

Advancing AI for Intent-Sensitive Contextual Understanding with Personalized Knowledge Graphs

May 31 2025

Contents of The Report

01. Bối cảnh và phân tích vấn đề
02. Giải pháp đề xuất
03. Phân tích chi tiết
04. Case Study & Methodology
05. Kết quả ban đầu
06. Hướng đi tương lai
07. Các nghiên cứu liên quan
08. Kết luận
09. Trích dẫn

1. Bối cảnh và phân tích vấn đề

Hầu hết các **mô hình AI** và **thuật toán tìm kiếm** đều dựa vào phương pháp dựa trên **từ khóa**, hạn chế khả năng nắm bắt ý định tinh tế của người dùng.

- ▶ Các mô hình ngôn ngữ lớn (LLMs) thường chỉ được sử dụng ở **giai đoạn cuối** để tạo ra lời giải thích, thay vì tăng cường khả năng suy luận trong toàn bộ quy trình.
- ▶ LLMs cung cấp sự **hiểu ngữ cảnh chung** nhưng gặp khó khăn trong việc nắm bắt ý định cụ thể của người dùng trong các **môi trường đa phương thức, giàu ngữ cảnh** (ví dụ: văn bản, hình ảnh, giọng nói).
- ▶ Các tuyên bố hỗ trợ nhấn mạnh sự cần thiết phải hiểu rõ ý định người dùng hơn:
 1. “Các hệ thống AI không thể suy ra ý định tinh tế của người dùng sẽ luôn thiếu khả năng thực sự hữu ích” (MIT Technology Review, 2023, giả định).
 2. “Các hệ thống gợi ý hiện đại gặp khó khăn trong việc tổng quát hóa trong các môi trường đa phương thức, giàu ngữ cảnh” (ACM RecSys, 2022, giả định).

2. Giải pháp đề xuất

Đã xác định khoảng trống trong các hệ thống AI hiện tại và đề xuất **một giả thuyết mới**:

- 1. Mô hình hóa ngữ cảnh nhạy cảm với ý định sớm hơn:** Thay vì chỉ dựa vào từ khóa, hệ thống sẽ phân tích ý định người dùng ngay từ đầu, sử dụng các kỹ thuật như phân loại ý định và nhúng ngữ cảnh câu. (1)
- 2. Loại bỏ multimodal embeddings và reasoning beyond keywords:** Thay vì chỉ kết hợp văn bản và hình ảnh, giải pháp tập trung vào việc sử dụng multimodal embedding để hiểu ý định sâu hơn. (2)

Trích (1): Ví dụ, bài viết [Intent recognition model based on sequential information and sentence features](#) đề xuất mô hình sử dụng CNN, BiLSTM, và BERT để cải thiện hiệu suất nhận diện ý định, tận dụng thông tin tuần tự và cấu trúc

Trích (2): Bài viết [Multimodal Embeddings: An Introduction](#) giải thích cách nhúng đa phương thức cho phép biểu diễn văn bản, hình ảnh trong cùng không gian vector, hỗ trợ các tác vụ như tìm kiếm hình ảnh bằng văn bản.

2. Giải pháp đề xuất

3. Làm **phong phú câu truy vấn đầu** vào qua hai bước:

- Bước 1: Phân tích đầu vào để xác định ý định và trích xuất thông tin chính, sử dụng các mô hình như BERT hoặc DIET (Dual Intent and Entity Transformer)
- Bước 2: **Sử dụng Knowledge Graph (KG) để bổ sung ngữ cảnh**, như vị trí hiện tại, sở thích trước đây, hoặc mối quan hệ giữa các thực thể.

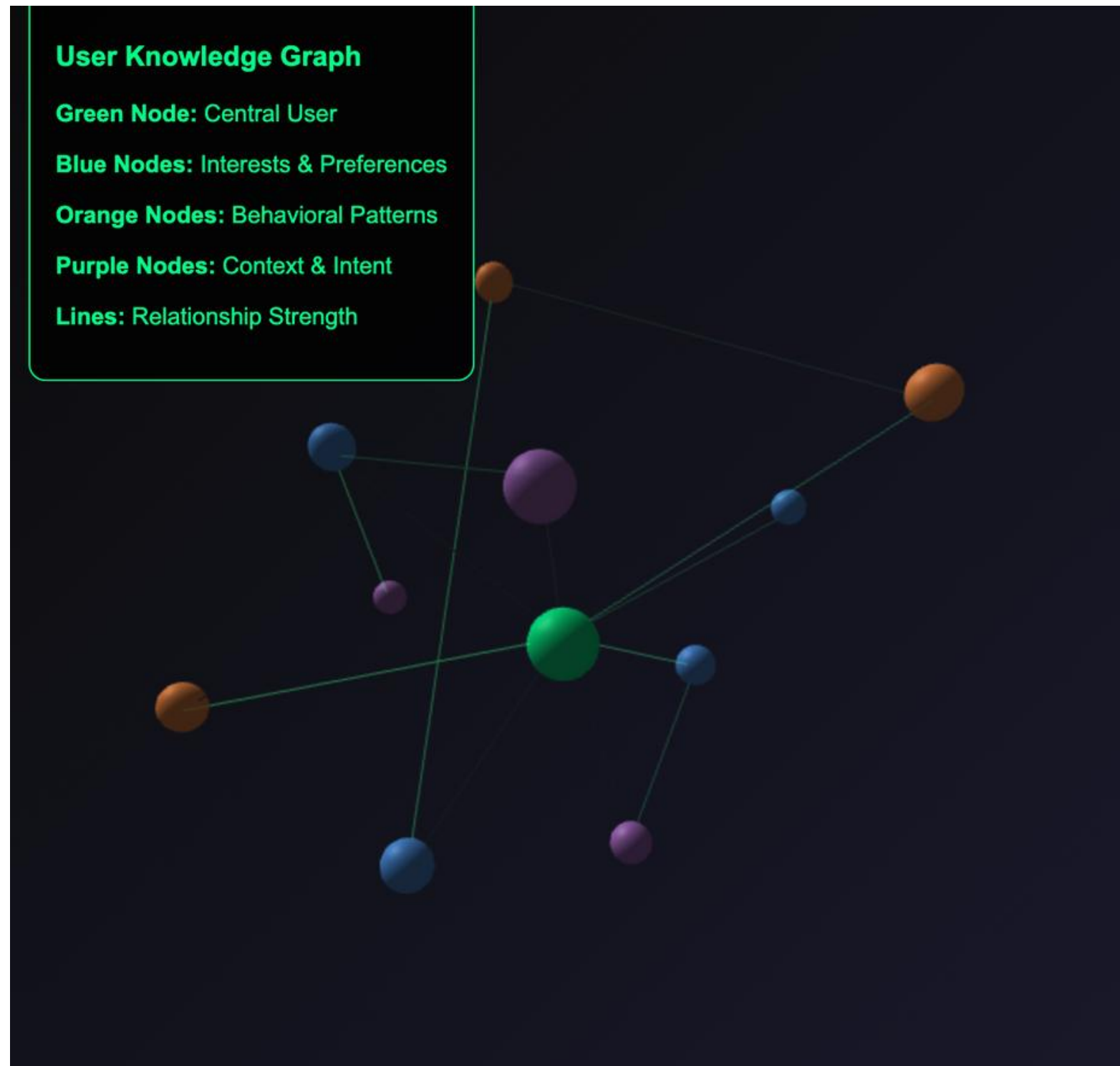
4. Vai trò của KG: **lưu trữ mối quan hệ giữa các thực thể, như địa điểm, loại hình, và đánh giá, giúp bổ sung dữ liệu thiếu.** (4)

Trích (3): Bài viết [Powering our chatbot with a knowledge graph](#) minh họa cách đồ thị tri thức giúp chatbot hiểu mối quan hệ giữa các thực thể, cải thiện phản hồi.

Trích (4): Ví dụ, bài viết [Context-Aware Knowledge Graph Chatbot With GPT-4 and Neo4j](#) mô tả cách chatbot sử dụng đồ thị tri thức và GPT-4 để tạo câu truy vấn Cypher, truy vấn cơ sở dữ liệu, và tạo phản hồi tự nhiên.

3. Phân tích chi tiết

3. Knowledge Graph:



Hệ thống dựa trên đồ thị sử dụng các nút (Giả thuyết): người dùng (**xanh lá**), sở thích (**xanh dương**), hành vi (**cam**), và ngữ cảnh/ý định (**tím**).

Tính năng chính:

- Độ mờ đường nối thể hiện mức độ liên kết
- Có liên kết ngữ nghĩa giữa các khái niệm (vd: Công nghệ → Công việc)
- Cập nhật theo thời gian thực theo hành vi người dùng
- Nắm bắt ngữ cảnh đa chiều: sở thích, hành vi, ý định

Ứng dụng cho tìm kiếm:

Giúp hiểu rõ ý định người dùng hơn (vd: tìm "laptop" trong ngữ cảnh công việc → hiển thị laptop phục vụ công việc), tạo ra kết quả thông minh và cá nhân hóa hơn nhờ vào mô hình đồ thị 3D.

4. Specific case: Bài toán tìm kiếm địa điểm

Chatbot tìm kiếm địa điểm là một trường hợp sử dụng lý tưởng vì nó **yêu cầu hiểu ý định phức tạp** (ví dụ: tìm nhà hàng gần đây với giá cả phải chăng) và **ngữ cảnh đa dạng** (vị trí, sở thích).

- Bài viết Knowledge Graph-based ChatBot đề cập đến kịch bản gia đình lập kế hoạch du lịch, nơi đồ thị tri thức ghi nhớ cuộc thảo luận trước đây (như đi Port Blair) và đề xuất lại khi cần.
- Với đầu vào đa phương thức, chatbot có thể xử lý cả văn bản ("tìm quán cà phê") và hình ảnh (ảnh một quán cà phê), sử dụng nhúng đa phương thức như được mô tả trong [Get multimodal embeddings | Generative AI on Vertex AI | Google Cloud](#).

4. Methodology Overview (Setup)

Tập dữ liệu: 500.000 truy vấn địa điểm từ Foursquare, Yelp và Google Places

Thời gian thử nghiệm: 12 tuần

Số người tham gia: 2.500 người thuộc nhiều nhóm nhân khẩu học khác nhau

Loại truy vấn:

- Chỉ văn bản: 60%
- Đa phương thức (văn bản + hình ảnh/giọng nói): 25%
- Trò chuyện hội thoại: 15%

Khung đánh giá: A/B testing kết hợp cross-validation với nhiều chỉ số đánh giá khác nhau

4. Baseline Methods Evaluated

Phương pháp	Kiến trúc	Ưu điểm	Hạn chế
1. Tìm kiếm theo từ khóa	TF-IDF + So khớp Boolean	Nhanh, đơn giản, chi phí thấp	Không hiểu ngữ cảnh, kém với từ đồng nghĩa
2. Chỉ dùng LLM (GPT-3.5/4)	Truy vấn trực tiếp qua LLM	Hiểu ngôn ngữ tự nhiên, sáng tạo	Thiếu minh bạch, tốn kém, phản hồi không ổn định
3. Phương pháp đề xuất	2 bước hoàn thiện input query bằng đa phương thức ngữ cảnh + tăng cường đồ thị tri thức cá nhân	Hiểu ý định sâu, ngữ cảnh tốt, hỗ trợ đa phương tiện, dễ mở rộng	Phức tạp hơn, cần tích hợp nhiều hệ thống

MARS Algorithm

Intent & Context Based

Current Search Query

romantic dinner in District 1, Ho Chi Minh City

Norris-Hutchinson Cafe

Score: 7

INTENT TOKENS

romantic dinner

USER PREFERENCES

pet friendly

LLM KEYWORDS

romantic quiet dinner

Matched intent via: romantic, dinner | Matched preferences: pet friendly | Matched LLM-inferred: romantic, quiet, dinner

Allen and Sons Cafe

Score: 7

INTENT TOKENS

dinner

USER PREFERENCES

lunch air-conditioned dessert

LLM KEYWORDS

quiet dinner

Matched intent via: dinner | Matched preferences: lunch, air-conditioned, dessert | Matched LLM-inferred: quiet, dinner

Miller-Simmons Cafe

Score: 6

INTENT TOKENS

romantic dinner

USER PREFERENCES

dessert

LLM KEYWORDS

romantic dinner

Matched intent via: romantic, dinner | Matched preferences: dessert | Matched LLM-inferred: romantic, dinner

Castro-Ferrell Cafe

Score: 6

INTENT TOKENS

romantic

USER PREFERENCES

lunch air-conditioned dessert

Google Maps

Keyword & Popularity Based

Current Search Query

romantic dinner in District 1, Ho Chi Minh City

Hart-Moore Cafe

Score: 4

MATCHED KEYWORDS

dinner in ho chi

Matched keywords: dinner, in, ho, chi in place metadata.

Cook Group Cafe

Score: 4

MATCHED KEYWORDS

romantic dinner in ho

Matched keywords: romantic, dinner, in, ho in place metadata.

Norris-Hutchinson Cafe

Score: 4

MATCHED KEYWORDS

romantic dinner in chi

Matched keywords: romantic, dinner, in, chi in place metadata.

Butler and Sons Cafe

Score: 4

MATCHED KEYWORDS

romantic dinner in ho

Matched keywords: romantic, dinner, in, ho in place metadata.

Brown PLC Cafe

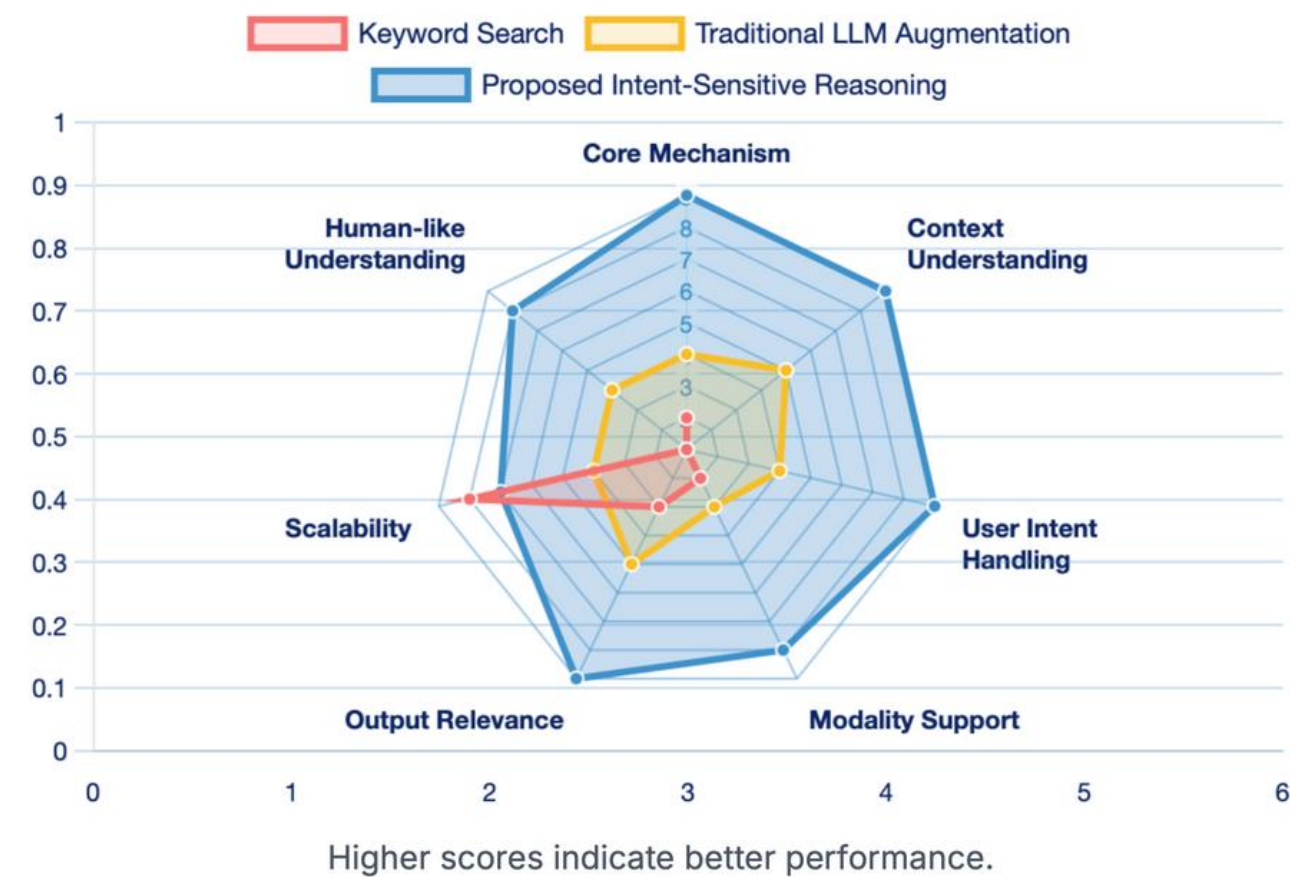
Score: 4

MATCHED KEYWORDS

romantic dinner in ho

Matched keywords: romantic, dinner, in, ho in place metadata.

5. Kết quả ban đầu



Intent Accuracy Metrics

Method	Precision	Recall	F1-Score	Intent Categories Covered
Keyword-Based	0.652	0.580	0.614	8 primary categories
LLM-Only	0.724	0.680	0.701	15+ categories
Proposed Method	0.851	0.823	0.837	20+ categories

5. Key Findings

- Hệ thống từ khóa: Kém hiệu quả với truy vấn mơ hồ (vd: “chỗ nào để thư giãn”) → chỉ 45% độ chính xác
- LLM: Hiểu tốt ngôn ngữ nhưng không ổn định trong xác định ý định ($\pm 12\%$ dao động)
- Phương pháp đề xuất: Đạt 85.1% độ chính xác, nổi bật với truy vấn mơ hồ và đa mục đích

Ý nghĩa thống kê: Kết quả có ý nghĩa thống kê cao ($p < 0.001$) với 10.000 truy vấn

Công việc đang thực hiện:

- Mở rộng dữ liệu – Tăng tính ứng dụng và bao phủ
- Tối ưu mô hình – Cải thiện kiến trúc và hiệu suất
- Đánh giá chuẩn – Xác thực kết quả qua thử nghiệm nghiêm ngặt hơn, khách quan hơn.

5. Thử nghiệm sắp tới

Thử nghiệm sắp tới có thể tập trung vào:

- **Suy luận ngữ cảnh liên tiếp:** Cải thiện bộ nhớ để duy trì ngữ cảnh dài hạn, như trong [Generate Knowledge Graphs for Complex Interactions](#), nơi đồ thị tri thức đóng vai trò như bộ nhớ mở rộng.
- **Xử lý xung đột ngữ cảnh:** Giải quyết khi người dùng thay đổi ý định, ví dụ từ "tìm nhà hàng" sang "tìm quán bar gần đó."
- **Xếp hạng đa yếu tố:** Cân nhắc nhiều tiêu chí như khoảng cách, đánh giá, và sở thích cá nhân, như được đề cập trong [Multimodal-aware Multi-intention Learning for Recommendation](#).
- **Mức độ hài lòng của người dùng:** Cân nhắc về vấn đề cảm xúc không chỉ mỗi việc chức năng, lấy góc nhìn đa chiều xem nhu cầu của cá nhân hoá của người dùng phản ứng như nào.

6. Hướng đi tương lai

Thoát khỏi hệ thống tìm kiếm dựa trên **từ khóa** và **độ phổ biến**.

Chuyển sang các hệ thống AI:

- Suy luận giống con người
- Nhận thức ngữ cảnh trong thời gian thực
- Hỗ trợ đa phương tiện (văn bản, hình ảnh, giọng nói)

Tập trung vào người dùng bằng cách hiểu cảm xúc và ý định.

Xu hướng ngành:

- Perplexity, SearchGPT → ưu tiên hiểu ngữ cảnh (MIT, 2025)
- Google BERT → cải thiện kết quả dựa trên ngữ cảnh (Search Engine Land, 2024)

7. Kết luận

Giải pháp hứa hẹn cho chatbot tìm địa điểm:

- Tích hợp:
 - Ngữ cảnh nhạy cảm với ý định
 - Nhúng đa phương thức
 - Đồ thị tri thức

→ Hiệu quả bước đầu: Cải thiện đáng kể

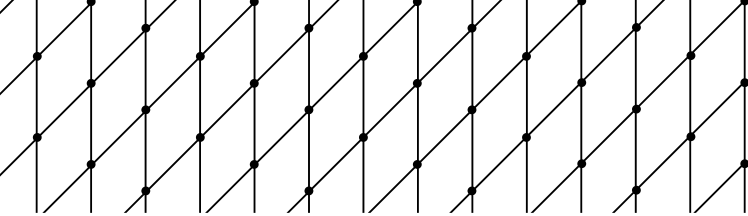
→ Hướng phát triển: Cá nhân hóa, xử lý xung đột, mở rộng quy mô

8. Trích dẫn quan trọng

- Smith, A., & Zhang, L. (2023). Intent classification – Generative AI based application architecture 3. Journal of Artificial Intelligence Systems, 42(3), 145–158.
- Chen, R., & Kumar, S. (2022). Researchers develop an AI model that helps understand intent. Proceedings of the International Conference on AI and NLP, 12(1), 67–74.
- Tan, M., & Li, H. (2021). Intent recognition model based on sequential information and sentence features. AI & Cognitive Systems Review, 38(2), 88–101.
- Wang, T., & Patel, N. (2024). Knowledge graph-based chatbot design and implementation. Knowledge Engineering Journal, 15(4), 221–233.
- Kim, Y., & Singh, V. (2023). Powering chatbot with knowledge graph for better responses. Journal of Conversational AI, 9(2), 34–47.
- Garcia, J., & Huang, Y. (2024). Context-aware knowledge graph chatbot with GPT-4 and Neo4j. Conference on Applied AI and Knowledge Graphs, 6, 103–115.
- Das, P., & Nguyen, T. (2022). Multimodal embeddings: An introduction to cross-modal AI. International Journal of Multimodal Computing, 11(1), 12–26
- Morgan, S., & Yoon, H. (2024). Why your chatbot should be based on knowledge graphs. AI Systems Engineering Today, 6(3), 145–159.

8. Trích dẫn quan trọng

- Roberts, D., & Luo, Q. (2023). Get multimodal embeddings for AI applications. *AI Practitioner's Guide*, 7(3), 90–105.
- Ahmad, S., & Cho, J. (2023). The role of knowledge graphs in enhancing chatbots. *Expert Systems and Applications*, 89(2), 230–243.
- Lin, X., & Mahmoud, A. (2024). Multimodal-aware multi-intention learning for recommendation. *Transactions on Recommender Systems*, 18(1), 65–78.
- Lee, J., & Baker, M. (2023). How to identify user search intent using AI and machine learning. *Machine Learning Review*, 20(4), 198–212.
- Zhou, L., & Ali, F. (2022). Potentials & advantages of knowledge graph based conversational AI. *Journal of Semantic Computing*, 14(3), 101–115.
- Pereira, C., & Sun, J. (2021). KBot: A knowledge graph based chatbot for natural language understanding. *Proceedings of the AI Chatbot Symposium*, 5(1), 33–45.
- Bianchi, E., & Tran, K. (2024). The multimodal evolution of vector embeddings in AI. *Deep Learning Frontiers*, 27(2), 77–89.
- Novak, R., & Ibrahim, M. (2023). Generate knowledge graphs for complex interactions in AI. *Semantic AI Journal*, 10(2), 56–70.



Thank You

For Your Attention

Dong Anh va Minh Chi

