13/3/2025	Bài báo tổng quan về ứng dụng trí tuệ nhân tạo trong phát sinh mã nguồn phần mềm, với trọng tâm là các mô hình ngôn ngữ lớn như GPT-4, CodeBERT, Codex và AlphaCode. Các mô hình này hỗ trợ lập trình viên tự động hóa nhiều tác vụ như sinh mã từ mô tả ngôn ngữ tự nhiên, phát hiện lỗi, tối ưu mã và cải thiện bảo trì phần mềm. Nghiên cứu áp dụng phương pháp PRISMA để phân tích tài liệu từ Web of Science giai đoạn 2021-2025, tập trung vào xu hướng và chủ đề quan trọng trong kỹ thuật phần mềm. Kết quả cho thấy sự gia tăng mạnh mẽ số lượng nghiên cứu vào năm 2024, đặc biệt từ các nguồn Elsevier và IEEE. Dù trí tuệ nhân tạo mang lại nhiều lợi ích, bài báo cũng đề cập đến các thách thức như độ chính xác của mã sinh, lỗi ảo giác, vấn đề bảo mật và quyền riêng tư. Việc tích hợp trí tuệ nhân tạo vào phát triển phần mềm đòi hỏi phương pháp tiếp cận phù hợp để khai thác tối đa tiềm năng công nghệ. Bài báo kết luận rằng nghiên cứu về trí tuệ nhân tạo trong kỹ thuật phần mềm vẫn còn nhiều khoảng trống, mở ra cơ hội phát triển trong tương lai.
Ngày hoàn thiện:	26/6/2025	
Ngày đăng:	28/6/2025	

TỪ KHÓA

Trí tuệ nhân tạo
Kỹ thuật phần mềm
Transformer
Trí tuệ nhân tạo
PRISMA

DOI: <https://doi.org/10.34238/tnu-jst.12305>

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

1. Giới thiệu

Trí tuệ nhân tạo tạo sinh (Generative artificial intelligence - GenAI) là một nhánh của trí tuệ nhân tạo, nó tập trung vào việc tạo ra các dữ liệu mới như văn bản, hình ảnh, âm thanh, mã nguồn... thay vì phân tích dữ liệu, thì nó sẽ học cách tạo ra các dữ liệu mới dựa vào các dữ liệu đã được đào tạo trước đó. Trí tuệ nhân tạo tạo sinh sử dụng các thuật toán của học máy để phân tích lượng dữ liệu lớn đã được đào tạo trước đó, hiểu các mẫu và tương quan trong dữ liệu đã được đào tạo, qua đó nó có thể tạo ra các dữ liệu tương tự trước đó mà không hoàn toàn giống nhau.

Trí tuệ nhân tạo tạo sinh mang lại nhiều lợi ích quan trọng, bao gồm khả năng tạo ra dữ liệu mới, tự động hóa các tác vụ lặp đi lặp lại, tiết kiệm thời gian, hỗ trợ ra quyết định, giải quyết các vấn đề phức tạp, cũng như thúc đẩy đổi mới sáng tạo trong nhiều lĩnh vực khác nhau. Tuy nhiên, trong bối cảnh hiện nay, việc phát triển và triển khai các mô hình trí tuệ nhân tạo, đặc biệt là các mô hình ngôn ngữ lớn (Large Language Models - LLMs) đang đặt ra những thách thức đáng kể. Cụ thể, chi phí đào tạo các mô hình này rất cao, đồng thời tiềm ẩn các rủi ro liên quan đến việc xử lý thông tin nhạy cảm trong các lĩnh vực như y tế, tài chính và thương mại điện tử.

Bài báo này nhằm mục đích nghiên cứu tổng quan về việc sử dụng GenAI trong lĩnh vực kỹ thuật phần mềm. Các nhà phát triển phần mềm, kiểm thử phần mềm có thể sử dụng mô hình ngôn ngữ lớn trong các pha phát triển, xây dựng. Sử dụng LLMs trong lĩnh vực kỹ thuật phần mềm là cơ hội để giúp cho các công ty phát triển phần mềm, các nhóm nghiên cứu triển khai được nhanh chóng, giảm thiểu nguồn nhân lực, tài chính.

Các mô hình ngôn ngữ lớn đang tạo ra bước tiến đột phá trong kỹ thuật phần mềm, đặc biệt trong các tác vụ liên quan đến mã nguồn như sinh mã, sửa lỗi, kiểm thử, đánh giá mã [1] – [3] và đánh giá tổng quan LLMs trong lĩnh vực giáo dục, pháp lý [4], [5]. Nhiều nghiên cứu gần đây đã khảo sát toàn diện khả năng của LLMs trong các bối cảnh này, từ các khung phân tích mô hình [6] đến phân loại thách thức và định hướng tương lai [3]. Các mô hình như ChatGPT và Copilot đã được ứng dụng rộng rãi nhờ khả năng hỗ trợ lập trình theo thời gian thực [7], [8]. Tuy nhiên, bên cạnh tiềm năng lớn, các vấn đề như độ chính xác, tính minh bạch và độ tin cậy vẫn là những thách thức cần giải quyết [6], [9], [10]. Các nghiên cứu đã chỉ ra rằng việc áp dụng LLMs trong các tác vụ quan trọng như sửa lỗi hay đánh giá mã đòi hỏi phải có các tiêu chuẩn đánh giá và xác thực rõ ràng [7]. Dù vậy, các kết quả thực nghiệm và đánh giá gần đây cho thấy LLMs vẫn là hướng đi đầy hứa hẹn trong phát triển phần mềm dựa trên trí tuệ nhân tạo (Artificial intelligence – AI) [1], [3]. Bài báo này nhằm tổng hợp và phân tích các xu hướng nghiên cứu tiêu biểu liên quan đến ứng dụng LLMs trong kỹ thuật phần mềm hiện nay.

Để giải quyết các vấn đề nêu trên, các nhóm nghiên cứu, nhà khoa học đã nêu, trình bày tiềm năng, cơ hội, khó khăn, thách thức đối với việc sử dụng phân tích tổng quan về trí tuệ nhân tạo tạo sinh trong lĩnh vực kỹ thuật phần mềm qua các câu hỏi:

Câu hỏi 1: Số lượng bài báo nghiên cứu trong cơ sở dữ liệu tạp chí, hội thảo đối với việc sử dụng trí tuệ nhân tạo tạo sinh trong lĩnh vực kỹ thuật phần mềm thay đổi như thế nào?

Câu hỏi 2: Những từ khóa nào xuất hiện thường xuyên nhất về sử dụng trí tuệ nhân tạo tạo sinh trong lĩnh vực kỹ thuật phần mềm của kho dữ liệu tạp chí, hội thảo?

Câu hỏi 3: Những chủ đề nghiên cứu quan trọng nhất trong việc sử dụng trí tuệ nhân tạo tạo sinh trong lĩnh vực kỹ thuật phần mềm là gì?

Câu hỏi 4: Những khoảng trống và lĩnh vực cho nghiên cứu trong tương lai là gì?

Việc trả lời các câu hỏi nghiên cứu ở phía trên sẽ giúp cho các nhà phát triển phần mềm có được những góc nhìn cơ bản để tiếp cận theo trí tuệ nhân tạo tạo sinh trong lĩnh vực kỹ thuật phần mềm. Các khoảng trống trong nghiên cứu hiện tại có thể mở ra những hướng đi mới cho các nhà nghiên cứu trong tương lai.

2. Phương pháp nghiên cứu

Bài viết sử dụng phương pháp nghiên cứu tổng quan PRISMA (Preferred Reporting Items for

Systematic Reviews and Meta-Analyses). Nhóm nghiên cứu sẽ áp dụng phương pháp tổng quan hệ thống PRISMA để thực hiện việc tìm kiếm và thống kê các bài báo khoa học từ các cơ sở dữ liệu như Scopus, Google Scholar, IEEE Xplore... bằng cách sử dụng các từ khóa cụ thể. Sau đó, các bài báo không đáp ứng tiêu chí lựa chọn sẽ được loại trừ khỏi danh sách nghiên cứu. Phương pháp PRISMA có các các tiêu chí đánh giá chất lượng trong nghiên cứu, minh họa số lượng các bài báo ở từng giai đoạn (tìm kiếm, sàng lọc, đủ điều kiện và lựa chọn phân tích), trích xuất các từ khóa, quốc gia để đưa ra được phân tích tổng quan chi tiết giúp nghiên cứu viên có thể đánh giá độ tin cậy của các nghiên cứu trước khi sử dụng chúng trong nghiên cứu của mình.

2.1. Nguồn tìm kiếm

Nghiên cứu này hoàn toàn dựa trên Web of Science (WoS) như một nguồn tìm kiếm chính để thu thập dữ liệu. WoS là một nền tảng cung cấp cơ sở dữ liệu học thuật mạnh mẽ, hỗ trợ tìm kiếm và truy xuất các tài liệu nghiên cứu khoa học có giá trị. Hệ thống này giúp các nhà nghiên cứu tiếp cận nhanh chóng với những công trình quan trọng, theo dõi xu hướng nghiên cứu hiện tại, đánh giá mức độ ảnh hưởng của bài báo dựa trên số lần trích dẫn, đồng thời mở rộng cơ hội hợp tác khoa học. Nhờ các công cụ phân tích hiện đại, WoS không chỉ hỗ trợ việc tìm kiếm mà còn giúp định hướng nghiên cứu một cách chính xác và kịp thời. Điều này đảm bảo rằng nghiên cứu có thể tổng hợp và đánh giá thông tin một cách toàn diện từ các nguồn dữ liệu chất lượng cao, phù hợp với việc phân tích thư mục và tổng quan tài liệu có hệ thống. Hạn chế của WoS là không bao gồm nguồn dữ liệu mở như Google Scholar hay các bài báo chưa được xuất bản chính thức.

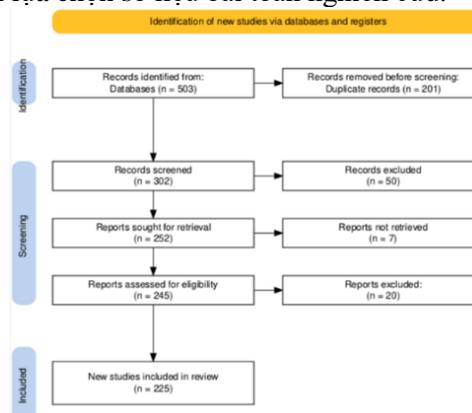
2.2. Tiêu chí tìm kiếm và điều kiện

Tác giả lựa chọn các bài báo để phân tích tổng quan từ cơ sở dữ liệu tạp chí, hội thảo trong lĩnh vực kỹ thuật phần mềm ứng dụng Generative AI. Tiêu chí lựa chọn bài báo gồm:

- i) Thuật ngữ tìm kiếm: ít nhất một thuật ngữ liên quan đến Generative AI (GenAI), Large language Models, Code, Software Engineering (SE) phải xuất hiện trong tiêu đề bài viết;
- ii) Thuật ngữ “Generative AI”, Software Engineering (SE), Large language model;
- iii) Loại trừ những bài báo được xuất bản từ 2020 trở về trước;
- iv) Thực hiện loại trừ những bài báo không liên quan giữa GenAI và Kỹ thuật phần mềm.

Qua quá trình tìm kiếm và sàng lọc dữ liệu với tổng số 503 bài báo, thực hiện tổng hợp các bài báo liên quan đến tiêu chí trên còn lại 225 bài đưa vào phân tích với thời gian xuất bản các bài báo, nghiên cứu từ năm 2021 đến 2025.

Hình 1 mô tả luồng thông tin qua các giai đoạn khác nhau để đánh giá hệ thống và sử dụng phương pháp PRISMA. Trong đó Identification là số lượng bài báo được tìm kiếm, Screening thực hiện sàng lọc theo các tiêu chí của tác giả, Eligibility – lựa chọn các bài báo đủ điều kiện truy cập, Included – giai đoạn lựa chọn số liệu bài toán nghiên cứu.

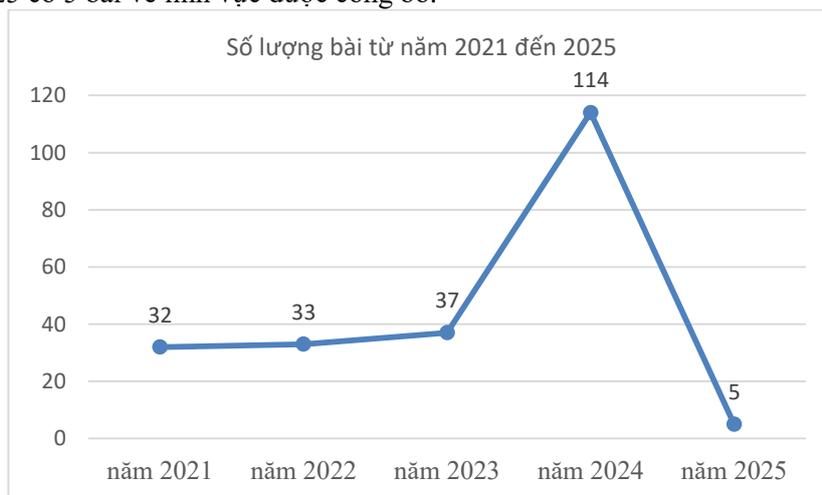


Hình 1. Biểu đồ thể hiện sự di chuyển của thông tin qua các giai đoạn khác nhau của một đánh giá hệ thống

3. Kết quả và bàn luận

3.1. Số lượng bài báo nghiên cứu trong cơ sở dữ liệu tạp chí, hội thảo đối với việc sử dụng LLMs trong lĩnh vực kỹ thuật phần mềm thay đổi như thế nào?

Biểu đồ Hình 2 cho thấy tổng quan phân phối số lượng các tài liệu sử dụng về mô hình ngôn ngữ lớn trong lĩnh vực kỹ thuật phần mềm từ năm 2021 đến 2025. Với tổng số bài được phân tích là 225, số lượng các bài báo được phân bố không đồng đều, các lĩnh vực này phát triển chậm trong giai đoạn từ 2021 đến 2023 với số lượng liên quan trong cơ sở dữ liệu tìm kiếm từ 32 đến 37 bài, đến năm 2024 các dữ liệu xuất bản liên quan đến từ khóa tăng đột biến lên đến 114 bài. Đầu năm 2025 có 5 bài về lĩnh vực được công bố.



Hình 2. Số lượng bài báo, ấn phẩm được xuất bản theo năm từ 2021 đến 2025

Bảng 1. Các bài báo về sử dụng trí tuệ nhân tạo sinh về lĩnh vực phần mềm, mã nguồn được xuất bản trên các tạp chí

Nguồn tạp chí	Số lượng bài báo
Elsevier	50
IEEE	47
Springer	36
Assoc Computing Machinery	17
Các tạp chí khác	75

Từ Bảng 1, ta thấy rằng Elsevier đóng góp nhiều bài viết nhất có liên quan đến mô hình ngôn ngữ lớn, kỹ thuật phần mềm, trí tuệ nhân tạo với 50 bài, chiếm tỷ lệ 22% bài được phân tích. Đứng thứ 2 là IEEE với 47 bài, chiếm 21%. Đứng thứ 3 là Springer với 36 bài, chiếm 16%. Các tạp chí, hội thảo còn lại gần tương đương nhau. Kết quả này cung cấp thông tin hữu ích cho các nhà nghiên cứu về LLMs trong các lĩnh vực về kỹ thuật phần mềm, AI, ChatGPT... Tỷ lệ 22% các công bố xuất hiện trên các tạp chí uy tín có thể được lý giải bởi ba nguyên nhân chính. Thứ nhất, các nghiên cứu trong lĩnh vực này có tính mới cao và chứa hàm lượng học thuật đáng kể. Thứ hai, phần lớn các nghiên cứu đang đóng vai trò quan trọng trong việc thúc đẩy các hướng phát triển của các nhóm nghiên cứu quy mô nhỏ. Cuối cùng, chất lượng học thuật cao của các nghiên cứu đã được các tạp chí khoa học ghi nhận thông qua việc công nhận những đóng góp thiết thực cho cộng đồng nghiên cứu.

Từ Bảng 2, ta có thể thấy rằng bài báo nghiên cứu “Deep Learning for Code Intelligence: Survey, Benchmark and Toolkit” đăng trên *Assoc Computing Machinery* của nhóm tác giả Wan và cộng sự [1] được các nhà khoa học quan tâm và trích dẫn nhiều nhất với số lượt trích dẫn là 309 trong cơ sở dữ liệu của WoS. Điều này cho thấy vị trí và tầm quan trọng của công trình của họ trong lĩnh vực này. Các bài viết khác nghiên cứu về LLMs trong lĩnh vực kỹ thuật phần mềm

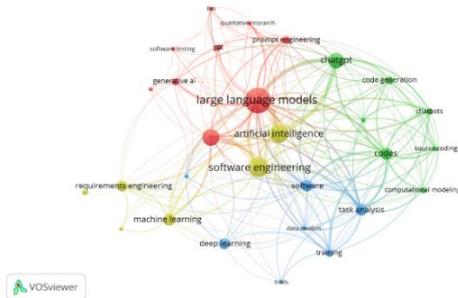
cũng được lượng trích dẫn lớn như “: A Model-Based Universal Analysis Framework for Large Language Models” [6]; “Decoding ChatGPT: A taxonomy of existing research, current challenges, and possible future directions” [7]. Bài viết của các tác giả còn lại có số lượng trích dẫn cao cho thấy sự quan tâm và tính mới của vấn đề này.

Bảng 2. Top 05 bài báo có số lượt được trích dẫn nhiều nhất trong cơ sở dữ liệu có liên quan đến lĩnh vực kỹ thuật phần mềm, trí tuệ nhân tạo tạo sinh

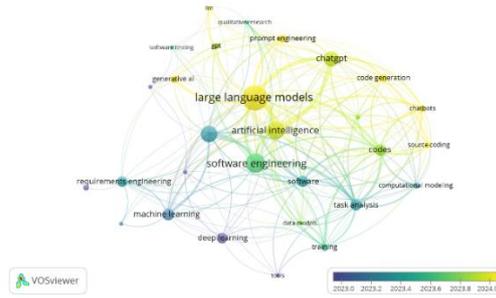
Tên bài báo	Lượt được trích dẫn
Deep Learning for Code Intelligence: Survey, Benchmark and Toolkit [1]	309
LUNA: A Model-Based Universal Analysis Framework for Large Language Models [6]	253
Decoding ChatGPT: A taxonomy of existing research, current challenges, and possible future directions [7]	227
A Systematic Review of AI-Enabled Frameworks in Requirements Elicitation [11]	184
Software Testing with Large Language Models: Survey, Landscape, and Vision [2]	164

3.2. Những từ khoá nào xuất hiện thường xuyên nhất về sử dụng trí tuệ nhân tạo tạo sinh trong lĩnh vực kỹ thuật phần mềm của kho dữ liệu tạp chí, hội thảo?

Hình 3 trình bày bản đồ trực quan về các xu hướng nghiên cứu, được tạo bằng phần mềm VOSviewer, nhằm phân tích mối quan hệ giữa lĩnh vực kỹ thuật phần mềm, mô hình ngôn ngữ lớn và trí tuệ nhân tạo trong giai đoạn từ năm 2021 đến 2025. Kết quả lập bản đồ cho thấy các cụm chủ đề trọng tâm, phản ánh mức độ ảnh hưởng của trí tuệ nhân tạo, đặc biệt là trí tuệ nhân tạo tạo sinh, đối với lĩnh vực kỹ thuật phần mềm...



Hình 3. Các lĩnh vực nghiên cứu liên quan giữa kỹ thuật phần mềm, trí tuệ nhân tạo tạo sinh và trí tuệ nhân tạo

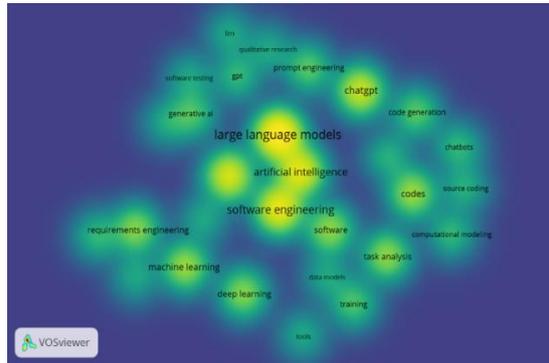


Hình 4. Sự phân bố theo năm các lĩnh vực nghiên cứu liên quan đến kỹ thuật phần mềm, trí tuệ nhân tạo tạo sinh và trí tuệ nhân tạo

Verify selected keywords

Selected	Keyword	Occurrences	Total link strength
<input checked="" type="checkbox"/>	large language models	40	117
<input checked="" type="checkbox"/>	artificial intelligence	27	83
<input checked="" type="checkbox"/>	software engineering	29	75
<input checked="" type="checkbox"/>	codes	16	70
<input checked="" type="checkbox"/>	chatgpt	22	66
<input checked="" type="checkbox"/>	task analysis	15	65
<input checked="" type="checkbox"/>	software	15	62
<input checked="" type="checkbox"/>	natural language processing	24	55
<input checked="" type="checkbox"/>	chatbots	7	42
<input checked="" type="checkbox"/>	training	9	42
<input checked="" type="checkbox"/>	code generation	10	40
<input checked="" type="checkbox"/>	prompt engineering	10	37
<input checked="" type="checkbox"/>	source coding	6	33
<input checked="" type="checkbox"/>	machine learning	16	32
<input checked="" type="checkbox"/>	software development management	6	30
<input checked="" type="checkbox"/>	computational modeling	7	29
<input checked="" type="checkbox"/>	data models	5	27
<input checked="" type="checkbox"/>	requirements engineering	14	23

Hình 5a. Từ khóa được trích xuất từ các nội dung tóm tắt của các bài báo liên quan đến Kỹ thuật phần mềm, trí tuệ nhân tạo, trí tuệ nhân tạo tạo sinh



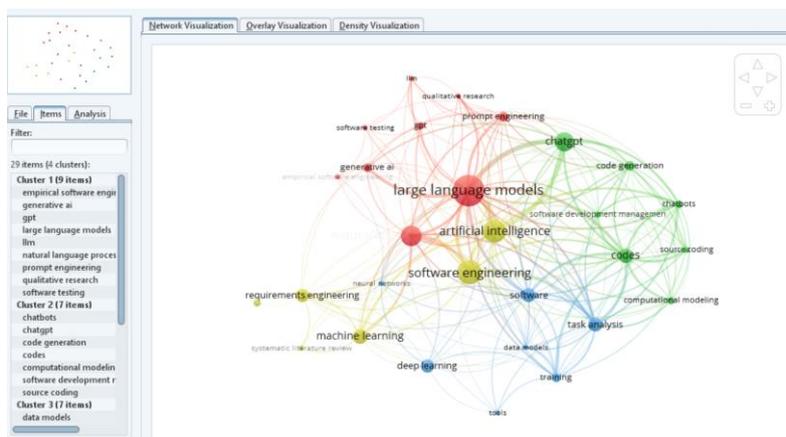
Hình 5b. Đám mây từ khóa được trích xuất từ các bài báo liên quan đến Kỹ thuật phần mềm, trí tuệ nhân tạo, trí tuệ nhân tạo tạo sinh

Hình 4 minh họa sự phân bố theo thời gian của các chủ đề nghiên cứu có mối liên kết chặt chẽ, phản ánh mức độ quan tâm và số lượng công trình khoa học trong các lĩnh vực như kỹ thuật phần mềm, mô hình ngôn ngữ lớn và trí tuệ nhân tạo. Đồng thời, Hình 4 cũng thể hiện sự kết nối với các chủ đề liên quan khác như lập trình, sinh mã tự động, phân tích nhiệm vụ và mã nguồn...

Hình 5a và 5b thể hiện những từ khóa xuất hiện thường xuyên nhất như mô hình ngôn ngữ lớn (LLMs) với tổng số 117 lần; Artificial intelligence với 83 lần; Software Engineering với 75 lần, code với 75 lần...

3.3. Những chủ đề nghiên cứu quan trọng nhất trong việc sử dụng trí tuệ nhân tạo tạo sinh trong lĩnh vực kỹ thuật phần mềm là gì?

Hình 6 nhấn mạnh việc tổng hợp các từ khóa đã được sử dụng trong các nghiên cứu giúp cho người đọc hiểu rõ hơn về những mối quan hệ giữa các từ khóa này. Từ bảng phân tích có các cụm tập trung nội dung Kỹ thuật phần mềm, ChatGPT, thẩm định, xác minh... như sau: Cụm số 1 (Cluster 1) với 9 mục, cụm số 2 (Cluster 2) với 7 mục, cụm số 3 (Cluster 3) với 7 mục,... Từ khoá "Large language models" có tần suất xuất hiện cao nhất với 117 lần, từ khóa "Artificial intelligence" đứng thứ 2 với 83 lần xuất hiện, từ khóa "Software Engineering" đứng thứ 3 với 75 lần xuất hiện... Những chủ đề này có thể là những hướng nghiên cứu quan trọng trong tương lai để tối ưu hóa sức mạnh của trí tuệ nhân tạo tạo sinh trong lĩnh vực Kỹ thuật phần mềm, phát sinh mã nguồn,....



Hình 6. Phân tích sự xuất hiện của từ khóa

3.4. Những khoảng trống và lĩnh vực cho nghiên cứu trong tương lai là gì?

Trí tuệ nhân tạo tạo sinh đã được nghiên cứu ở mức độ trung bình từ 2021 đến 2023 nhưng đến năm 2024, 2025 đã dần dần phát triển mạnh, nhưng vẫn còn một số hạn chế và thách thức cần phải vượt qua. Bảng 3 chỉ ra tần suất các hạn chế, thách thức của trí tuệ nhân tạo tạo sinh, kỹ thuật phần mềm được nêu trong các ấn phẩm.

Bảng 3. Tần suất các hạn chế được chỉ ra trong các ấn phẩm

Hạn chế/ thách thức	Ấn phẩm tham khảo
Không nắm bắt được một số bảo mật nhất định	[2], [11]
Không khái quát hóa và lập trình các ngôn ngữ khác	[11]
Dung lượng bộ nhớ	[12]
Đánh giá chất lượng	[13], [14]
Nội dung sai lệch	[8], [15], [16]
Phương pháp tiếp cận	[17] - [19]
Mối quan tâm về khả năng tái sản xuất	[2]

Những hạn chế này cho thấy vẫn còn nhiều thách thức trong việc phát triển và ứng dụng AI,

đòi hỏi các giải pháp tối ưu hơn để nâng cao hiệu quả và độ tin cậy của hệ thống. Qua các thông tin nêu trên có thể thấy đề tài “Tổng quan về ứng dụng trí tuệ nhân tạo sinh trong quá trình phát sinh mã nguồn phần mềm” là một hướng nghiên cứu mới, cấp thiết.

Nhóm nghiên cứu cũng đề xuất một số giải pháp để nâng cao hiệu quả ứng dụng trí tuệ nhân tạo sinh trong quá trình phát sinh mã nguồn phần mềm nói riêng cũng như lĩnh vực kỹ thuật phần mềm nói chung là: Cải tiến các thuật toán bảo mật AI, áp dụng kỹ thuật mã hóa dữ liệu, xác thực đa yếu tố và tăng cường cơ chế giám sát an toàn thông tin; Tối ưu hóa mô hình bằng phương pháp nén trọng số (quantization, pruning) và sử dụng kiến trúc mô hình nhẹ hơn như distillation models; Phát triển mô hình đa ngôn ngữ bằng cách mở rộng tập dữ liệu huấn luyện, kết hợp kỹ thuật fine-tuning trên nhiều ngữ cảnh lập trình khác nhau; Chuẩn hóa quy trình huấn luyện và thử nghiệm, sử dụng các nền tảng AI có khả năng tái tạo kết quả và công bố mã nguồn mở để đảm bảo tính minh bạch; Xây dựng các bộ tiêu chí đánh giá khách quan, sử dụng phương pháp kiểm thử chéo (cross-validation) và tích hợp đánh giá của chuyên gia để xác minh độ chính xác.

4. Kết luận

Bài viết này cung cấp một nghiên cứu tổng quan về ứng dụng trí tuệ nhân tạo sinh trong lĩnh vực công nghệ phần mềm, tập trung phân tích và đánh giá các bài báo, tài liệu liên quan đến việc sử dụng GenAI trong giai đoạn từ năm 2021 đến năm 2025. Kết quả nghiên cứu cho thấy, từ năm 2021 đến 2023, số lượng nghiên cứu về GenAI còn hạn chế, chỉ chiếm khoảng 45%, nhưng đã có sự gia tăng đột biến vào năm 2024 đạt 51%, đặc biệt trong các lĩnh vực như kỹ thuật phần mềm, kiểm thử, ChatGPT và kiểm định.

Các tạp chí và hội thảo hàng đầu trong lĩnh vực này bao gồm Elsevier và IEEE, trong đó bài báo "Deep Learning for Code Intelligence: Survey, Benchmark and Toolkit" được trích dẫn 309 lần, trở thành một trong những nghiên cứu có ảnh hưởng lớn nhất về trí tuệ nhân tạo sinh trong kỹ thuật phần mềm. Ngoài ra, các nghiên cứu về trí tuệ nhân tạo sinh trong các lĩnh vực như kỹ thuật phần mềm, kiểm định, giáo dục và ChatGPT chủ yếu được công bố tại các hội nghị và tạp chí có uy tín.

Tuy nhiên, vẫn còn nhiều khía cạnh của trí tuệ nhân tạo sinh trong kỹ thuật phần mềm và ChatGPT chưa được phân tích sâu trong các nghiên cứu trước đây. Bài tổng quan này không chỉ giúp nhận diện những khoảng trống nghiên cứu mà còn đóng vai trò quan trọng trong việc định hướng các nghiên cứu tương lai về trí tuệ nhân tạo sinh trong lĩnh vực công nghệ phần mềm.

Lời cảm ơn

Nghiên cứu này được hỗ trợ bởi đề tài T2025-07-03 được thực hiện tại Trường Đại học Công nghệ Thông tin và Truyền thông – Đại học Thái Nguyên, Thái Nguyên, Việt Nam.

TÀI LIỆU THAM KHẢO/ REFERENCES

- [1] Y. Wan *et al.*, “Deep learning for code intelligence: Survey, benchmark and toolkit,” *ACM Comput. Surv.*, vol. 56, no. 12, pp. 1–41, 2024.
- [2] J. Wang, Y. Huang, C. Chen, Z. Liu, S. Wang, and Q. Wang, “Software testing with large language models: Survey, landscape, and vision,” *IEEE Transactions on Software Engineering*, vol. 50, no. 4, pp. 911 – 936, 2024.
- [3] Z. Zheng *et al.*, “Towards an understanding of large language models in software engineering tasks,” *Empirical Softw. Engg.*, vol. 30, no. 2, Dec. 2024, doi: 10.1007/s10664-024-10602-0.
- [4] V. V. Nguyen *et al.*, “Revolutionizing Education: An Extensive Analysis of Large Language Models Integration,” *IRJSTEM*, vol. 4, no. 4, pp. 10-21, 2024.
- [5] H. K. Nguyen, V. V. Nguyen, T. V. Nguyen, and H. C. Nguyen, “Phi-3 Meets Law: Fine-tuning Mini Language Models for Legal Document Understanding,” *Research, Development and Application on Information and Communication Technology*, vol. 2024, no. 3, pp. 136–142, 2024.

-
- [6] D. Song *et al.*, “LUNA: A Model-Based Universal Analysis Framework for Large Language Models,” *IEEE Transactions on Software Engineering*, vol. 50, no. 7, pp. 1921–1948, 2024, doi: 10.1109/TSE.2024.3411928.
- [7] S. S. Sohail *et al.*, “Decoding ChatGPT: A taxonomy of existing research, current challenges, and possible future directions,” *Journal of King Saud University-Computer and Information Sciences*, vol. 35, no. 8, 2023, Art. no. 101675.
- [8] N. Kiesler, D. Lohr, and H. Keuning, “Exploring the potential of large language models to generate formative programming feedback,” in *2023 IEEE Frontiers in Education Conference (FIE)*, IEEE, 2023, pp. 1–5.
- [9] Z. Fan, X. Gao, M. Mirchev, A. Roychoudhury, and S. H. Tan, “Automated repair of programs from large language models,” in *2023 IEEE/ACM 45th International Conference on Software Engineering (ICSE)*, IEEE, 2023, pp. 1469–1481.
- [10] J. Lu, L. Yu, X. Li, L. Yang, and C. Zuo, “LLaMA-Reviewer: Advancing Code Review Automation with Large Language Models through Parameter-Efficient Fine-Tuning,” in *2023 IEEE 34th International Symposium on Software Reliability Engineering (ISSRE)*, IEEE, 2023, pp. 647–658.
- [11] V. Siddeshwar, S. Alwidian, and M. Makrehchi, “A Systematic Review of AI-Enabled Frameworks in Requirements Elicitation,” *IEEE Access*, vol. 12, pp. 154310–154336, 2024.
- [12] J. Wei *et al.*, “Emergent abilities of large language models,” *arXiv preprint arXiv:2206.07682*, 2022.
- [13] J. Sallou, T. Durieux, and A. Panichella, “Breaking the silence: the threats of using llms in software engineering,” in *Proceedings of the 2024 ACM/IEEE 44th International Conference on Software Engineering: New Ideas and Emerging Results*, 2024, pp. 102–106.
- [14] G. Xiao, J. Lin, M. Seznec, H. Wu, J. Demouth, and S. Han, “Smoothquant: Accurate and efficient post-training quantization for large language models,” in *International Conference on Machine Learning*, PMLR, 2023, pp. 38087–38099.
- [15] Y. Liu *et al.*, “Summary of chatgpt-related research and perspective towards the future of large language models,” *Meta-Radiology*, vol. 1, no. 2, 2023, Art. no. 100017.
- [16] P. Vaithilingam, T. Zhang, and E. L. Glassman, “Expectation vs. experience: Evaluating the usability of code generation tools powered by large language models,” in *Chi Conference on Human Factors in Computing Systems Extended Abstracts*, 2022, pp. 1–7.
- [17] I. Singh *et al.*, “Progprompt: Generating situated robot task plans using large language models,” in *2023 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, 2023, pp. 11523–11530.
- [18] S. Makridakis, F. Petropoulos, and Y. Kang, “Large language models: Their success and impact,” *Forecasting*, vol. 5, no. 3, pp. 536–549, 2023.
- [19] Y. Fu *et al.*, “Gpt4aigchip: Towards next-generation ai accelerator design automation via large language models,” in *2023 IEEE/ACM International Conference on Computer Aided Design (ICCAD)*, IEEE, 2023, pp. 1–9.