

## Bài báo nghiên cứu

**TÍCH HỢP TRI THỨC DẠNG ONTOLOGY  
VÀ ĐỒ THỊ TRI THỨC CHO HỆ THỐNG CHATBOT  
HỖ TRỢ TRUY VẤN KIẾN THỨC TRONG GIÁO DỤC**

*Nguyễn Việt Hưng<sup>1</sup>, Lê Thị Ngọc Thảo<sup>1</sup>, Nguyễn Văn Hậu<sup>1</sup>,  
Nguyễn Đắc Long<sup>1</sup>, Trần Phong Nhã<sup>2</sup>, Nguyễn Đình Hiến<sup>3\*</sup>*

<sup>1</sup>Trường Đại học Sư phạm Thành phố Hồ Chí Minh, Việt Nam

<sup>2</sup>Phân hiệu tại Thành phố Hồ Chí Minh, Trường Đại học Giao thông Vận tải, Việt Nam

<sup>3</sup>Trường Đại học Công nghệ Thông tin, Đại học Quốc gia Thành phố Hồ Chí Minh, Việt Nam

\*Tác giả liên hệ: Nguyễn Đình Hiến – Email: [hiennnd@uit.edu.vn](mailto:hiennnd@uit.edu.vn)

Ngày nhận bài: 10-02-2024; ngày nhận bài sửa: 27-3-2024; ngày duyệt đăng: 29-3-2024

**TÓM TẮT**

Ngày nay, việc học trực tuyến ngày càng phổ biến và đa dạng. Người học có thể thông qua các thiết bị điện tử và các nền tảng internet để học tập, tìm kiếm tài liệu, tra cứu kiến thức. Trong nghiên cứu này, chúng tôi đề xuất một mô hình tích hợp ontology tri thức quan hệ – toán tử và đồ thị tri thức. Mô hình có thể tối ưu khả năng biểu diễn kiến thức các môn học kết hợp với đặc tả các quan hệ giữa các thành phần tri thức. Trên cơ sở tri thức được xây dựng, các vấn đề truy vấn kiến thức được giải quyết để đáp ứng các yêu cầu cho việc xây dựng hệ thống hỗ trợ học tập thông minh. Giải pháp đề xuất được ứng dụng để xây dựng hệ thống hỗ trợ truy vấn kiến thức môn học Cơ sở dữ liệu dưới dạng chatbot. Hệ thống có thể hỗ trợ việc truy vấn nội dung kiến thức môn học, theo phân loại kiến thức và theo các dạng bài tập trong môn học. Hệ thống xây dựng cũng được đánh giá và so sánh với các mô hình ngôn ngữ lớn hiện nay trên phương diện hỗ trợ học tập môn học Cơ sở dữ liệu.

**Từ khóa:** chatbot; cơ sở dữ liệu; hệ thống thông minh; các hệ cơ sở tri thức; công nghệ tri thức

**1. Giới thiệu**

Trong nền công nghiệp 4.0, lĩnh vực công nghệ thông tin đóng một vai trò quan trọng (Elayyan, 2021), do đó, việc đào tạo nhân lực tiếp cận với công nghệ thông tin là nhu cầu rất cấp thiết. Các hình thức học tập và giảng dạy khác nhau được áp dụng để phục vụ cho việc học, trong đó, học tập trực tuyến là hình thức học có nhiều ưu điểm (Akhter et al., 2021). Với hình thức học này, người học có thể tự học, tự tìm hiểu kiến thức trên các nền tảng trực tuyến mà không cần gặp trực tiếp giảng viên của mình, từ đó nâng cao khả năng tự học. Tuy nhiên, để có thể tự học một cách hiệu quả thì cần phải có một môi trường tra cứu kiến thức chuẩn khoa học và có độ chính xác cao.

*Cite this article as:* Nguyen Viet Hung, Le Thi Ngoc Thao, Nguyen Van Hau, Nguyen Dac Long, Tran Phong Nha, & Nguyen Dinh Hien (2024). Integrating knowledge based on ontology and knowledge graph for chatbot in education query. *Ho Chi Minh City University of Education Journal of Science*, 21(5), 771-784.

Trong chương trình đào tạo ngành Công nghệ Thông tin, môn Cơ sở dữ liệu là một môn học quan trọng (Foster & Godbole, 2022). Môn học cung cấp cho người học các kiến thức cơ bản về cơ sở dữ liệu (CSDL), hệ quản trị CSDL và ngôn ngữ truy vấn dữ liệu. Qua môn học, người học nắm và biết cách sử dụng thành thạo các lệnh truy vấn dữ liệu, nắm vững việc sử dụng hệ quản trị cơ sở dữ liệu để thực hiện các thao tác thiết kế và quản lý CSDL, hoặc phát triển ứng dụng có liên quan.

Chatbot được xem là những thiết bị hay phần mềm có giao diện người dùng hỗ trợ giao tiếp qua ngôn ngữ tự nhiên (Dale, 2016; Følstad & Brandtzaeg, 2017). Bên cạnh đó, với sự phát triển của các mô hình ngôn ngữ lớn hiện nay (Large Language Models – LLMs), các ứng dụng chatbot ngày càng trở nên phổ biến (Agrawal et al., 2022). Chatbot có thể ứng dụng trong việc xây dựng các hệ thống hỗ trợ trong học tập (Davies, 2020). Tuy nhiên, các hệ thống chưa được tích hợp một cơ sở tri thức đầy đủ, đặc biệt là kiến thức của các môn học, và chưa được tích hợp khả năng tương tác với người học mang tính sư phạm trong việc hướng dẫn học tập. Do đó, hệ thống chưa đáp ứng các yêu cầu của một hệ hỗ trợ học tập. Các nghiên cứu đã đề xuất giải pháp cho việc xây dựng cơ sở tri thức cho chatbot trong hệ thống học tập. Tuy nhiên, các phương pháp này chưa thiết lập được việc biểu diễn các quan hệ giữa các đối tượng trong miền kiến thức một cách toàn vẹn.

Đồ thị tri thức sử dụng mô hình dữ liệu dựa trên đồ thị để nắm bắt kiến thức trong các kịch bản ứng dụng liên quan đến việc tích hợp, quản lý và trích xuất giá trị từ nhiều nguồn dữ liệu khác nhau ở quy mô lớn (Zhang & Li, 2024; Noy et al., 2019). Đồ thị cung cấp sự trừu tượng ngắn gọn và trực quan cho cho nhiều lĩnh vực khác nhau, trong đó các cạnh và đường dẫn nắm bắt các mối quan hệ khác nhau, có khả năng phức tạp giữa các thực thể của một miền (Angles & Gutiérrez, 2008). Vì vậy, tích hợp đồ thị tri thức và các mô hình tri thức quan hệ giúp cho việc biểu diễn các dạng tri thức quan hệ trở nên chính xác và toàn vẹn hơn.

Trong bài báo này, chúng tôi đề xuất một phương pháp biểu diễn kiến thức môn học dựa trên sự kết hợp mô hình tri thức quan hệ – toán tử (Nguyen et al., 2020) và đồ thị tri thức (Abu-Salih & Alotaibi, 2024). Phương pháp có thể áp dụng để tổ chức cơ sở tri thức cho môn học Cơ sở dữ liệu và hỗ trợ việc thiết kế hệ thống tra cứu kiến thức dạng chatbot. Hệ thống có thể hỗ trợ trong việc tra cứu nội dung theo bài học, theo phân loại kiến và theo các dạng bài tập trong môn học. Chatbot được xây dựng có thể hỗ trợ cho sinh viên học môn học Cơ sở dữ liệu thuộc chương trình đào tạo ngành Công nghệ Thông tin. Hệ thống xây dựng cũng được đánh giá và so sánh với các mô hình LLMs hiện nay, như ChatGPT, Google Gemini và Microsoft Copilot, trong việc hỗ trợ học tập với môn học Cơ sở dữ liệu.

## **2. Nội dung nghiên cứu**

### **2.1. Công trình nghiên cứu liên quan**

Hiện nay có nhiều phương pháp thiết kế cho các hệ thống tìm kiếm thông tin (Do et al., 2020; Liu et al., 2018; Hagedorn et al., 2020). Ontology là giải pháp hữu hiệu để biểu diễn các miền tri thức quan hệ (Nguyen et al., 2022). Ontology được sử dụng trong việc biểu diễn các miền tri thức khác nhau như trong giáo dục (Nguyen et al., 2021), các văn bản luật (Ngo et al.,

2024). Các giải pháp ontology sẽ được tích hợp các phương pháp tìm kiếm để xây dựng hệ thống tìm kiếm, tra cứu kiến thức theo ngữ nghĩa (Nguyen et al., 2021; Truong et al., 2022).

Tri thức về các quan hệ và toán tử có thể được biểu diễn bởi ontology Rela-Ops (Nguyen et al., 2022). Mô hình này là sự kết hợp các thành phần tri thức như khái niệm, quan hệ và các toán tử giữa các khái niệm, luật suy diễn của miền tri thức. Mô hình Rela-Ops model cũng được vận dụng để xây dựng các hệ thống hướng dẫn giải bài tập thông minh trong giáo dục (Nguyen et al., 2020). Ontology Rela-Ops model cũng được nghiên cứu kết hợp thành phần tri thức hàm để tổ chức cơ sở tri thức cho các môn học trong việc truy vấn kiến thức (Truong et al., 2022). Các hệ thống truy vấn cũng được thiết kế trên cơ sở biểu diễn mối quan hệ thông qua đồ thị tri thức (Nguyen et al., 2023). Tuy nhiên, trong các hệ thống này, sự liên hệ giữa cấu trúc đồ thị tri thức và mô hình Rela-Ops chưa được biểu diễn một cách đầy đủ. Mối liên hệ chỉ được thực hiện thông qua các loại quan hệ nội tại của miền tri thức.

Phương pháp tìm kiếm theo chỉ mục (index-based search) là một kỹ thuật giúp nhanh chóng tìm thấy tất cả các kết quả thích hợp từ cơ sở dữ liệu theo các từ khóa đã nhập (Guemmat & Ouahabi, 2018). Tuy nhiên, các phương pháp này không thể hỗ trợ việc tìm kiếm theo ngữ nghĩa. Vì vậy, để đáp ứng nhu cầu tìm kiếm theo ngữ nghĩa, cũng như tra cứu kiến thức trong miền tri thức, các giải pháp sử dụng cơ sở tri thức được nghiên cứu và xây dựng (Yoo & Jeong, 2020). Giải pháp này giúp cho người dùng có thể truy vấn thông tin cần thiết một cách chính xác hơn.

Nghiên cứu này đề xuất việc tích hợp đồ thị tri thức vào một mô hình chatbot để xây dựng một hệ thống trả lời kiến thức môn Cơ sở dữ liệu thông minh có thể tương tác với người dùng. Mô hình tri thức được xây dựng sử dụng ontology Rela-Ops model để tổ chức cơ sở tri thức của môn học CSDL và tích hợp đồ thị tri thức để biểu diễn các kiến thức đó trong hệ thống tra cứu kiến thức. Trên cơ sở tri thức này, hệ thống được thiết kế có thể hỗ trợ sinh viên tốt hơn trong việc tìm kiếm nội dung, thông qua việc tra cứu kiến thức theo bài, tra cứu theo các loại của thành phần tri thức với các dạng bài tập tương ứng, và gợi ý các kiến thức liên quan. Hệ thống có thể tương tác với người dùng để hướng dẫn việc học các kiến thức mà người học đang quan tâm.

## 2.2. Mô hình quan hệ biểu diễn kiến thức môn học Cơ sở dữ liệu dựa trên đồ thị tri thức

### 2.2.1. Mô hình tri thức quan hệ và toán tử

Ontology Rela-Ops model là mô hình biểu diễn kiến thức kết hợp giữa các quan hệ và toán tử được xây dựng theo tiếp cận hướng đối tượng, trong đó mỗi khái niệm là một lớp các đối tượng. Các thành phần của mô hình là các tập hợp có cấu trúc và các thuộc tính tương ứng.

**Định nghĩa 3.1.** (Nguyen et al., 2020): Ontology tri thức quan hệ và toán tử, *Rela - Ops model*, là một bộ gồm 4 thành phần:

#### (C, R, Ops, Rules)

trong đó, **C** là tập hợp các khái niệm, mỗi khái niệm là một lớp đối tượng. Mỗi khái niệm  $c \in \mathbf{C}$  có cấu trúc (*Name, Content, Attrs, Rel\_In, Keys*), với *Name* là tên của khái niệm, *Content* là diễn giải nội dung của khái niệm, *Attrs* là tập các thuộc tính của khái niệm  $c$ , *Rel\_In* là các

quan hệ nội tại giữa các thuộc tính trong  $c$ , và  $Keys$  là các từ khóa liên quan đến khái niệm.  $\mathbf{R}$  là tập hợp các quan hệ giữa các khái niệm, đối tượng và thuộc tính, các quan hệ này bao gồm các quan hệ đặc biệt, “*is-a*” và “*has-a*”, và các quan hệ hai ngôi khác.  $\mathbf{Ops}$  là một tập hợp các toán tử.  $\mathbf{Rules}$  là các luật suy diễn của miền tri thức. Các luật này là các tính chất, định lí, hay công thức trong miền tri thức.

2.2.2. Cấu trúc đồ thị tri thức cho mô hình *Rela – Ops*

Đồ thị tri thức là tập hợp các mô tả được liên kết với nhau về các khái niệm, thực thể và mối quan hệ trong một miền tri thức (Hogan et al., 2021). Cơ sở tri thức được biểu diễn dưới dạng cấu trúc đồ thị và gồm hai thành phần chính là tập đỉnh và tập cạnh.

**Định nghĩa 3.2.** Đồ thị tri thức cho miền tri thức  $\mathcal{K} = (\mathbf{C}, \mathbf{R}, \mathbf{Ops}, \mathbf{Rules})$  là một tập hợp:

$$KG = (V, E)$$

trong đó:

+  $V$  là tập các đỉnh của đồ thị biểu diễn tri thức cho miền tri thức  $\mathcal{K}$  bao gồm khái niệm, đối tượng của miền tri thức và mối quan hệ giữa các thuộc tính bên trong.

+  $E$  là tập các cạnh biểu diễn mối quan hệ một chiều giữa các đỉnh trong  $V$ . Tập  $E$  được chia làm hai loại như sau:  $E = E_{\text{internal}} \cup E_{\text{external}}$ , với:

$$E_{\text{internal}} = \bigcup_{c \in \mathbf{C}} E^c \text{ với } E^c = \{r(u, v) \mid u, v \in c.Attrs, r \in c.Rel\_In\}$$

$$E_{\text{external}} = \{rel(w, t) \mid w, t \in \mathcal{K}.C, rel \in \mathcal{K}.R\}$$

$E_{\text{internal}}$  là các quan hệ nội tại giữa các thuộc tính trong cùng một khái niệm, và  $E_{\text{external}}$  là quan hệ giữa các khái niệm và thuộc tính của miền tri thức. Ví dụ: Tri thức  $A$  liên quan với tri thức  $B$  thì ta có:  $v(A) - (related\_to) \rightarrow v(B)$ .

Hình 1 là một phần đồ thị tri thức được xây dựng cho kiến thức môn học CSDL được tổ chức theo mô hình *Rela – Ops*.



**Hình 1.** Đồ thị tri thức biểu diễn kiến thức môn học CSDL được đặc tả theo ontology *Rela – Ops model*

2.2.3. Tích hợp cấu trúc của đồ thị tri thức và mô hình *Rela – Ops*

Mô hình tri thức quan hệ toán tử có thể biểu diễn đầy đủ các kiến thức trong miền tri thức. Đồ thị tri thức có ưu điểm trong việc thể hiện các quan hệ giữa các thành phần trong

miền tri thức. Trong mục này, chúng tôi sẽ trình bày mô hình tích hợp để tổ chức cơ sở tri thức theo ontology Rela-Ops model kết hợp với cấu trúc đồ thị tri thức. Mô hình tích hợp này có thể tối ưu việc biểu diễn tri thức tổng quát của mô hình tri thức Rela-Ops model cùng với khả năng biểu diễn các quan hệ của đồ thị tri thức để thực hiện các truy vấn kiến thức. Đồng thời, mô hình tích hợp này cũng hỗ trợ trong việc thêm mới và cập nhật kiến thức.

**Định nghĩa 3.3.** Mô hình tích hợp ontology Rela – Ops model và cấu trúc đồ thị tri thức, gọi là *KG Rel-Ops model*, gồm các thành phần như sau:

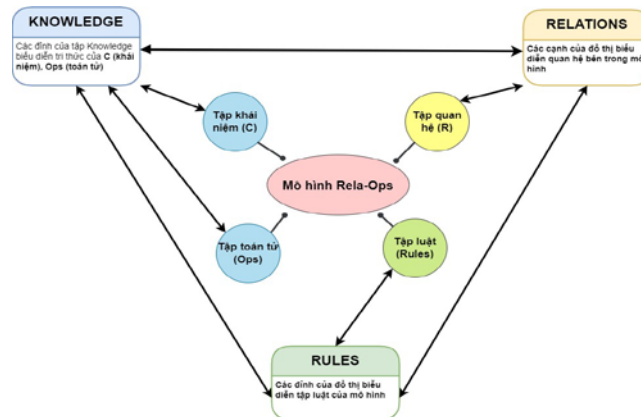
$$\mathcal{K} = (\mathbf{C}, \mathbf{R}, \mathbf{Ops}, \mathbf{Rules}) \oplus \mathbf{K}_{KG}$$

trong đó,  $(\mathbf{C}, \mathbf{R}, \mathbf{Ops}, \mathbf{Rules})$  là cấu trúc ontology Rela-Ops model như trong Định nghĩa 3.1,  $\mathbf{K}_{KG}$  là đồ thị tri thức như Định nghĩa 3.2 với mỗi đỉnh trong đồ thị  $\mathbf{K}_{KG}$  được chia làm hai loại: *Knowledge* và *Rules*.

+ *Knowledge*: là các đỉnh lưu trữ tri thức ở dạng ngôn ngữ tự nhiên. Các tri thức này sẽ là kết quả hiển thị cho người dùng.

+ *Rules*: là các đỉnh lưu trữ tập tri thức gồm có luật, quan hệ và hàm bên trong miền tri thức.

Tập cạnh trong đồ thị  $\mathbf{K}_{KG}$  biểu diễn các quan hệ giữa các đỉnh. Mỗi cạnh là một bộ gồm (*Attributes, Prop, Start, End*), trong đó, *Attributes* là tập các thuộc tính của quan hệ, *Props* là các tính chất của quan hệ, *Start* là đỉnh bắt đầu của quan hệ, và *End* là đỉnh kết thúc của quan hệ.



**Hình 2.** Cấu trúc của mô hình tri thức tích hợp KG Rel-Ops model

#### 2.2.4. Tổ chức cơ sở tri thức cho kiến thức môn Cơ sở dữ liệu

Kiến thức môn Cơ sở dữ liệu được thu thập từ (Silberschatz et al., 2019). Tri thức này có thể được tổ chức bởi mô hình KG Rel-Ops model. Cơ sở tri thức được xây dựng để thiết kế hệ thống Chatbot truy vấn kiến thức thông minh cho môn học CSDL. Trong đó, các thành phần được tổ chức như sau:

- **C**: Mỗi khái niệm  $c \in \mathbf{C}$  có cấu trúc  $c = (Name, Content, Attrs, Keys)$ .

Ví dụ 3.1. Khái niệm ENTITY có cấu trúc như sau:

ENTITY: {

*Name* = “thực\_thể”,

*Content* = “Thực thể là bất cứ các đối tượng, sự vật hay sự việc. Một thực thể có thể là địa điểm, người, đối tượng, sự kiện hoặc một khái niệm, lưu trữ dữ liệu trong cơ sở dữ liệu. Đặc điểm của các thực thể là phải có một thuộc tính và một khóa duy nhất.”,

```

Attrs = {
  attribute: tính chất của thực thể,
  Chapter: “Chương 2”
},
Rel_In = {};
Keys = [“thực thể”, “entity”, “entities”]
}
    
```

- **R**: là tập quan hệ giữa các khái niệm trong **C**.

Ví dụ 3.2. Quan hệ “belong\_to”: “thuộc\_tính\_kết\_hợp” belong\_to “thuộc\_tính”.

Quan hệ “related\_to” biểu diễn sự liên quan lẫn nhau giữa hai khái niệm: “thực\_thể” related\_to “thuộc\_tính”.

- **Ops** = {phép chọn, phép giao, phép hợp, phép tích Descartes, ...}

• **Rules**: Với mỗi luật  $r \in Rules$  là một luật suy diễn, có dạng:  $u(r) \rightarrow v(r)$ , với  $u(r)$ ,  $v(r)$  là các tập sự kiện.

Ví dụ 3.3.  $\forall c_1, c_2 \in C$ :

+ rule1:  $\{c_1 \text{ related\_to } c_2\} \rightarrow \{c_2 \text{ related\_to } c_1\}$

+ rule2:  $\{c_1 \text{ belong\_to } c_2\} \rightarrow \{c_2.Attrs \subseteq c_1.Attrs\}$

### 2.3. Các vấn đề và giải thuật

Truy vấn kiến thức dựa trên câu truy vấn đầu vào của người dùng là yếu tố quyết định của một chatbot thông minh. Với câu truy vấn đầu vào nhận được, hệ thống tiến hành trích xuất từ khóa chính. Từ các cụm từ khóa có được, cơ sở tri thức của hệ thống tiến hành phân tích và truy xuất nhằm tìm kiếm mối quan hệ giữa các cụm từ khóa và thành phần kiến thức, cuối cùng đưa ra kết quả truy vấn tương ứng. Các vấn đề cho việc thiết kế hệ thống truy vấn kiến thức hỗ trợ trong học tập:

**Vấn đề 1. Phân loại mục đích câu truy vấn:** với câu truy vấn đầu vào nhận được, hệ thống tiến hành sử dụng bộ công cụ phân tách câu truy vấn thành các cụm từ khóa. Từ các cụm từ khóa có được, hệ thống lựa chọn ra nội dung tri thức phù hợp và lấy kết quả kiến thức từ cơ sở tri thức.

**Vấn đề 2. Truy xuất nội dung kiến thức dựa trên từ khóa:** với tập từ khóa được trích xuất từ truy vấn, phương pháp so sánh sự giống nhau giữa nghĩa của từ khóa với nội dung trong cơ sở tri thức được đề xuất. Phương pháp này xác định nội dung kiến thức cần thiết cho truy vấn được nhập vào.

**Vấn đề 3. Tìm kiếm kiến thức liên quan:** khi tìm kiếm một kiến thức, người học cũng cần có một số kiến thức liên quan đến kiến thức đó. Các kiến thức liên quan giúp người học hiểu thêm hơn nội dung tìm kiếm.

#### 2.3.1. Phân loại mục đích câu truy vấn

Trên thực tế, có nhiều loại câu truy vấn khác nhau. Trong mô hình này chúng tôi đề cập đến bốn loại truy vấn thường gặp trong truy vấn kiến thức môn học:

- Truy vấn khái niệm (*ask.define*);
- Truy vấn bài tập (*ask.assignment*);

- Truy vấn thuộc tính, tính chất (*ask.properties, ask.using...*);
- Truy vấn so sánh giữa hai khái niệm (*ask.compare*).

Phương pháp phân loại ý định được sử dụng trong hệ thống là Pattern Matching. Từ output của Aspect Extractor, hệ thống sẽ lọc ra các normal tags rồi thực hiện so sánh với danh sách các alias của các ý định. Thuật toán dùng để so sánh là BM25. BM25 là một thuật toán xếp hạng được cải tiến từ nền tảng là TF-IDF (Term frequency – Inverse document frequency), trong đó sự tương tác được xem như một đại lượng xác suất. BM25 sẽ đánh trọng số của mỗi văn bản dựa trên một truy vấn tương ứng. Cụ thể, với một truy vấn  $Q$ , và  $q_1, \dots, q_n$ ,  $q, \dots, q_n$  là trọng số của văn bản  $d$ . Khi đó, ta có:

$$BM25(D, Q) := \sum_{i=1}^n idf(q_i, d) \frac{F(q_i, D) \cdot (k+1)}{F(q_i) + k \cdot \left(1 - b + b \cdot \frac{|D|}{d_{avg}}\right)} \quad (1)$$

trong đó: +  $F(q_i, D)$ : số lượng xuất hiện của token  $q_i$  trong văn bản  $D$  ( $1 \leq i \leq n$ );  
 +  $|D|$ : số lượng token trong  $D$ ;  
 +  $d_{avg}$ : trung bình số lượng token trung bình trong một văn bản;  
 +  $b$  và  $k$ : các tham số của BM25.

Bước phân loại câu truy vấn này giúp hệ thống khai thác được nội dung câu trả lời phù hợp. Hệ thống sử dụng bộ công cụ PyVi để tách câu truy vấn đầu vào thành các cụm từ khóa. Từ các cụm từ khóa tách được, hệ thống tiến hành sử dụng thuật toán xếp hạng BM25 để trả về loại câu truy vấn phù hợp.

- Câu truy vấn khái niệm: Sau khi hệ thống trích xuất các cụm từ khóa từ câu truy vấn mà người dùng nhập vào sẽ được nhận dạng và phân loại câu truy vấn của hệ thống.

*Ví dụ:* Với câu truy vấn đầu vào là “Tôi muốn biết về khóa ngoại”. Sau khi tách, hệ thống lấy được cụm từ “muốn biết về” và phân loại câu truy vấn trên là loại “*ask.define*”.

- Câu truy vấn thuộc tính, tính chất và bài tập: các câu truy vấn này có cách phân loại giống với câu truy vấn khái niệm, với mỗi câu đầu vào của người dùng sẽ được xử lý dựa vào các cụm từ khóa và được thuật toán xếp hạng chọn ra loại câu truy vấn phù hợp.

*Ví dụ:* với câu “Cho tôi bài tập về đại số quan hệ”, cụm từ “bài tập” sẽ có yếu tố quyết định trả về loại câu truy vấn là “*ask.assignment*”.

- Câu truy vấn so sánh: hệ thống phân loại câu truy vấn dạng này sẽ trả về loại câu truy vấn so sánh và hai cụm từ khóa là tên của hai kiến thức mà người dùng muốn so sánh. Hệ thống sẽ so sánh từng thuộc tính của hai kiến thức và hiện thị sự khác biệt từng phần giữa hai đối tượng được so sánh.

### 2.3.2. Truy xuất nội dung tìm kiếm dựa trên từ khóa

Hệ thống tìm kiếm nội dung phù hợp dựa trên việc so khớp từ khóa được tách với từ khóa trong cơ sở tri thức. Bộ từ khóa trong cơ sở tri thức được thu thập dựa trên việc trích xuất từ khóa các câu truy vấn liên quan đến môn Cơ sở dữ liệu. Hệ thống sẽ tiến hành xếp hạng và truy xuất nội dung và cho ra kết quả tương ứng.

Cho miền tri thức  $\mathcal{K} = (\mathbf{C}, \mathbf{R}, \mathbf{Ops}, \mathbf{Rules}) \oplus \mathbf{KKG}$  và câu truy vấn  $q$ , thuật toán sau sẽ lấy ra được nội dung tri thức phù hợp với câu truy vấn  $q$ .

**Input:** Miền tri thức  $\mathcal{K}$  và câu truy vấn  $q$ .

**Output:** Nội dung kiến thức trong  $\mathcal{K}$  phù hợp với câu truy vấn  $q$ .

#### Thuật giải 4.1

##### Bước 1.

**Trích xuất** từ khóa có nghĩa gần với từ khóa  $W$  trong câu truy vấn  $q$

$W := \text{Meaning}(q)$

Knowledge := {} //tập kết quả truy vấn

If (tồn tại từ khóa  $W$  trong miền tri thức  $\mathcal{K}$ )

$\text{Search} :=$  tập hợp các thành phần của  $\mathcal{K}$  liên quan đến kiến thức.

**Bước 2.** Trích xuất tri thức từ các phần tử trong  $\text{Search}$  dựa trên từ khóa trong  $W$ .

**Cập nhật** Knowledge.

**Bước 3.** Hợp nhất các thực tại của mô hình tri thức và so sánh từ khóa của chúng bằng cách phương pháp giải quyết Vấn đề 1 trong mục 4.1.

**Cập nhật** Knowledge.

**Bước 4.** Trả về kết quả là nội dung trong Knowledge.

#### 2.3.3. Tìm kiếm kiến thức liên quan

Các kiến thức liên quan giúp sinh viên và người dùng hiểu đầy đủ thêm về chủ đề mà họ tìm kiếm. Với câu truy vấn đầu vào ngoài việc trả lời nội dung kiến thức phù hợp thì hệ thống sẽ tìm kiếm các kiến thức liên quan đến nội dung đó.

Cho miền tri thức  $\mathcal{K} = (\mathbf{C}, \mathbf{R}, \mathbf{Ops}, \mathbf{Rules}) \oplus \mathbf{KKG}$  và câu truy vấn đầu vào  $q$ , thuật toán sau sẽ trả về tất cả tri thức có liên quan đến tri thức phù hợp với  $q$ .

**Input:** Miền tri thức  $\mathcal{K}$  và câu truy vấn  $q$ .

**Output:** Tập kiến thức liên quan đến kiến thức phù hợp với  $q$ .

#### Thuật giải 4.2:

**Bước 1:**  $\text{Knowledge} :=$  Tập nội dung kiến thức phù hợp với  $q$ . # Được xác định trong Thuật giải 4.1

##### Bước 2:

for each  $k$  in  $\text{Knowledge}$  do

**Xác định** tập kiến thức liên quan đến  $k$ .

$\text{Related}(k) := \{k' \mid \exists \text{rel} \in \mathcal{K}.\mathbf{R}, \text{rel}(k, k')\}$

**Bước 3:** **Hiển thị** kết quả trong  $\text{Related}(\text{Knowledge})$  dưới dạng một tập các đối tượng kiến thức.

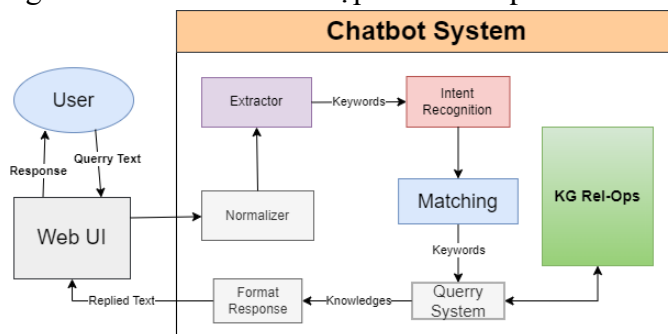
$\text{Related}(\text{Knowledge}) := \bigcup_{k \in K} \text{Related}(k)$

#### 2.4. Thiết kế Chatbot thông minh truy vấn kiến thức môn Cơ sở dữ liệu



2.4.1. Kiến trúc của hệ thống Chatbot hỗ trợ truy vấn

Kiến trúc của môn Cơ sở dữ liệu được thu thập từ các sách (Elmasri & Navathe, 2016; Silberschat et al., 2019). Các kiến trúc được thu thập này được phân loại theo các cách khác nhau như theo chương, đơn vị, chủ đề, bài tập và các phương pháp giải. Nội dung của môn này được lưu trữ bằng mô hình tri thức tích hợp KG Rel-Ops model.



Hình 3. Kiến trúc của hệ thống Chatbot thông minh hỗ trợ truy vấn

Hệ thống gồm các phần chính:

- + Giao diện người dùng: Tiếp nhận các câu truy vấn và hiển thị câu trả lời cho của hệ thống;
- + Normalizer: Chuẩn hóa các từ giúp cho việc phân loại và so khớp;
- + Extractor: Trả về tập các từ khóa được tách từ bởi công cụ tách từ PyVi;
- + Intent recognition: phân loại câu truy vấn từ tập các cụm từ bằng thuật toán xếp hạng;
- + Matching: Xử lý việc so khớp giữa từ khóa có trong cơ sở tri thức và nhận tập tri thức tương ứng;
- + Mô hình KG Rel-Ops model: là mô hình tích hợp cấu trúc đồ thị tri thức ontology Rela-Ops như trong Định nghĩa 3.3.

Hình 4 là giao diện của hệ thống chatbot hỗ trợ truy vấn kiến trúc môn học Cơ sở dữ liệu.



Hình 4. Giao diện của hệ thống Chatbot truy vấn kiến trúc môn Cơ sở dữ liệu

2.4.2. Thử nghiệm hệ thống

Hệ thống chatbot được thử nghiệm và đánh giá độ chính xác dựa trên nội dung trả về khi thực hiện truy vấn. Thử nghiệm được thực với 2923 câu truy vấn, với các chủ đề truy vấn được thể hiện trong Bảng 1. Hệ thống đạt độ chính xác cao với các câu truy vấn về tính chất thuộc tính (trên 90%), tuy nhiên, với các câu truy vấn về việc so sánh thì chatbot chưa trả lời chính xác (khoảng 70%). Điều này, do câu truy vấn dạng so sánh có nhiều yêu cầu khác nhau nên chatbot chưa thể xử lý được một cách đầy đủ, để từ đó có thể truy xuất kiến

thức một cách phù hợp. Về mặt tổng thể, những câu truy vấn trả lời chưa chính xác là do hệ thống chưa phân tích đầy đủ ngữ nghĩa của câu truy vấn được nhập vào.

**Bảng 1. Kết quả thực nghiệm trên số câu truy vấn thử nghiệm**

Loại câu truy vấn	Số lượng câu truy vấn	Số câu chính xác	Tỉ lệ
Khái niệm	2,322	1,993	85,8%
Luật/Quy tắc	72	67	93,1%
Tính chất/Thuộc tính	182	165	90,7%
Bài tập	67	58	86,6%
So sánh	280	197	70,8%
<b>Tổng</b>	<b>2,923</b>	<b>2,480</b>	<b>84,8%</b>

**2.4.3. So sánh với các hệ thống mô hình Ngôn ngữ lớn (LLMs)**

Hiện nay, có nhiều mô hình ngôn ngữ lớn thực hiện các tác vụ như một chatbot, chẳng hạn ChatGPT của Open AI (<https://chat.openai.com/>), Google Gemini (<https://gemini.google.com/chat>), Microsoft Copilot (<https://copilot.microsoft.com>). Các mô hình này đều là một hệ thống chatbot đa tính năng, có thể thực hiện nhiều tác vụ khác nhau, trong đó cũng có tính năng truy vấn kiến thức về lĩnh vực Cơ sở dữ liệu. Mô hình ChatGPT và Copilot đều sử dụng mô hình ngôn ngữ lớn là GPT 3.5 (Generative Pre-trained Transformer) (GPT 4.0 được sử dụng cho các phiên bản nâng cao). Đây là một mô hình học máy mạnh mẽ được đào tạo trên một lượng lớn dữ liệu để thực hiện nhiều công việc ngôn ngữ tự nhiên, chẳng hạn như tạo văn bản, dịch ngôn ngữ, trả lời câu hỏi, và nhiều ứng dụng khác. Gemini được xây dựng dựa trên một mô hình PaLM 2 (<https://ai.google/discover/palm2/>). PaLM là một mô hình Transformer được đào tạo trên một tập dữ liệu khổng lồ gồm văn bản và mã nguồn, tương tự như GPT 3.5.

Trong mục này, hệ thống chatbot truy vấn kiến thức môn học Cơ sở dữ liệu được so sánh với ChatGPT, Gemini và Copilot theo ba tiêu chí đối với một hệ thống hỗ trợ học tập thông minh: Tính chính xác, Tính sự phạm và Sự hữu ích, tiện dụng (Nguyen et al., 2018; Almeida et al., 2019). Kết quả so sánh được thể hiện trong Bảng 2.

**Bảng 2. So sánh với các mô hình ngôn ngữ lớn**

Tiêu chí	ChatGPT (mô hình GPT 3.5)	Gemini (mô hình PaLM 2)	Copilot (mô hình GPT 3.5)	Hệ thống chatbot
Tính chính xác	<ul style="list-style-type: none"> <li>Hệ thống cung cấp câu trả lời dựa trên lượng dữ liệu lớn đã học.</li> <li>Tuy nhiên kiến thức sẽ không được cập nhật liên tục</li> <li>Thỉnh thoảng ChatGPT có thể tạo ra câu trả lời không chính xác hoặc không đầy đủ, đặc biệt là khi truy vấn phức tạp hoặc chuyên sâu.</li> </ul>	<ul style="list-style-type: none"> <li>Được học từ bộ dữ liệu lớn của Google. Hệ thống có thể trả lời các câu hỏi đặt ra.</li> <li>Độ chính xác của câu trả lời vẫn chưa được đảm bảo vì hệ thống vẫn đang trong phiên bản thử nghiệm.</li> </ul>	<ul style="list-style-type: none"> <li>Hệ thống được học từ bộ dữ liệu lớn và đa dạng.</li> <li>Copilot có 3 chức năng trả lời cho người dùng lựa chọn là: More Creative, More Balance, More Precise. Đối với chức năng More Precise câu trả lời được lựa chọn các nguồn chính xác và độ chính xác cao nhưng vẫn dựa hoàn toàn vào dữ liệu đã học.</li> </ul>	<ul style="list-style-type: none"> <li>Dữ liệu được thu thập từ các tài liệu chuẩn như giáo trình, các khóa học của các trường đại học, kiến thức của giảng viên viên. Từ đó, thiết lập cơ sở tri thức cho hệ thống một cách đầy đủ và chính xác.</li> <li>Câu trả lời của hệ thống bám đạo độ chính xác về mặt chuyên môn.</li> </ul>

<p>Tính sự phạm</p>	<ul style="list-style-type: none"> <li>• Người học có thể tham khảo nội dung của hệ thống nhưng phải có chọn lọc vì có thể nội dung không chính xác.</li> <li>• Hệ thống không có hình ảnh trực quan.</li> <li>• Hệ thống chưa quan tâm đến yếu tố sư phạm, và chi tương tác theo các nội dung được tìm thấy.</li> </ul>	<ul style="list-style-type: none"> <li>• Người học có thể sử dụng kiến thức của hệ thống</li> <li>• Hệ thống tương tác với người học dựa trên các yếu tố tâm lí sư phạm, giúp gợi mở khả năng tìm tòi kiến thức thông qua việc gợi ý các kiến thức liên quan.</li> </ul>	
<p>Sự hữu ích, tiện dụng</p>	<ul style="list-style-type: none"> <li>• Giao diện dễ sử dụng</li> <li>• Tốc độ trả lời ổn định.</li> <li>• Hệ thống không có hình ảnh trực quan.</li> </ul>	<ul style="list-style-type: none"> <li>• Giao diện hiện đại dễ nhìn.</li> <li>• Tốc độ trả lời trung bình, thường chậm hơn so với ChatGPT.</li> <li>• Hệ thống không có hình ảnh trực quan.</li> </ul>	<ul style="list-style-type: none"> <li>• Giao diện dễ sử dụng, tốc độ trả lời nhanh, có gợi ý các câu hỏi liên quan.</li> <li>• Chương trình có hỗ trợ hình ảnh trực quan với từng nội dung kiến thức.</li> </ul>

### 3. Kết luận và hướng phát triển

Trong nghiên cứu này, một mô hình tích hợp ontology tri thức quan hệ – toán tử và đồ thị tri thức, gọi là KG Rel-Ops model, được đề xuất. Mô hình đề xuất có thể biểu diễn đầy đủ các kiến thức của môn học và tăng khả năng thích ứng làm cho cấu trúc của mô hình linh hoạt hơn, có thể áp dụng để biểu diễn các miền tri thức khác nhau. Bên cạnh đó, cấu trúc đồ thị giúp cho việc truy vấn trên cơ sở tri thức như tìm kiếm, so sánh thông tin hay gợi ý các kiến thức liên quan cũng được nghiên cứu và giải quyết. Giải pháp đề xuất được ứng dụng cơ sở tri thức cho môn học Cơ sở dữ liệu và hỗ trợ việc thiết kế hệ thống tra cứu kiến thức dạng chatbot. Hệ thống có thể hỗ trợ trong việc tra cứu nội dung theo bài học, theo phân loại kiến thức và theo các dạng bài tập trong môn học. Hệ thống có thể truy vấn thông tin nhanh chóng, và kết quả trả về được hiển thị đầy đủ và đáp ứng các yêu cầu về sư phạm trong việc hướng dẫn học tập. Hệ thống xây dựng cũng được đánh giá và so sánh với các mô hình LLMs hiện nay, như ChatGPT, Google Gemini, và Microsoft Copilot, trong việc hỗ trợ học tập với môn học Cơ sở dữ liệu.

Trong tương lai, hệ thống chatbot thông minh có thể phát triển thêm nhiều chức năng để hỗ trợ người dùng trong việc tìm kiếm kiến thức. Hệ thống chatbot truy vấn kiến thức sẽ giúp người dùng có thể học tập, kiểm tra tự đánh giá trình độ kiến thức của mình. Thông qua mô hình KG Rel-Ops model, cơ sở tri thức có thể kết hợp một số miền kiến thức (Kusairi, 2020) để công cụ truy vấn có thể hỗ trợ việc nghiên cứu các môn học ngành Công nghệ Thông tin ở mức cơ bản tại các trường đại học. Ngoài ra, chatbot còn mang đến một môi trường học tập trực tuyến hấp dẫn hơn, thúc đẩy việc học tập của học sinh, sinh viên (Okonkwo & Ade-Ibijola, 2022; Weber et al., 2021). Bên cạnh đó, nghiên cứu để xây dựng các hệ thống hướng dẫn hỗ trợ việc hướng dẫn giải bài tập sẽ giúp cho người học có thể học tập hiệu quả hơn trong việc hiểu và vận dụng kiến thức trong các bài tập của môn học (Trinh et al., 2024; Nguyen, 2023).

- ❖ **Tuyên bố về quyền lợi:** Các tác giả xác nhận hoàn toàn không có xung đột về quyền lợi.
- ❖ **Lời cảm ơn:** Nghiên cứu này được tài trợ bởi Nguồn ngân sách khoa học và công nghệ Trường Đại học Sư phạm Thành phố Hồ Chí Minh trong đề tài mã số CS.2023.19.19.

### TÀI LIỆU THAM KHẢO

- Abu-Salih, B., & Alotaibi, S. (2024). A systematic literature review of knowledge graph construction and application in education. *Heliyon*, 2024, e25383. <https://doi.org/10.1016/j.heliyon.2024.e25383>
- Agrawal, M., Heggelmann, S., Lang, H., Kim, Y., & Sontag, D. (2022). Large language models are few-shot clinical information extractors. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP 2022)*, (pp. 1998-2022). <https://doi.org/10.18653/v1/2022.emnlp-main.130>
- Akhter, S., Javed, M. K., Shah, S. Q., & Javaid, A. (2021). Highlighting the advantages and disadvantages of E-learning. *Psychol. Educ*, 58(5), 1607-1614.
- Almeida, M., Ferreira, J., & Ferreira, F. A. (2019). Developing a multi-criteria decision support system for evaluating knowledge transfer by higher education institutions. *Knowledge Management Research & Practice*, 17(4), 358-372. <https://doi.org/10.1080/14778238.2018.1534533>
- Angles, R., & Gutiérrez, C. (2008). Survey of graph database models. *ACM Comp. Surv.*, 40(1), Article 1. <https://doi.org/10.1145/1322432.1322433>
- Dale, R. (2016). The return of the chatbots. *Natural Language Engineering*, 22(5), 811-817. <https://doi.org/10.1017/S1351324916000243>
- Davies, J. N., Verovko, M., Verovko, O., & Solomakha, I. (2020). Personalization of e-learning process using AI-powered chatbot integration. In *Proceedings of International scientific-practical conference, MODS 2020*, (pp. 209-216). Springer. [https://doi.org/10.1007/978-3-030-58124-4\\_20](https://doi.org/10.1007/978-3-030-58124-4_20)
- Do, N. V., Nguyen, H. D., & Hoang, L. N. (2020). Some techniques for intelligent searching on ontology-based knowledge domain in e-learning. In *Proceedings of 12<sup>th</sup> International Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2020)*, (vol. 2, pp. 313-320). <https://doi.org/10.5220/0010174403130320>
- Elayyan, S. (2021). The future of education according to the fourth industrial revolution. *Journal of Educational Technology and Online Learning*, 4(1), 23-30. <https://doi.org/10.31681/jetol.737193>
- Elmasri, R., & Navathe, S. B. (2016). *Fundamentals of Database Systems*. Global Edition.
- Foster, E., & Godbole, S. (2022). *Database systems: a pragmatic approach*. Taylor & Francis.
- Følstad, A., & Brandtzaeg, P. (2017). Chatbots – The new world of HCI. *ACM Interactions*, 24, 38-42. <https://doi.org/10.1145/3085558>
- Guemmat, K., & Ouahabi, S. (2018). A literature review of indexing and searching techniques implementation in educational search engines. *International Journal of Information and Communication Technology Education (IJICTE)*, 14(2), 72-83. <https://doi.org/10.4018/IJICTE.2018040106>

- Hagedorn, T., Bone, M., Kruse, B., Grosse, I., & Blackburn, M. (2020). Knowledge representation with ontologies and semantic web technologies to promote augmented and artificial intelligence in systems engineering. *Insight*, 23(1), 15-20. <https://doi.org/10.1002/inst.12279>
- Hogan, A., Blomqvist, E., Cochez, M., D'Amato, C., De Melo, G., Gutierrez, C., Kirrane, S., Gayo, J. E. L., Navigli, R., Neumaier, S., Ngomo, A.-C. N., Polleres, A., Rashid, S. M., Rula, A., Schmelzeisen, L., Sequeda, J., Staab, S., & Zimmermann, A. (2020). Knowledge graphs. *ACM Computing Surveys*, 54(4), Article 71. <https://doi.org/10.1145/3447772>
- Kusairi, S. (2020). A Web-Based Formative Feedback System Development by Utilizing Isomorphic Multiple Choice Items to Support Physics Teaching and Learning. *Journal of Technology and Science Education*, 10(1), 117-126. <https://doi.org/10.3926/jotse.781>
- Liu, Q., Huang, Z., Huang, Z., Liu, C., Chen, E., Su, Y., & Hu, G. (2018). Finding similar exercises in online education systems. In *Proceedings of the 24<sup>th</sup> ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD 2018)*, (pp. 1821-1830). <https://doi.org/10.1145/3219819.3219960>
- Nguyen, H., Do, N., Tran, N., Pham, H. (2018). Criteria of a Knowledge model for an Intelligent Problems Solver in Education. In *Proceedings of 10<sup>th</sup> IEEE International Conference on Knowledge and Systems Engineering (KSE 2018)*, (pp. 288-293). <https://doi.org/10.1109/KSE.2018.8573434>
- Nguyen, H. D., Do, N. V., Pham, V. T., Selamat, A., & Herrera-Viedma, E. (2020). A method for knowledge representation to design Intelligent Problems Solver in mathematics based on Rela-Ops model. *IEEE Access*, 8, 76991-77012. <https://doi.org/10.1109/ACCESS.2020.2988929>
- Nguyen, H., Tran, T. V., Pham, X. T., Huynh, A. T., & Do, N. V. (2021). Ontology-based integration of knowledge base for building an intelligent searching chatbot. *Sensors & Materials*, 33(9), 3101-3123. <https://doi.org/10.18494/SAM.2021.3264>
- Nguyen, H. D., Do, N. V., & Pham, V. T. (2022). A methodology for designing knowledge-based systems and applications. In *Applications of Computational Intelligence in Multi-Disciplinary Research*, (pp. 159-185). Academic Press. <https://doi.org/10.1016/B978-0-12-823978-0.00001-0>
- Nguyen, H. D., Truong, D., Vu, S., Nguyen, D., Nguyen, H., & Tran, N. T. (2023). Knowledge Management for Information Querying System in Education via the Combination of Rela-Ops Model and Knowledge Graph. *Journal of Cases on Information Technology (JCIT)*, 25(1), Article 13. <http://doi.org/10.4018/JCIT.324113>
- Nguyen, H. (2023). Intelligent System in Education: Requirements and Design Method, *Journal of Electronic Voltage and Application (JEVA)*, 4(2), 12-19. <https://doi.org/10.30880/jeva.2023.04.02.002>
- Ngo, H. Q., Nguyen, H. D., & Le-Khac, N. A. (2024). Ontology Knowledge Map Approach Towards Building Linked Data for Vietnamese Legal Applications. *Vietnam Journal of Computer Science*, 11(02), 323-342. <https://doi.org/10.1142/S2196888824500015>
- Noy, N., Gao, Y., Jain, A., Narayanan, A., Patterson, A., & Taylor, J. (2019). Industry-scale knowledge graphs: Lessons and challenges. *Communications of the ACM*, 62(8), 36-43. <https://doi.org/10.1145/3331166>
- Okonkwo, C. W., & Ade-Ibijola, A. (2022). Chatbots applications in education: A systematic review. *Computers and Education: Artificial Intelligence*, 2, Article 100033. <https://doi.org/10.1016/j.caeai.2021.100033>
- Silberschatz, A., Korth, H., & Sudarshan, S. (2019). *Database System Concepts (7<sup>th</sup> ed.)*. McGraw-Hill.

- Trinh, T. H., Wu, Y., Le, Q. V., He, H., & Luong, T. (2024). Solving olympiad geometry without human demonstrations. *Nature*, 625, 476-482. <https://doi.org/10.1038/s41586-024-07115-7>
- Truong, D. M., Nguyen, H. D., Vu, S., Pham, V. T., & Nguyen, D. (2022). Construct an intelligent querying system in education based on ontology integration. In *Proceedings of the 2<sup>nd</sup> IEEE International Conference on Computing (ICOCO 2022)*, (pp. 340-345). IEEE. <https://doi.org/10.1109/ICOCO56118.2022.10031735>
- Weber, F., Wambsganss, T., Rüttimann, D., & Söllner, M. (2021). Pedagogical Agents for Interactive Learning: A Taxonomy of Conversational Agents in Education. *Proceedings of 42<sup>nd</sup> International Conference on Information Systems (ICIS 2021)* (pp. 1-17). Austin.
- Yoo, S., & Jeong, O. (2020). Auto-growing knowledge graph-based intelligent chatbot using BERT. *ICIC Express Lett*, 14(1), 67-73. <https://doi.org/10.24507/icicel.14.01.67>
- Zhang, F., & Li, X. (2024). Knowledge-enhanced online doctor recommendation framework based on knowledge graph and joint learning. *Information Sciences*, 662, Article 120268. <https://doi.org/10.1016/j.ins.2024.120268>

---

## INTEGRATING KNOWLEDGE BASED ON ONTOLOGY AND KNOWLEDGE GRAPH FOR CHATBOT IN EDUCATION QUERY

Nguyễn Việt Hưng<sup>1</sup>, Lê Thị Ngọc Thảo<sup>1</sup>, Nguyễn Văn Hậu<sup>1</sup>,  
Nguyễn Đắc Long<sup>1</sup>, Trần Phong Nha<sup>2</sup>, Nguyễn Đình Hiền<sup>3\*</sup>

<sup>1</sup>Ho Chi Minh City University of Education, Vietnam

<sup>2</sup>Campus in Ho Chi Minh City, University of Transport and Communications, Vietnam

<sup>3</sup>University of Information Technology, Vietnam National University Ho Chi Minh City, Vietnam

\*Corresponding author: Nguyễn Đình Hiền – Email: [hiennnd@uit.edu.vn](mailto:hiennnd@uit.edu.vn)

Received: February 10, 2024; Revised: March 27, 2024; Accepted: March 29, 2024

### ABSTRACT

Online learning has become increasingly prevalent in recent years, allowing learners to utilize electronic devices and internet platforms for studying, accessing materials, and acquiring knowledge. This study proposes a model that integrates relational knowledge ontology, operators, and knowledge graphs to optimize the representation of subject knowledge combined with specifications of relationships between knowledge components. Based on the constructed knowledge base, the model solves knowledge query problems to meet the requirements for building intelligent learning support systems. The proposed solution is applied to develop a knowledge query support system for the Database course in the form of a chatbot. This chatbot can facilitate querying course knowledge content according to knowledge classification and types of exercises within the course. The constructed system is evaluated and compared with current large language models in terms of supporting the study of the Database course.

**Keywords:** chatbot; database; intelligent system; knowledge-based systems; knowledge engineering