

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG



Phan Anh Dũng

**ỨNG DỤNG MÔ HÌNH DỮ LIỆU ĐỒ THỊ
TRONG PHÁT TRIỂN MẠNG THÔNG TIN SỨC KHỎE**

Chuyên ngành: Khoa học máy tính

Mã số: 8.48.01.01

TÓM TẮT LUẬN VĂN THẠC SĨ

HÀ NỘI – NĂM 2023

Luận văn được hoàn thành tại:

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

Người hướng dẫn khoa học:

.....

(Ghi rõ học hàm, học vị)

Phản biện 1:

.....

Phản biện 2:

.....

Luận văn sẽ được bảo vệ trước Hội đồng chấm luận văn thạc sĩ tại Học viện Công nghệ Bưu chính Viễn thông

Vào lúc: giờ ngày tháng năm

Có thể tìm hiểu luận văn tại:

- Thư viện của Học viện Công nghệ Bưu chính Viễn thông.

MỞ ĐẦU

1. Tính cấp thiết của đề tài

Trong những năm gần đây, cùng với sự phát triển kinh tế thì nhu cầu tư vấn khám chữa bệnh của người dân gia tăng nhanh chóng. Các công tác xã hội được khuyến khích phát triển trong lĩnh vực y tế nhằm hỗ trợ các y bác sĩ giảm bớt áp lực công việc, nâng cao hiệu quả điều trị. Các mạng thông tin hiện hành thiếu các nội dung, các công cụ cần thiết chuyên sâu để đáp ứng các nhu cầu giao tiếp, kết nối về y tế. Vì vậy, nhu cầu có một mạng thông tin sức khỏe chuyên biệt để làm nơi giao lưu, trao đổi thông tin là cấp thiết.

2. Tổng quan vấn đề nghiên cứu

Vấn đề nghiên cứu của đề tài: ứng dụng mô hình dữ liệu đồ thị và khoa học dữ liệu dựa trên đồ thị (graph data science) trong phát triển mạng thông tin sức khỏe.

Đầu tiên, cần nghiên cứu về mạng thông tin. Cần làm rõ thế nào là mạng thông tin? Các đặc trưng riêng biệt cần có của mạng thông tin trong lĩnh vực y tế hay còn gọi là mạng thông tin sức khỏe. Xác định đối tượng, mục tiêu sử dụng của mạng thông tin này trong việc chia sẻ thông tin nhằm nâng cao chất lượng chăm sóc y tế. Các vấn đề khó khăn thực tiễn đặt ra trong sự phát triển tin học y tế của các khối bệnh viện Việt Nam hiện nay. Sau đó, luận văn cần xây dựng ra được cơ sở lý luận của việc sử dụng mô hình dữ liệu đồ thị bằng cách nghiên cứu các mô hình, các hướng tiếp cận nhằm giải quyết bài toán thiết lập mạng thông tin. Các mô hình dữ liệu quan hệ, hướng đối tượng và đồ thị có ưu điểm, nhược điểm là gì? Khả năng phát triển, mở rộng về sau của các mô hình đó có thuận lợi không, có phù hợp với nhu cầu đặt ra của mạng thông tin không? Sau đó quá trình tiến hành, thử nghiệm có thuận lợi không?

Luận văn cần đặt ra các nghiên cứu hệ quản trị dữ liệu đồ thị. Việc tìm kiếm các nền tảng liên quan đến dữ liệu đồ thị với các công cụ, cộng đồng hỗ trợ mạnh mẽ là rất quan trọng. Luận văn cần khảo sát một số nền tảng hệ quản trị CSDL để lựa chọn ra nền tảng phù hợp với các tiêu chí đặt ra (Neo4j, Jena, GraphDB...). Hệ quản trị CSDL này cần lưu trữ các mô hình dữ liệu đồ thị trên đó và có các công cụ truy vấn, trực quan mạnh mẽ hỗ trợ phát triển các ứng dụng.

Tiếp sau đó, luận văn sẽ nghiên cứu về việc phát triển mạng thông tin sức khỏe cụ thể là xây dựng website và cơ sở dữ liệu để cung cấp các truy xuất thông tin về sức khỏe, gồm các đối tượng thông tin được liên kết với nhau: bác sĩ, phòng khám, hệ thống thông tin ngành,

chuyên ngành, các cấp y tế dự phòng v.v... Làm rõ hệ thống này sử dụng mô hình và hệ thống dữ liệu nào, có ưu điểm, hạn chế gì. Phần tiếp theo của luận văn sẽ là phần áp dụng mô hình dữ liệu đồ thị xây dựng nên hệ thống truy xuất thông tin sức khỏe này. Sau cùng là sự so sánh, đánh giá khả năng, tính phù hợp của mô hình dữ liệu đồ thị với các mô hình dữ liệu khác trong việc phát triển mạng thông tin sức khỏe.

3. Mục đích nghiên cứu

Về mặt lý luận: Nghiên cứu các mô hình dữ liệu phổ biến: mô hình dữ liệu hướng đối tượng, mô hình dữ liệu quan hệ và mô hình dữ liệu đồ thị. Hiểu rõ cơ sở lý thuyết về mô hình dữ liệu đồ thị và khoa học dữ liệu đồ thị và các ứng dụng của nó.

Về thực tiễn: Phân tích, so sánh, đánh giá, thử nghiệm trên các mô hình dữ liệu với tập dữ liệu đầu vào ban đầu. Sử dụng, vận hành Neo4j, xây dựng ứng dụng linh hoạt, hoạt động với các chức năng cơ bản minh họa cho tính linh động của mô hình. Xây dựng một ứng dụng web và CSDL cung cấp thông tin ngành y tế.

4. Đối tượng và phạm vi nghiên cứu

Đối tượng nghiên cứu: Các mô hình dữ liệu đặc biệt là mô hình dữ liệu đồ thị, khoa học dữ liệu. Hệ quản trị CSDL đồ thị Neo4j. Các công cụ để lập trình ứng dụng trong phát triển mạng thông tin. Ứng dụng trong phát triển mạng thông tin y tế và sức khỏe.

Phạm vi nghiên cứu: Nghiên cứu lý thuyết mô hình dữ liệu đồ thị, khoa học dữ liệu đồ thị và một số hệ quản trị CSDL đồ thị (Neo4j, Jena, GraphDB...). Xây dựng mô hình, kiểm thử phần mềm trong phạm vi một cơ quan y tế địa phương tuyến cơ sở.

5. Phương pháp nghiên cứu

Phương pháp nghiên cứu lý thuyết: Nghiên cứu cơ sở lý thuyết về các mô hình dữ liệu. Đọc và phân tích các tài liệu về mô hình dữ liệu đồ thị và các nghiên cứu liên quan về khoa học dữ liệu đồ thị. Nghiên cứu các thuật toán, phương pháp luận trong việc phát triển mạng thông tin ứng dụng hệ thống quản trị dữ liệu đồ thị.

Phương pháp thực nghiệm: Xây dựng, thử nghiệm và đánh giá độ hiệu quả của các mô hình dữ liệu. Xây dựng hệ thống mạng thông tin sức khỏe dựa trên mô hình dữ liệu đồ thị. Kiểm thử tính năng, đánh giá chất lượng sản phẩm.

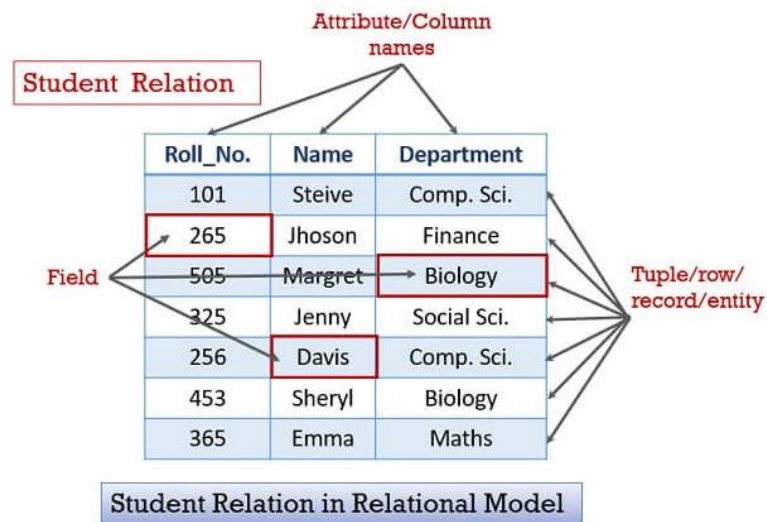
CHƯƠNG 1. CƠ SỞ LÝ THUYẾT

1.1. Mô hình và khoa học dữ liệu

1.1.1. Mô hình dữ liệu quan hệ.

a. Khái niệm mô hình dữ liệu quan hệ:

Mô hình dữ liệu (MHDL) quan hệ (RDM – Relational Data Model) biểu diễn cơ sở dữ liệu dưới dạng một tập hợp các quan hệ, tương ứng với bảng giá trị trong đó mỗi quan hệ có các thuộc tính (attributes) và bộ giá trị (tuples) tương ứng với các cột và hàng. Mỗi bộ giá trị tượng trưng cho một thực thể hoặc mỗi quan hệ trong thế giới thực và tên của quan hệ cùng các thuộc tính cung cấp thông tin về ý nghĩa của từng bộ giá trị.



Hình 1- 1: Minh họa mô hình dữ liệu quan hệ [11]

b. Cấu trúc và thành phần của một mô hình dữ liệu quan hệ:

Một MHDL quan hệ bao gồm các thành phần sau: Bảng (Table), Cột (Column), Dòng (Row), Khóa chính (Primary key), Khóa ngoại (Foreign key), Ràng buộc (Constraint).

c. Các phương thức truy vấn dữ liệu trong mô hình dữ liệu quan hệ:

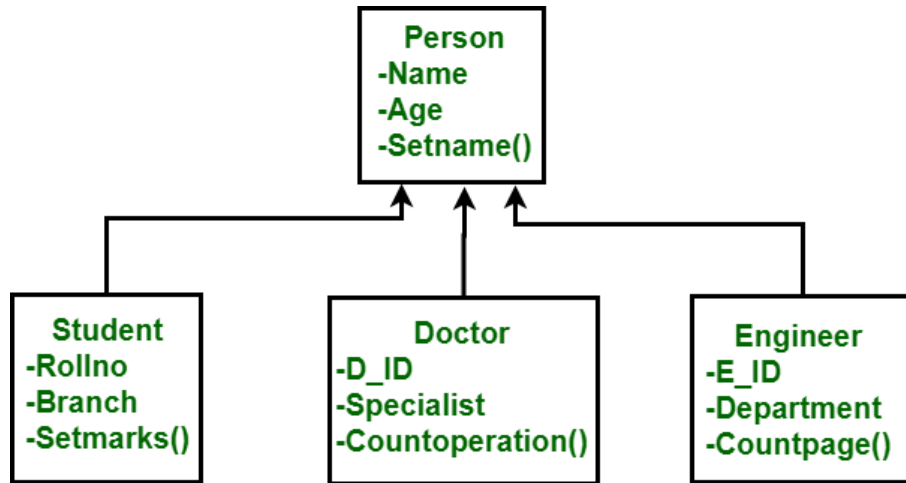
SELECT, INSERT, UPDATE, DELETE, JOIN, GROUP BY, ORDER BY, DISTINCT

d. Ví dụ minh họa

1.1.2. Mô hình dữ liệu hướng đối tượng

a. Khái niệm mô hình dữ liệu hướng đối tượng

Mô hình dữ liệu hướng đối tượng (Object-oriented data model) là một MHDL trong đó dữ liệu được biểu diễn dưới dạng các đối tượng (object), mỗi đối tượng là một thực thể của một lớp (class) cụ thể trong hệ thống. MHDL hướng đối tượng đưa ra một cách tiếp cận khác biệt so với MHDL quan hệ, trong đó các thực thể và mối quan hệ giữa chúng được biểu diễn bằng các đối tượng và phương thức của chúng.



Hình 1- 2: Minh họa mô hình dữ liệu hướng đối tượng.

Một số đặc điểm của MHDL hướng đối tượng bao gồm: Đối tượng (Object), Lớp (Class), Kế thừa (Inheritance), Đa hình (Polymorphism), Trừu tượng (Abstraction).

b. Cấu trúc và thành phần của một mô hình dữ liệu hướng đối tượng

MHDL hướng đối tượng bao gồm các thành phần sau:

Lớp (Class), Đối tượng (Object), Thuộc tính (Attribute/Property), Phương thức (Method), Kế thừa (Inheritance), Đa hình (Polymorphism), Đóng gói (Encapsulation),

c. Mối quan hệ giữa các đối tượng trong mô hình dữ liệu hướng đối tượng

Trong MHDL hướng đối tượng, các đối tượng có thể tương tác với nhau thông qua các mối quan hệ.

- Mối quan hệ một - nhiều (one-to-many)
- Mối quan hệ nhiều - nhiều (many-to-many)
- Mối quan hệ một - một (one-to-one)

d. Các phương thức truy vấn dữ liệu trong mô hình dữ liệu hướng đối tượng

Ngôn ngữ truy vấn đối tượng (Object Query Language - OQL), Ngôn ngữ truy vấn đối tượng dựa trên phương pháp (Method-based Object Query Language - MOQL), Ngôn ngữ truy vấn đối tượng phi cấu trúc (Object Query Language - Unstructured - OQLUS), Ngôn ngữ truy vấn đối tượng lồng nhau (Nested Object Query Language - NOQL), Ngôn ngữ truy vấn đối tượng chủ động (Active Object Query Language - AOQL)

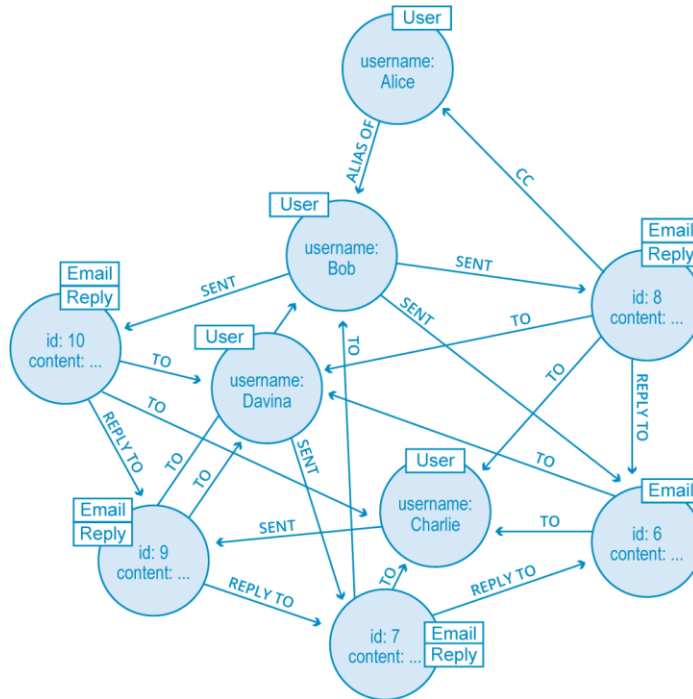
e. Ví dụ minh họa

1.1.3. Mô hình dữ liệu đồ thị.

a. Khái niệm mô hình dữ liệu đồ thị

Mô hình dữ liệu đồ thị là một kiểu MHDL được sử dụng trong khoa học dữ liệu để biểu diễn và mô tả các mối quan hệ giữa các đối tượng thông qua các đỉnh (vertex) và cạnh (edge) trên đồ thị.

Một đồ thị là một tập hợp các đỉnh và cạnh được sắp xếp theo một số quy tắc. Mỗi đỉnh đại diện cho một đối tượng và các cạnh đại diện cho các mối quan hệ giữa các đối tượng đó. Các đối tượng và mối quan hệ có thể được đặc trưng bằng các thuộc tính (property) và trọng số (weight), tùy thuộc vào ứng dụng cụ thể.



Hình 1- 3: Minh họa mô hình dữ liệu đồ thị [12]

b. Cấu trúc và thành phần của một mô hình dữ liệu đồ thị

MHDL đồ thị bao gồm các thành phần sau: Đỉnh (Vertex), Cạnh (Edge), Đồ thị (Graph), Thuộc tính (Attribute), Đường đi (Path), Đồ thị con (Subgraph), Độ dài đường đi (Path length).

c. Các loại đồ thị

MHDL đồ thị có các loại đồ thị sau:

Đồ thị vô hướng (Undirected graph), Đồ thị có hướng (Directed graph), Đồ thị trọng số (Weighted graph), Đồ thị vô hướng có trọng số (Undirected weighted graph). Đồ thị có hướng có trọng số (Directed weighted graph).

d. Các phương thức truy vấn dữ liệu trong mô hình dữ liệu đồ thị

Trong MHDL đồ thị, có các phương thức truy vấn dữ liệu sau:

DFS (Depth First Search), BFS (Breadth First Search), Shortest Path Algorithms, Traversal Algorithms, Matching Algorithms.

e. Ví dụ minh họa

1.1.4. Phân tích, đánh giá các mô hình dữ liệu

Để phân tích và đánh giá một MHDL, ta cần xem xét các yếu tố dưới đây:

Tính nhất quán (Consistency), Hiệu suất (Performance), Khả năng mở rộng (Scalability), Khả năng bảo mật (Security), Tính linh hoạt (Flexibility).

1.1.5. Phương pháp mô hình hóa dữ liệu

Mô hình hóa dữ liệu là quá trình biến đổi dữ liệu từ dạng phi cấu trúc hoặc dạng cấu trúc chưa phù hợp thành các cấu trúc dữ liệu phù hợp với MHDL được chọn.

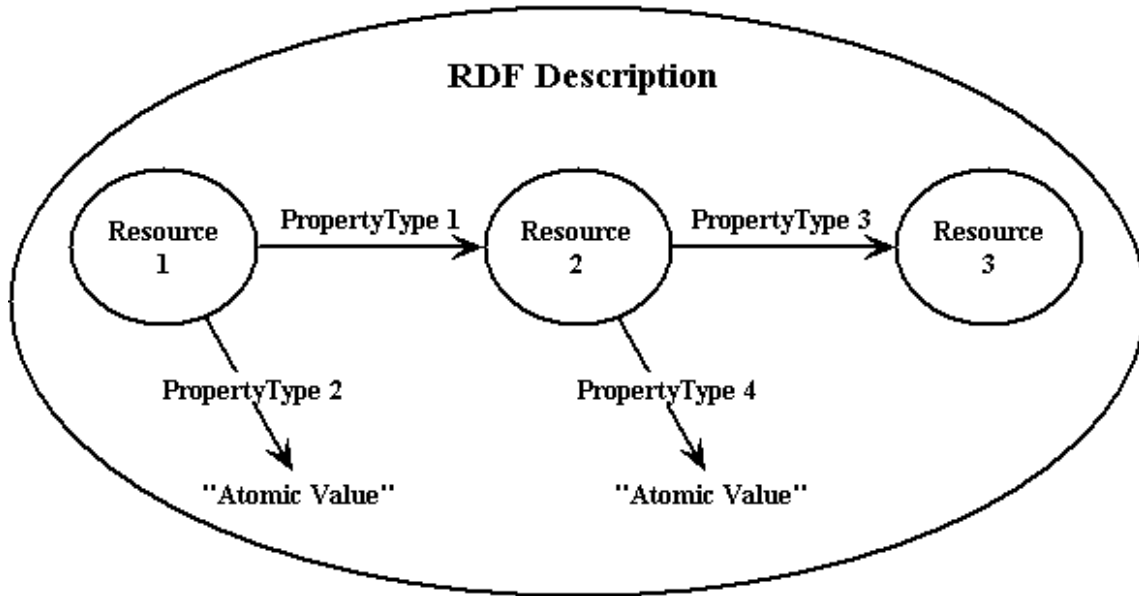
Để thực hiện mô hình hóa dữ liệu, ta có thể thực hiện theo các bước sau:

- Bước 1 - Thu thập và khảo sát dữ liệu: thu thập các nguồn dữ liệu.
- Bước 2 - Chuẩn bị dữ liệu.
- Bước 3 - Thiết kế mô hình dữ liệu.
- Bước 4 - Triển khai mô hình dữ liệu.
- Bước 5 - Kiểm tra và đánh giá mô hình dữ liệu.
- Bước 6 - Duy trì và cập nhật mô hình dữ liệu.

1.1.6. So sánh các mô hình dữ liệu đồ thị: RDF, Knowledge Graph và Property Graph

a. Trình bày về các mô hình dữ liệu đồ thị:

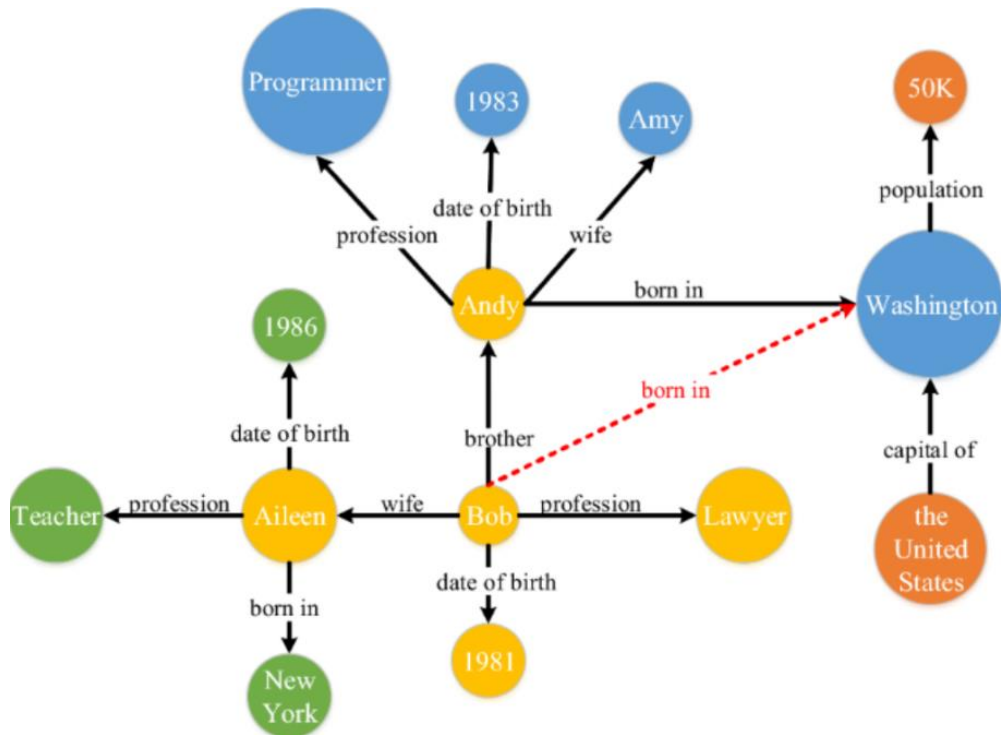
Resource Description Framework (RDF): là một MHDL đồ thị để mô tả thông tin về các tài nguyên trên web



Hình 1- 4: Ví dụ về RDF [13]

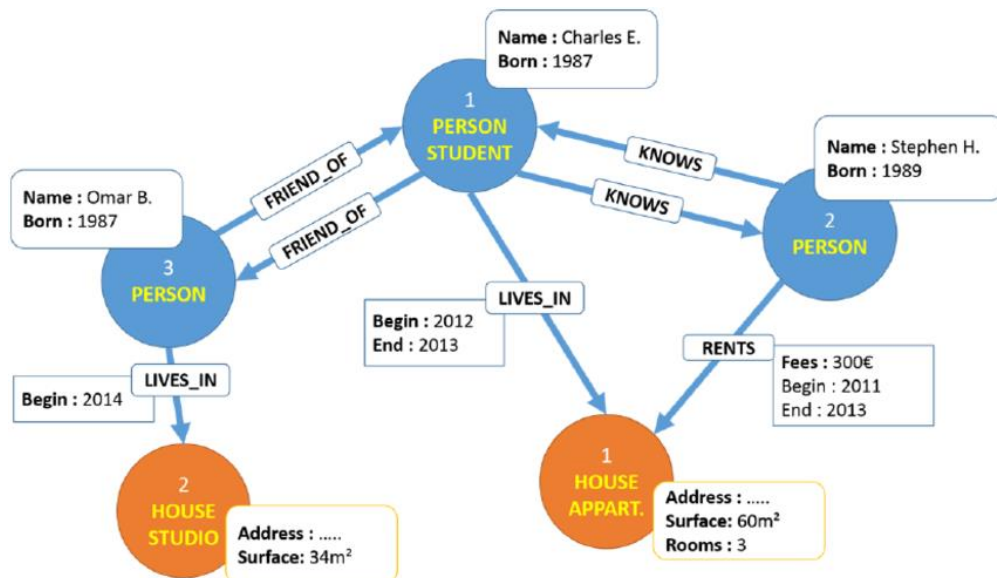
Mỗi tài nguyên được đại diện bằng một URI (Uniform Resource Identifier), và thuộc tính được mô tả bằng các cặp URI và giá trị. Các giá trị có thể là các giá trị nguyên thủy như chuỗi hoặc số, hoặc cũng có thể là các URI của tài nguyên khác.

Knowledge Graph: là một MHDL đồ thị, tập trung vào mô hình hóa tri thức của thế giới thực bằng cách xây dựng các mối quan hệ giữa các thực thể khác nhau. Điều này cho phép mô hình hóa và tương tác với tri thức dưới dạng đồ thị có cấu trúc, giúp cho việc tìm kiếm và truy vấn thông tin trở nên dễ dàng và nhanh chóng hơn.



Hình 1- 5: Ví dụ về Knowledge Graph [14]

Property Graph: là một MHDL đồ thị sử dụng để mô hình hóa các mối quan hệ giữa các đối tượng. Trong Property Graph, các đối tượng được biểu diễn bởi các nút (node), các mối quan hệ giữa các đối tượng được biểu diễn bởi các cạnh (edge), và các thuộc tính được gán cho các đối tượng và các mối quan hệ.



Hình 1- 6: Ví dụ về Property Graph [15]

b. So sánh các mô hình dữ liệu

Tính nhất quán (Consistency): là khả năng giữ cho dữ liệu được đồng bộ và không bị xung đột trong các bảng, đối tượng hoặc đồ thị.

Hiệu suất (Performance): hiệu suất của các MHDL phụ thuộc vào nhiều yếu tố, bao gồm kích thước của dữ liệu, tần suất truy vấn và tính năng của phần mềm hỗ trợ.

Khả năng mở rộng (Scalability): là một trong những yếu tố quan trọng trong việc thiết kế và triển khai các hệ thống dữ liệu.

Khả năng bảo mật (Security): Khả năng bảo mật trong các MHDL quan hệ, hướng đối tượng và đồ thị phụ thuộc vào cách triển khai của chúng.

Tính linh hoạt (Flexibility): là một khía cạnh quan trọng khi đánh giá sự phù hợp của các MHDL cho một ứng dụng cụ thể.

1.1.7. Khoa học dữ liệu đồ thị

a. Khái niệm về khoa học dữ liệu đồ thị

Khoa học dữ liệu đồ thị (Graph Data Science) là một lĩnh vực của khoa học dữ liệu tập trung vào nghiên cứu và phân tích dữ liệu đồ thị. Dữ liệu đồ thị là một loại dữ liệu bao gồm các đỉnh (nodes) và các cạnh (edges) nối giữa các đỉnh, đại diện cho mối quan hệ giữa các đối tượng.

b. Ứng dụng của khoa học dữ liệu đồ thị

Khoa học dữ liệu đồ thị đang được áp dụng rộng rãi trong nhiều lĩnh vực, đặc biệt là trong các lĩnh vực có liên quan đến mạng xã hội, mạng thông tin, y tế, tài chính, v.v. Các ứng dụng của khoa học dữ liệu đồ thị bao gồm: Mạng xã hội, Mạng thông tin, Y tế, Tài chính...

1.2. Các hệ thống quản trị dữ liệu đồ thị

1.2.1. GraphDB

GraphDB là một hệ thống quản trị cơ sở dữ liệu đồ thị được phát triển bởi công ty Ontotext, hỗ trợ lưu trữ, truy vấn và xử lý dữ liệu đồ thị lớn.

Các tính năng của GraphDB bao gồm:

Lưu trữ dữ liệu đồ thị, Các loại câu truy vấn đa dạng, Hỗ trợ các tính năng truy vấn thông minh, Hỗ trợ các tính năng phân tích dữ liệu đồ thị, Hỗ trợ khả năng mở rộng, Hỗ trợ giao diện đồ họa.

Cách sử dụng GraphDB để lưu trữ và truy xuất dữ liệu đồ thị

Một số ứng dụng của GraphDB trong thực tế: Tìm kiếm thông tin khoa học, Quản lý tri thức, Hệ thống thông tin y tế, Hệ thống phân tích dữ liệu, Hệ thống gợi ý,

1.2.2. Neo4j

a. Giới thiệu về Neo4j

Neo4j là một hệ quản trị cơ sở dữ liệu đồ thị (Graph Database) mã nguồn mở được xây dựng trên nền tảng Java, có khả năng lưu trữ, xử lý và truy xuất dữ liệu đồ thị với tốc độ nhanh và hiệu quả cao.

Các tính năng của Neo4j bao gồm:

- Lưu trữ và quản lý dữ liệu đồ thị với khả năng mở rộng lên tới hàng tỉ đỉnh và cạnh.
- Hỗ trợ truy vấn dữ liệu đồ thị bằng ngôn ngữ truy vấn Cypher
- Cung cấp nhiều tính năng độc đáo để truy xuất và xử lý dữ liệu đồ thị
- Hỗ trợ các công nghệ mới như blockchain và machine learning, giúp cải thiện tính bảo mật và khả năng phân tích dữ liệu.
- Cung cấp các giao diện lập trình ứng dụng (API) đa dạng và dễ sử dụng, bao gồm các API dành cho Java, .NET, Python, JavaScript, và nhiều ngôn ngữ khác.

b. Cách sử dụng Neo4J để lưu trữ và truy xuất dữ liệu đồ thị

c. Một số ứng dụng của Neo4J trong thực tế

1.2.3. Jena

a. Giới thiệu về Jena và các tính năng của nó

Apache Jena là một framework mã nguồn mở được sử dụng để phát triển các ứng dụng semantic web, đặc biệt là trong lĩnh vực Linked Data. Jena hỗ trợ việc xử lý RDF (Resource Description Framework), OWL (Web Ontology Language) và các tri thức liên quan.

Các tính năng chính của Jena bao gồm:

- Hỗ trợ các định dạng RDF và OWL

- Các API cho việc tạo, thêm, xóa và truy vấn các triple store.
- Các công cụ truy vấn SPARQL (SPARQL Protocol and RDF Query Language)
- Các API cho việc xây dựng các ứng dụng semantic
- Hỗ trợ các công cụ để xây dựng các ứng dụng web dựa trên dữ liệu đồ thị.
- Các tính năng nâng cao cho phân tích dữ liệu như reasoner và các công cụ hỗ trợ phát hiện tri thức bất đồng nhất.

b. Cách sử dụng Jena để lưu trữ và truy xuất dữ liệu đồ thị

c. Một số ứng dụng của Jena trong thực tế

Linked Data, Phân tích văn bản, Bảo mật và quản lý định danh, Quản lý tri thức, Mạng xã hội

1.2.4. So sánh và đánh giá

GraphDB:

Bảng 1- 1: Ưu điểm và nhược điểm của GraphDB

Ưu điểm	Nhược điểm
<ul style="list-style-type: none"> - Cung cấp tính năng lập chỉ mục dữ liệu và truy vấn đồ thị nhanh chóng. - Có khả năng mở rộng linh hoạt bằng cách thêm các node và cluster mới. - Hỗ trợ RDF và SPARQL, đây là 2 tiêu chuẩn quan trọng trong việc lưu trữ và truy vấn dữ liệu đồ thị. - Có cộng đồng phát triển lớn và hỗ trợ chính thức từ công ty Ontotext. 	<ul style="list-style-type: none"> - Có giá cả đắt đỏ hơn so với các hệ thống khác. - Hạn chế về tính linh hoạt và tùy biến trong việc lập trình ứng dụng.

Neo4j:*Bảng 1- 2: Ưu điểm và nhược điểm của Neo4j*

Ưu điểm	Nhược điểm
<ul style="list-style-type: none"> - Có tính năng lập chỉ mục dữ liệu và truy vấn đồ thị nhanh chóng. - Hỗ trợ truy vấn đồ thị thông qua ngôn ngữ Cypher, cho phép người dùng thao tác với dữ liệu đồ thị một cách dễ dàng. - Có cộng đồng phát triển lớn và hỗ trợ chính thức từ công ty Neo4j. 	<ul style="list-style-type: none"> - Có giá cả đắt đỏ hơn so với các hệ thống khác. - Hạn chế về tính linh hoạt và tùy biến trong việc lập trình ứng dụng.

Jena:*Bảng 1- 3: Ưu điểm và nhược điểm của Jena*

Ưu điểm	Nhược điểm
<ul style="list-style-type: none"> - Là một trong những hệ thống quản trị dữ liệu đồ thị miễn phí và mở mã nguồn. - Hỗ trợ RDF và SPARQL, đây là 2 tiêu chuẩn quan trọng trong việc lưu trữ và truy vấn dữ liệu đồ thị. - Có tính năng linh hoạt trong việc tạo các câu truy vấn SPARQL. 	<ul style="list-style-type: none"> - Có tính năng lập chỉ mục và truy vấn đồ thị chậm hơn so với các hệ thống khác. - Khả năng mở rộng chưa tốt.

1.3. Kết luận chương

Trong chương này, luận văn đã trình bày các cơ sở lý thuyết quan trọng liên quan đến nghiên cứu về mạng thông tin sức khỏe. Đầu tiên, luận văn đã nêu các mô hình dữ liệu quan hệ, hướng đối tượng và đồ thị, và nhận thức được ưu điểm và hạn chế của các CSDL ứng với từng mô hình. Luận văn đã tiến hành so sánh các dạng đồ thị phổ biến và chỉ ra được ưu điểm, nhược điểm của từng dạng. Từ đó có cơ sở để lựa chọn loại cơ sở dữ liệu phù hợp trong bài toán về mạng thông tin sức khỏe.

CHƯƠNG 2. BÀI TOÁN PHÁT TRIỂN MẠNG THÔNG TIN SỨC KHỎE VÀ THIẾT KẾ HỆ THỐNG

2.1. Giới thiệu bài toán mạng thông tin sức khỏe

2.1.1. Bài toán phát triển mạng thông tin sức khỏe

Phạm vi hệ thống: Mạng thông tin sức khỏe nhằm tạo ra một hệ thống quản lý thông tin liên quan đến sức khỏe của các cá nhân và các dịch vụ y tế

Đối tượng tham gia hệ thống: Bệnh nhân, Bác sỹ và nhân viên y tế, Quản trị hệ thống

Quy trình nghiệp vụ của hệ thống: Đăng ký và xác thực người dùng, Quản lý thông tin bệnh nhân, Tra cứu thông tin bệnh nhân, Phân phối công việc.

Yêu cầu cần đạt của hệ thống: Bảo mật, Truy cập dễ dàng, Truy xuất thông tin, Tích hợp hệ thống,

2.1.2. Các chức năng cơ bản của mạng thông tin sức khỏe

Các chức năng cơ bản của mạng thông tin sức khỏe bao gồm:

Thu thập thông tin sức khỏe của người dùng, Chia sẻ thông tin sức khỏe, Cập nhật thông tin sức khỏe, Phân tích thông tin sức khỏe, Tra cứu thông tin sức khỏe,

Tính năng phân phối công việc sẽ giúp phân phối các trường hợp bệnh án cho các bác sỹ dựa trên chuyên môn, kinh nghiệm, thời gian rảnh của họ và nhiều yếu tố khác. Hệ thống sẽ tự động tìm kiếm bác sỹ phù hợp nhất để giải quyết các trường hợp bệnh án, giúp tối ưu hóa công việc của bác sỹ và đảm bảo chất lượng chăm sóc cho bệnh nhân.

2.1.3. Các hướng tiếp cận và giải quyết bài toán

Các hướng tiếp cận và giải quyết bài toán sử dụng các mô hình khác nhau như:

- Mô hình quan hệ
- Mô hình đồ thị
- Mô hình hướng đối tượng

Các giải pháp đang được triển khai thực tế trong việc lập mạng thông tin sức khỏe:

Hệ thống quản lý thông tin bệnh án điện tử, Mạng xã hội y tế, Các ứng dụng di động liên quan đến sức khỏe, Công nghệ truyền thông và thông tin y tế.

Các tiêu chí để so sánh, đánh giá các mô hình

Tính mở rộng, Tính linh hoạt, Tính tương thích, Tính dễ sử dụng, Hiệu năng, Tính bảo mật, Tính tương tác

So sánh 3 mô hình về tính mở rộng

So sánh 3 mô hình về tính linh hoạt

So sánh 3 mô hình về tính tương thích, tính dễ sử dụng

So sánh 3 mô hình về hiệu suất

So sánh 3 mô hình về tính bảo mật

So sánh 3 mô hình dữ liệu về tính tương tác

2.1.4. Đề xuất giải pháp giải quyết

Với cách tiếp cận này, ta sẽ sử dụng mô hình đồ thị để biểu diễn các quan hệ giữa các đối tượng trong hệ thống sức khỏe, từ đó xây dựng một mạng thông tin sức khỏe hoàn chỉnh và hiệu quả.

Để đưa ra giải pháp tối ưu, cần xem xét các yếu tố sau:

- Phù hợp với yêu cầu của bài toán
- Sử dụng mô hình đồ thị
- Sử dụng công nghệ mới nhất
- Thiết kế hệ thống tối ưu
- Tối ưu hóa quá trình truy vấn

Dựa trên ưu và nhược điểm của các mô hình quan hệ, hướng đối tượng và đồ thị, cùng với yêu cầu của bài toán lập mạng thông tin sức khỏe, mô hình đồ thị được đưa ra là mô hình tối ưu.

2.2. Áp dụng phát triển mạng thông tin sức khỏe

2.2.1. Các bước xây dựng hệ thống dữ liệu theo mô hình dữ liệu đồ thị

Để xây dựng hệ thống dữ liệu theo MHDLD đồ thị cho mạng thông tin sức khỏe, các bước chính bao gồm:

- Thiết kế schema
- Triển khai cơ sở dữ liệu
- Nhập dữ liệu
- Tối ưu hóa và tinh chỉnh
- Phát triển ứng dụng web

2.2.2. Thiết kế CSDL đồ thị trên Neo4j

a. Xác định các đối tượng và mối quan hệ giữa chúng trong mạng thông tin sức khỏe

Mạng thông tin sức khỏe có các chức năng cơ bản đã đặt ra ở mục 2.1.1. Bước đầu mạng thông tin sẽ có chức năng “Thu thập thông tin sức khỏe”, chức năng “Chia sẻ thông tin sức khỏe”. Vì vậy luận văn đặt ra thiết kế các đối tượng trong mạng như sau:

STT	Đối tượng	Mô tả
1	AdminHeThong	AdminHeThong có quyền điều khiển toàn bộ hệ thống và quản lý các tài khoản đăng nhập vào hệ thống.
2	AdminBenhVien	AdminBenhVien quản lý hoạt động của các bác sỹ và phòng khám trong bệnh viện.
3	BacSy	BacSy là người chăm sóc sức khỏe cho bệnh nhân, có liên quan đến KhoaKhamBenh và DichVuKhamBenh.
4	BenhNhan	BenhNhan là người bệnh được khám và chữa trị tại bệnh viện, có thông tin HoSoBenhAn được tạo ra khi bệnh nhân đăng ký khám bệnh.
5	BenhVien	BenhVien là nơi cung cấp dịch vụ y tế, có mối quan hệ với các đối tượng khác trong hệ thống.
6	ChuyenMon	ChuyenMon là các chuyên môn y khoa khác nhau có liên quan đến các KhoaKhamBenh.
7	DichVuKhamBenh	DichVuKhamBenh là các dịch vụ khám bệnh mà bệnh viện cung cấp cho bệnh nhân, có liên quan đến KhoaKhamBenh và BacSy.
8	HoSoBenhAn	HoSoBenhAn là tài liệu lưu trữ thông tin về sức khỏe của bệnh nhân.

9	KhoaKhamBenh	KhoaKhamBenh là các khoa chuyên môn trong bệnh viện, có liên quan đến các ChuyenMon, DichVuKhamBenh và BacSy.
10	LichSuCongViec	LichSuCongViec là lịch sử làm việc của các nhân viên trong bệnh viện, bao gồm cả BacSy và AdminBenhVien.
11	TaiKhoan	TaiKhoan là thông tin đăng nhập của các đối tượng trong hệ thống, được quản lý bởi AdminHeThong.
12	TinhHuyenXa	TinhHuyenXa là các đơn vị hành chính trong địa phương, có thể được liên kết đến địa điểm của BenhVien.
13	BaseModel	Đối tượng dùng chung chứa các thuộc tính cơ bản
14	ThongTinNguoiDung	ThongTinNguoiDung là thông tin cá nhân của các đối tượng trong hệ thống.

Bảng 2- 1: Các đối tượng trong mạng thông tin sức khỏe

b. Thiết kế schema cho CSDL đồ thị.

Dựa trên các đối tượng và mối quan hệ giữa chúng, ta có thể thiết kế schema cho CSDL đồ thị như sau:

Node: AdminHeThong, AdminBenhVien, BacSy, BenhNhan, BenhVien, ChuyenMon, DichVuKhamBenh, HoSoBenhAn, KhoaKhamBenh, LichSuCongViec, TaiKhoan, TinhHuyenXa.

Relationship:

- CO_THONG_TIN: quan hệ giữa các node "AdminHeThong", "AdminBenhVien", "BacSy", "BenhNhan", "TaiKhoan" với node "ThongTinNguoiDung"
- CO_BASE: quan hệ giữa các node "AdminHeThong", "AdminBenhVien", "BacSy", "BenhNhan", "BenhVien" và "TaiKhoan" với node "BaseModel"
- LAM_VIEC_TAI: quan hệ giữa các node "BacSy", "DichVuKhamBenh" và "KhoaKhamBenh" với node "BenhVien"
- CO_CHUYEN_MON: quan hệ giữa các node "BacSy" với node "ChuyenMon"
- CO_DIA_CHI: quan hệ giữa các node "BenhVien" và "TinhHuyenXa"
- DUOC_KHAM_BENH: quan hệ giữa các node "BenhNhan", "DichVuKhamBenh" và "BacSy"

- CO_HO_SO: quan hệ giữa các node "BenhNhan" và "HoSoBenhAn"
- CO_CONG_VIEC: quan hệ giữa các node "BacSy" và "LichSuCongViec"

c. Tạo các nodes và relationships trong CSDL đồ thị.

d. Tối ưu hóa CSDL đồ thị để đạt hiệu suất cao nhất.

2.3. Xây dựng hệ thống phần mềm truy vấn và trực quan hóa

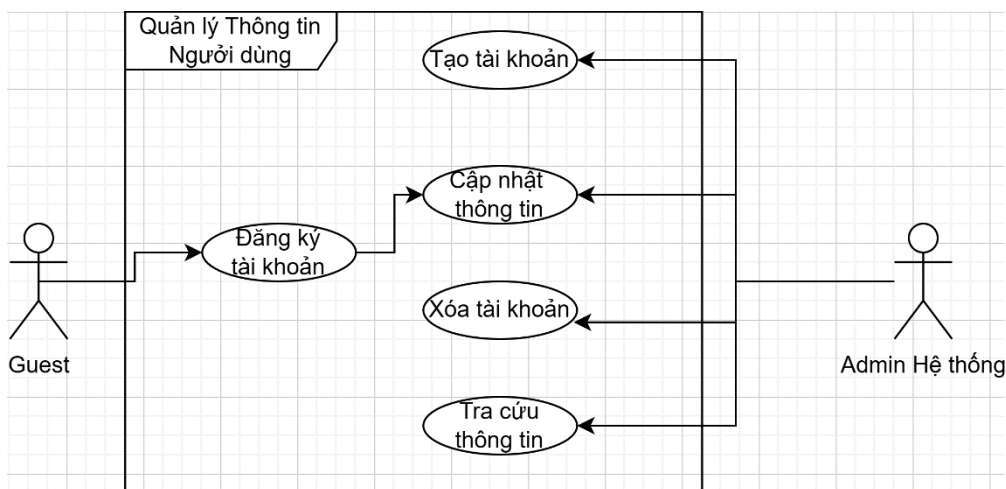
2.3.1. Lập Use case.

Để xây dựng hệ thống phần mềm truy vấn và trực quan hóa cho mạng thông tin sức khỏe, luận văn đưa ra một số use case sau:

a. Usecase: Quản lý thông tin người dùng:

Mô tả: Use case này mô tả việc quản lý thông tin người dùng trong hệ thống, bao gồm tạo mới người dùng, cập nhật thông tin người dùng, xóa người dùng khỏi hệ thống và xem danh sách người dùng trong hệ thống.

Người sử dụng chính: Admin hệ thống, Guest

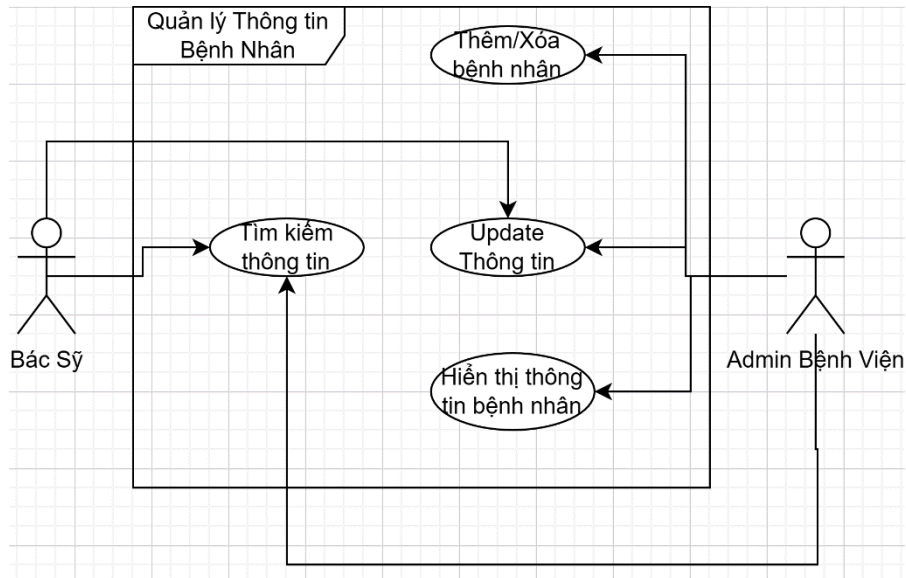


Hình 2- 1: Usecase Quản lý Thông tin Người dùng

b. Usecase: Quản lý thông tin bệnh nhân:

Mô tả: Use case này cho phép admin bệnh viện, bác sỹ có thể quản lý thông tin của các bệnh nhân trong hệ thống của 1 bệnh viện.

Người tương tác: admin bệnh viện, bác sỹ

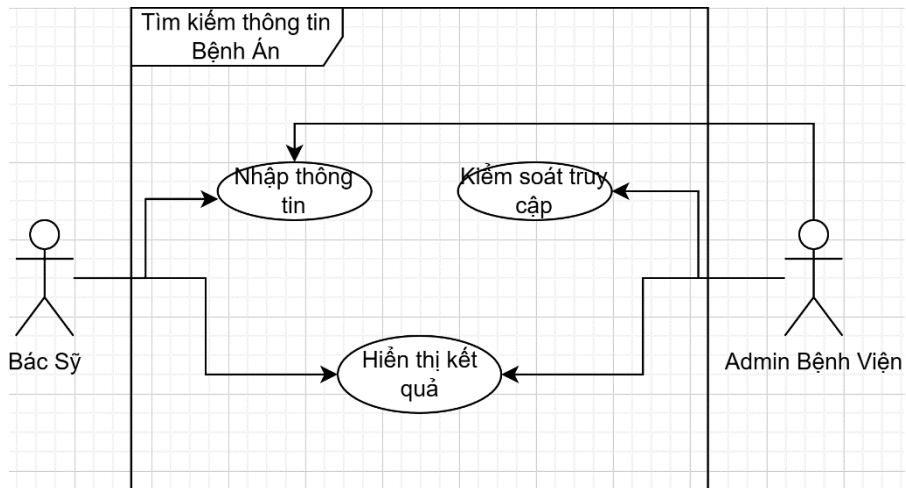


Hình 2- 2: Usecase Quản lý thông tin Bệnh nhân

c. Usecase: Tìm kiếm thông tin bệnh án của bệnh nhân

Mô tả: Use case này cho phép người dùng tìm kiếm thông tin về bệnh án của một bệnh nhân trong hệ thống.

Người dùng: Bác sĩ, Admin bệnh viện.

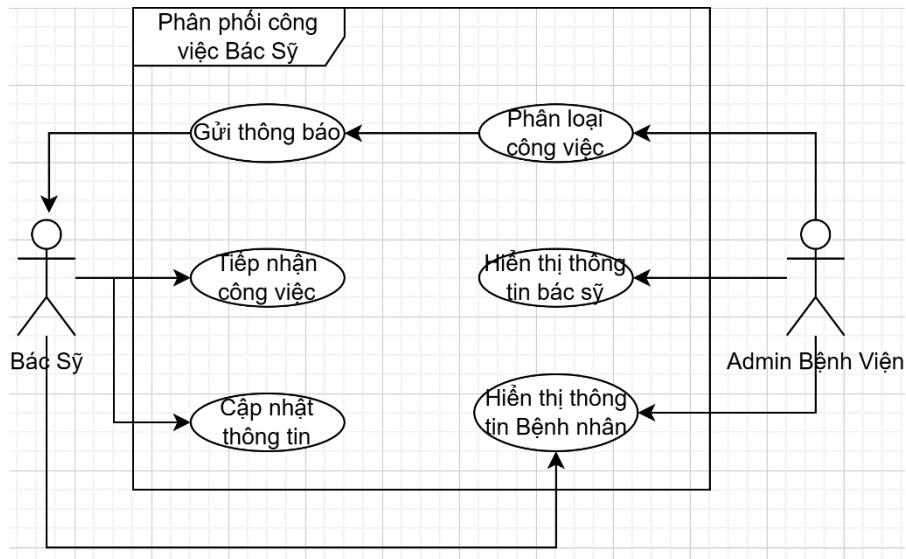


Hình 2- 3: Usecase Tìm kiếm thông tin bệnh án

d. Usecase: Phân phối công việc (cho Bác sĩ)

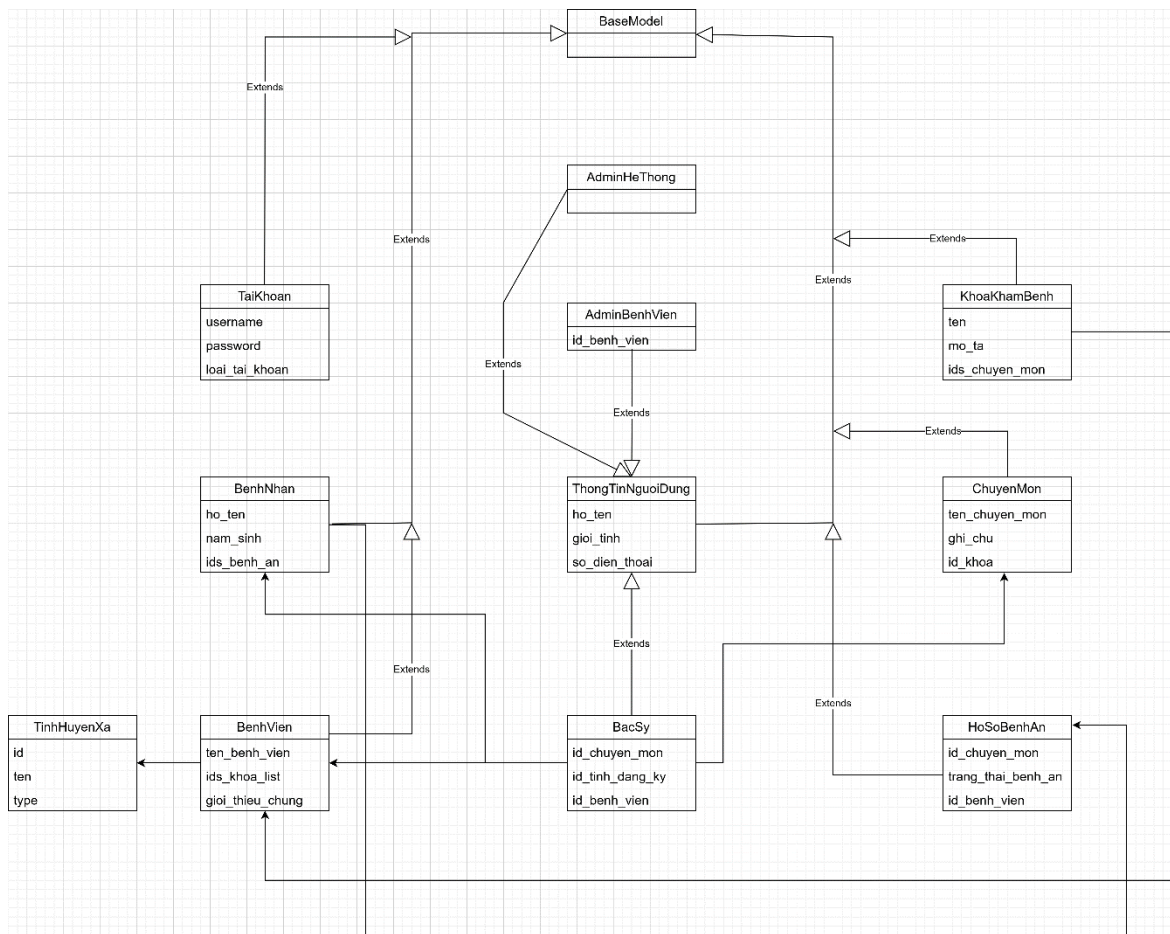
Mô tả: Use case này mô tả quá trình phân phối công việc cho các bác sĩ trong bệnh viện.

Người tương tác: Admin bệnh viện, bác sĩ



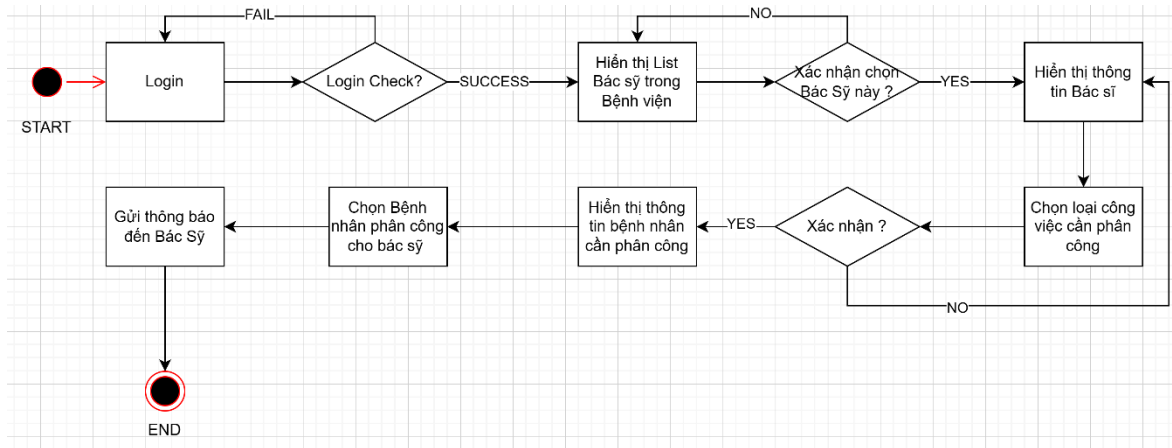
Hình 2- 4: Usecase Phân phối công việc cho Bác sỹ

2.3.2. Lập biểu đồ lớp.



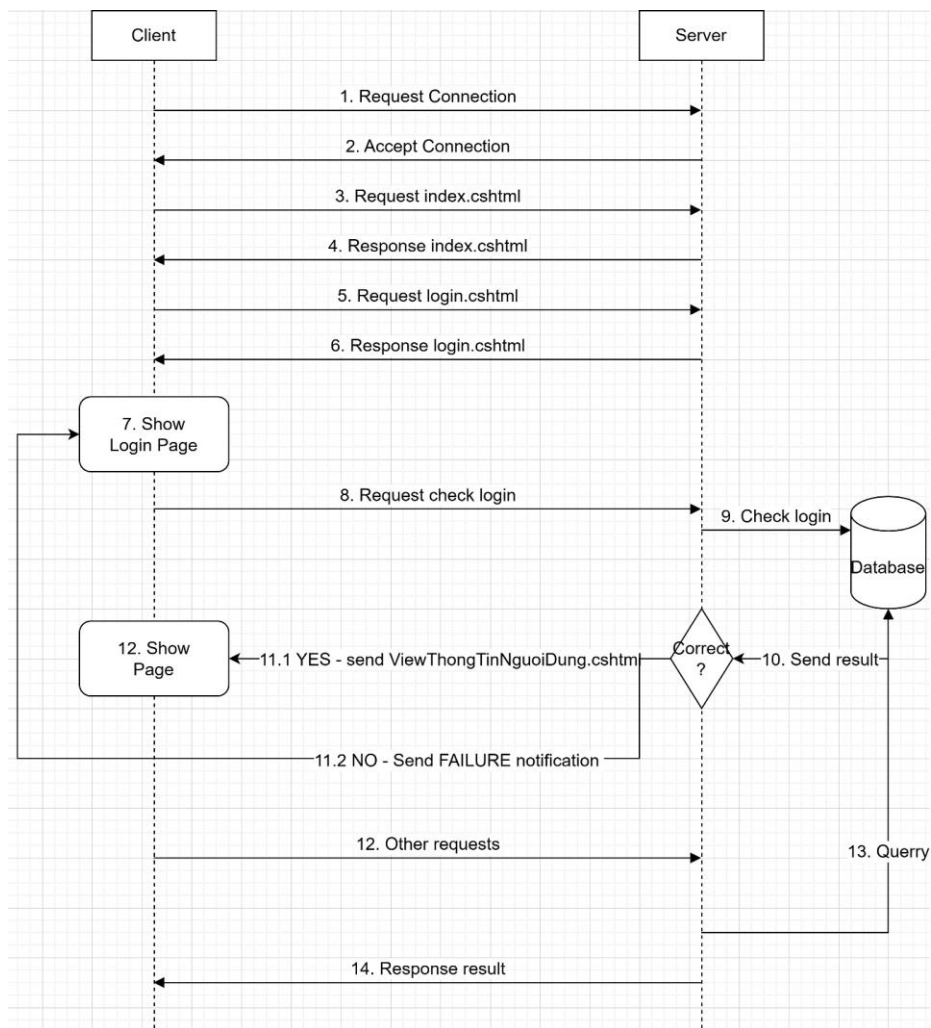
Hình 2- 5: Biểu đồ lớp

2.3.3. Lập biểu đồ hoạt động.



Hình 2- 6: Biểu đồ hoạt động Phân công công việc cho Bác sỹ

2.3.4. Lập biểu đồ tuần tự.



Hình 2- 7: Biểu đồ tuần tự quá trình Login

2.4. Kết luận chương

Chương 2 tập trung vào bài toán phát triển mạng thông tin sức khỏe và thiết kế hệ thống cho bài toán này. Luận văn đã trình bày về bài toán phát triển mạng thông tin sức khỏe sẽ giải quyết trong luận văn, các chức năng của nó, và quy trình nghiệp vụ của hệ thống. Luận văn cũng đã tìm hiểu về việc xây dựng hệ thống phần mềm truy vấn và trực quan hóa để hỗ trợ việc quản lý và tìm kiếm thông tin sức khỏe hiệu quả.

CHƯƠNG 3. KẾT QUẢ VÀ ĐÁNH GIÁ

3.1. Thu thập và chia sẻ thông tin sức khỏe

Thu thập thông tin sức khỏe

Chia sẻ thông tin

3.2. Kịch bản kiểm thử hệ thống

Trong phần này, luận văn trình bày chi tiết về kịch bản kiểm thử được thiết kế để đánh giá hệ thống mạng thông tin sức khỏe theo các yêu cầu quan trọng như bảo mật, truy cập dễ dàng, truy xuất thông tin và tích hợp hệ thống. Dưới đây là một mô tả chi tiết về các bước kiểm thử và hoạt động tương ứng của từng yêu cầu:

- Kiểm thử bảo mật
- Kiểm thử truy cập dễ dàng
- Kiểm thử truy xuất thông tin:
- Kiểm thử tích hợp hệ thống:

3.3. Triển khai cài đặt và kết quả

Để triển khai dự án mạng thông tin sức khỏe, dự án sử dụng Neo4j làm cơ sở dữ liệu đồ thị và C# để phát triển website quản lý.

3.4. Lập trình các module

Sau khi đã triển khai cài đặt CSDL đồ thị trên Neo4j, tiếp theo là lập trình các module để cung cấp các chức năng cho website mạng thông tin sức khỏe.

Dưới đây là các module:

VolumnSystemBL (Business Logic), VolumnSystemDTO (Data Transfer Object), VolumnSystemModel, VolumnSystemRepository, VolumnSystemUtils (Utilities), VolumnSystemWeb.

3.5. Thực nghiệm hệ thống

Kết quả chạy thử nghiệm

a. Đăng nhập hệ thống

BenhVienTot.com

Đăng nhập vào hệ thống

Tên đăng nhập

* abcd...

Mật khẩu

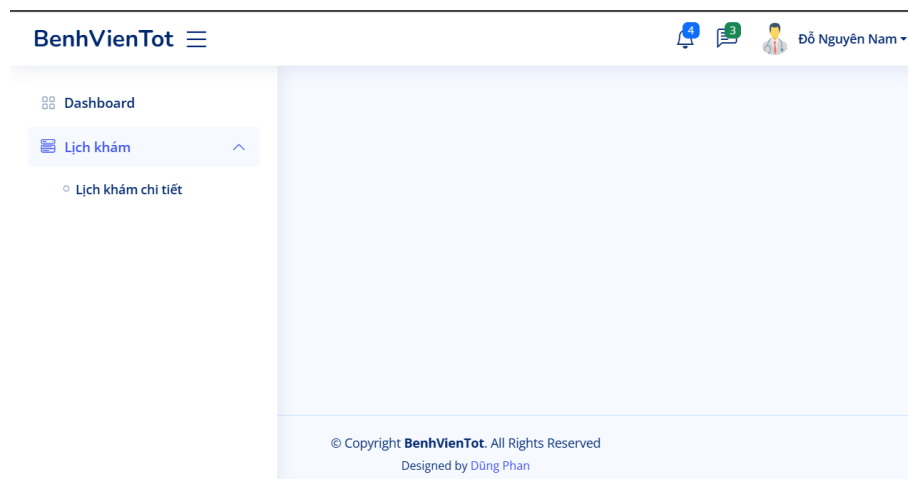
Đăng nhập

Hoặc [tạo tên đăng nhập mới](#)

Powered by [Dũng Phan](#)

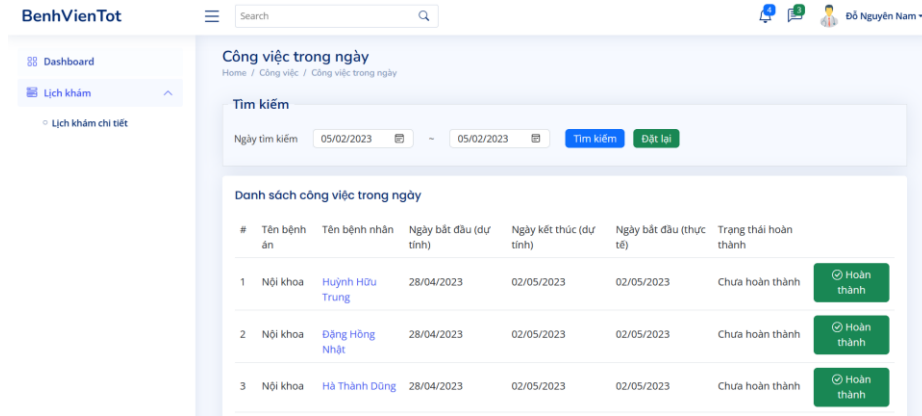
Hình 3- 1: Màn hình đăng nhập hệ thống

b. Công việc bác sỹ



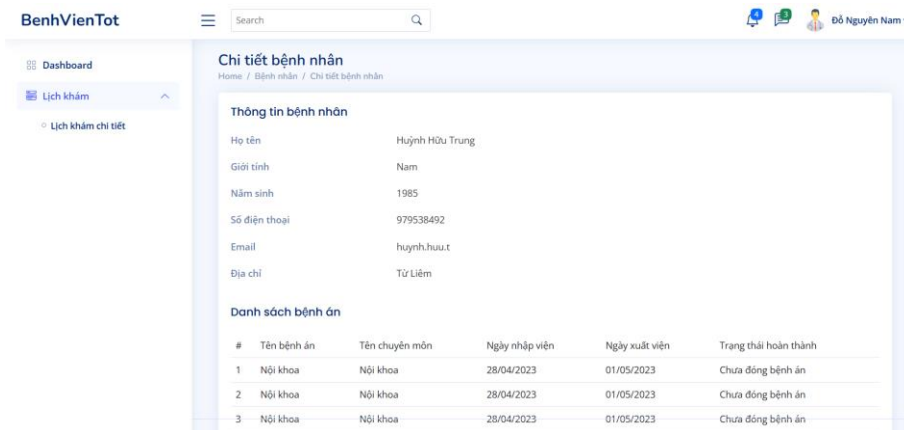
Hình 3- 2: Màn hình công việc bác sỹ

c. Lịch khám chi tiết



Hình 3- 3: Màn hình lịch khám chi tiết (cho Bác sỹ)

d. Chi tiết thông tin bệnh nhân



Hình 3- 4: Màn hình chi tiết thông tin bệnh nhân

3.6. Đánh giá kết quả hệ thống

Kết quả hệ thống ứng với các kịch bản kiểm thử

- Đánh giá bảo mật
- Đánh giá giao diện người dùng:
- Đánh giá truy xuất thông tin:
- Đánh giá tích hợp hệ thống

3.7. Kết luận chương

Trong chương 3, luận văn đã tiến hành trình bày các cách thức thu thập và chia sẻ thông tin sức khỏe. Thông tin này được sử dụng trong luận văn đảm bảo các yêu cầu ban đầu về dữ liệu. Luận văn đã đưa ra kịch bản kiểm thử hệ thống để đánh giá sơ bộ hiệu quả, khả

năng vận hành của hệ thống có đáp ứng với các yêu cầu đề ra ban đầu hay không. Sau đó luận văn đã tiến hành lập trình các module và thực nghiệm hệ thống.

Các chức năng đã được triển khai đáp ứng được nhu cầu quản lý và cung cấp thông tin liên quan đến sức khỏe cho người dùng, đặc biệt là cho bác sỹ và các cơ sở y tế. Hệ thống cho phép quản lý các thông tin liên quan đến bệnh nhân, bác sỹ, cơ sở y tế, kết quả khám bệnh, chẩn đoán, hồ sơ bệnh án và các thông tin khác, giúp cho việc theo dõi tình trạng sức khỏe của bệnh nhân và đưa ra các phương án điều trị phù hợp.

Tuy nhiên, vẫn còn một số hạn chế trong quá trình triển khai và sử dụng hệ thống, như khả năng tương thích với các hệ thống khác, độ tin cậy của dữ liệu và hiệu suất khi có số lượng lớn người dùng truy cập cùng lúc. Để cải thiện và nâng cao hiệu quả của hệ thống, cần tiếp tục nghiên cứu và phát triển các tính năng mới, cải tiến các khâu quản lý và bảo mật dữ liệu, đồng thời tối ưu hóa hệ thống để đáp ứng nhu cầu của người dùng trong thời gian ngắn nhất.

KẾT LUẬN

Trong đề tài này, luận văn đã thực hiện xây dựng một hệ thống mạng thông tin sức khỏe dựa trên CSDL đồ thị Neo4j và ngôn ngữ lập trình C#. Hệ thống này được phát triển để giải quyết các vấn đề trong việc quản lý thông tin sức khỏe và phân phối công việc trong lĩnh vực y tế. Luận văn đã thiết kế một cấu trúc CSDL đồ thị phù hợp cho việc lưu trữ thông tin sức khỏe, bao gồm các đối tượng như bệnh nhân, bác sỹ, thuốc và bệnh tật. Đồng thời, luận văn đã phát triển các chức năng để thêm, sửa đổi, xóa thông tin trên CSDL đồ thị.

Hệ thống cũng được thiết kế với các tính năng quản lý người dùng và phân phối công việc cho bác sỹ. Các chức năng này bao gồm quản lý tài khoản người dùng, phân công công việc và theo dõi tiến độ công việc.

Để triển khai hệ thống, luận văn sử dụng Neo4j làm CSDL đồ thị và lập trình bằng ngôn ngữ C#. Kết quả đạt được là hệ thống có thể quản lý được thông tin sức khỏe và phân phối công việc cho bác sỹ.

Tuy nhiên, hệ thống vẫn còn một số hạn chế như chưa hỗ trợ tính năng đa ngôn ngữ, chưa có tính năng tìm kiếm thông tin bệnh nhân theo nhiều tiêu chí khác nhau. Để cải thiện hệ thống, luận văn sẽ tiếp tục phát triển các tính năng mới và tối ưu hóa hiệu suất hệ thống.

Tổng kết lại, đề tài này đã đạt được mục tiêu ban đầu trong việc xây dựng hệ thống mạng thông tin sức khỏe để quản lý thông tin sức khỏe và phân phối công việc trong lĩnh vực y tế. Hệ thống này có tiềm năng để phát triển và ứng dụng rộng rãi trong thực tế.

TÀI LIỆU THAM KHẢO

Tài liệu tham khảo Tiếng Việt:

- [1] Bộ Y tế (2011), *Kỷ yếu Hội thảo quốc tế về chuẩn hóa hệ thống thông tin y tế*, Quảng Nam.
- [2] Bộ Y tế (2008), *Báo cáo về “Số lượng người điều trị ngoại chấn 1 năm”*.
- [3] Cục Quản lý Khám, chữa bệnh - Bộ Y tế (2008), *Số liệu kiểm tra 932 bệnh viện năm 2008*.
- [4] Trần Xuân Chức, Trần Văn Tuyên, Hoàng Văn Tiến, Nguyễn Sơn Hải, Trần Thị Diệu Trinh (2014), “Giải pháp ứng dụng thu thập và cung cấp thông tin y tế chủ động tới cộng đồng”, Hội thi Tin học khối cán bộ, công chức trẻ toàn quốc lần II-2014.
- [5] Lê Hồng Hà, Trần Xuân Chức, Kiều Mai (2015), “Y tế di động và triển vọng phát triển tại Việt Nam”, *Kỷ yếu Hội nghị ứng dụng Công nghệ thông tin trong ngành y tế lần thứ 7*, Bộ Y tế, tr. 114 - 116.
- [6] Nguyễn Huy Khánh, (2012). *Các công nghệ lập trình hiện đại*, Đại học Khoa học Tự nhiên TP. Hồ Chí Minh.
- [7] Nguyễn Hoàng Phương, Phí Văn Thâm, Nguyễn Tuấn Khoa (2008), *Kỷ yếu hội thảo khoa học: Ứng dụng Công nghệ thông tin trong quản lý bệnh viện*, Trung tâm tin học, Bộ Y tế.
- [8] Nguyễn Đức Thuận, Vũ Duy Hải, Trần Anh Vũ (2006), *Hệ thống thông tin y tế*, Nhà xuất bản Bách khoa Hà Nội.
- [9] WB, WHO, UNICEF, JICA (2013), *Bao phủ chăm sóc sức khỏe toàn dân: Việt Nam – Các sáng kiến quốc gia, thách thức và vai trò của hoạt động hợp tác quốc tế*.
- [10] WB (2014), *Phát triển kỹ năng: Xây dựng lực lượng lao động cho một nền kinh tế thị trường hiện đại ở Việt Nam*.

Tài liệu tham khảo Tiếng Anh:

- [11] Neha T, (2019 November), *Relational Data Model* [Online]. Available: <https://binaryterms.com/relational-data-model.html>

- [12] Bryce Merkl Sasaki, (2018 Jul), Graph Databases for Beginners: Data Modeling Pitfalls to Avoid [Online], Available: <https://neo4j.com/blog/data-modeling-pitfalls/>
- [13] Atakan Güney, (2019 Oct), Introduction to Resource Description Framework and SPARQL (RDF 101) [Online]. Available: <https://medium.com/@atakanguney94/introduction-to-resource-description-framework-and-sparql-rdf-101-5857f4a6a8a6>
- [14] Ajitesh Kumar, (2022 Augst), Knowledge Graph Concepts & Machine Learning: Examples [Online]. Available: <https://vitalflux.com/knowledge-graph-concepts-machine-learning-examples/>
- [15] Raouf Bouhali, Anne Laurent (2015), Exploiting RDF Open Data Using NoSQL Graph Databases, University of Montpellier
- [16] Alberz Akrawi (2010), Social Network System, SE-100 44 Stockholm, Sweden.
- [17] Alexander Richter, Michael Koch (2014), Functions of Social Networking Services.
- [18] Gunther Eysenbach, MD, MPH (2008), Medicine 2.0 Proceedings, Toronto, Canada.
- [19] Honigman (2010), Friending, Tweeting, Blogging & LinkingIn: Legal Considerations for Health Care Providers.
- [20] Ian Robinson, Jim Webber, and Emil Eifrem (2013), Graph Databases, O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472.
- [21] Ian Robinson (2010), RESTful Application Protocols.
- [22] IDG Communication, Inc., (2018), *State of Digital Business Transformation*.
- [23] Marcos Boyington, Po Chen, Grace Kum, Van Le-Pham, Eric Morales, Jake Warmerdam, Cheuk (Anna) Yu, Jingren Zhou (2004), Design Specifications for Social Networking System.
- [24] Mark Needham, Amy E. Hodler, (2019), Graph Algorithms Practical Examples in Apache Spark & Neo4j, O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472.
- [25] The Neo4j Team neo4j.org (2013), *The Neo4j Manual*.