

# Tìm hiểu và nghiên cứu về ontology và ứng dụng trong công nghệ tri thức

## I. Khái niệm về ontology

### 1. Khái niệm

Trong khoa học máy tính, một ontology là một mô hình dữ liệu biểu diễn một lĩnh vực và được sử dụng để suy luận về các đối tượng trong lĩnh vực đó và mối quan hệ giữa chúng. Ontology cung cấp một bộ từ vựng chung bao gồm các khái niệm, các thuộc tính quan trọng và các định nghĩa về các khái niệm và các thuộc tính này. Ngoài bộ từ vựng, ontology còn cung cấp các ràng buộc, đôi khi các ràng buộc này được coi như các giả định cơ sở về ý nghĩa mong muốn của bộ từ vựng, nó được sử dụng trong một miền mà có thể được giao tiếp giữa người và các hệ thống ứng dụng phân tán hỗn tạp khác.

Các ontology được sử dụng như là một biểu mẫu trình bày tri thức về thế giới hay một phần của nó. Các ontology thường miêu tả:

- \* Các cá thể: Các đối tượng cơ bản, nền tảng
- \* Các lớp: Các tập hợp, hay kiểu của các đối tượng
- \* Các thuộc tính: Thuộc tính, tính năng, đặc điểm, tính cách, hay các thông số mà các đối tượng có và có thể đem ra chia sẻ.
- \* Các mối liên hệ: Các con đường mà các đối tượng có thể liên hệ tới một đối tượng khác.

Bộ từ vựng ontology được xây dựng trên cơ sở tầng RDF và RDFS, cung cấp khả năng biểu diễn ngữ nghĩa mềm dẻo cho tài nguyên Web và có khả năng hỗ trợ lập luận.

### 2. Các phần tử trong ontolog

#### 2.1 Các cá thể (Individuals) - Thể hiện

Các cá thể là các thành phần cơ bản, nền tảng của một ontology. Các cá thể trong một ontology có thể bao gồm các đối tượng cụ thể như con người, động vật, cái bàn... cũng như các cá thể trừu tượng như các thành viên hay các từ. Một ontology có thể không cần bất kỳ một cá thể nào, nhưng một trong những lý do chính của một ontology là để cung cấp một ngữ nghĩa của việc phân lớp các cá thể, mặc dù các cá thể này không thực sự là một phần của ontology.

#### 2.2 Các lớp (Classes) - Khái niệm

Các lớp là các nhóm, tập hợp các đối tượng trừu tượng. Chúng có thể chứa các cá thể, các lớp khác, hay là sự phối hợp của cả hai.

Các ontology biến đổi tùy thuộc vào cấu trúc và nội dung của nó: Một lớp có thể chứa các lớp con, có thể là một lớp tổng quan (chứa tất cả mọi thứ), có thể là lớp chỉ chứa những cá thể riêng lẻ, Một lớp có thể xếp gộp vào hoặc bị xếp gộp vào bởi các lớp khác. Mối quan hệ xếp gộp này được sử dụng để tạo ra một cấu trúc có thứ bậc các lớp, thường là với một lớp thông dụng nhất kiểu Thing ở trên đỉnh và các lớp rất rõ ràng kiểu 2002, Ford ở phía dưới cùng.

### ***2.3 Các thuộc tính (Properties)***

Các đối tượng trong ontology có thể được mô tả thông qua việc khai báo các thuộc tính của chúng. Mỗi một thuộc tính đều có tên và giá trị của thuộc tính đó. Các thuộc tính được sử dụng để lưu trữ các thông tin mà đối tượng có thể có. Ví dụ, đối với một cá nhân có thể có các thuộc tính: Họ\_tên, ngày\_sinh, quê\_quán, số\_cmnd... Giá trị của một thuộc tính có thể có các kiểu dữ liệu phức tạp.

### ***2.4 Các mối quan hệ (Relation)***

Một trong những ứng dụng quan trọng của việc sử dụng các thuộc tính là để mô tả mối liên hệ giữa các đối tượng trong ontology. Một mối quan hệ là một thuộc tính có giá trị là một đối tượng nào đó trong ontology.

Một kiểu quan hệ quan trọng là kiểu quan hệ xếp gộp (subsumption). Kiểu quan hệ này mô tả các đối tượng nào là các thành viên của các lớp nào của các đối tượng. Hiện tại, việc kết hợp các ontology là một tiến trình được làm phần lớn là thủ công, do vậy rất tốn thời gian và đắt đỏ. Việc sử dụng các ontology là cơ sở để cung cấp một định nghĩa thông dụng của các thuật ngữ cốt lõi có thể làm cho tiến trình này trở nên dễ quản lý hơn. Hiện đang có các nghiên cứu dựa trên các kỹ thuật sản sinh để nối kết các ontology, tuy nhiên lĩnh vực này mới chỉ hiện hữu về mặt lý thuyết.

## **3. Ngôn ngữ OWL**

OWL (The Web Ontology Language) là một ngôn ngữ gần như XML dùng để mô tả các hệ cơ sở tri thức. OWL là một ngôn ngữ đánh dấu dùng để xuất bản và chia sẻ dữ liệu trên Internet thông qua những mô hình dữ liệu gọi là “ontology”. Ontology mô tả một lĩnh vực (domain) và diễn tả những đối tượng trong lĩnh vực đó cùng những mối quan hệ giữa các đối tượng này. OWL là phần mở rộng về từ vựng của RDF và được kế thừa từ ngôn ngữ DAML+OIL Web ontology – một dự án được hỗ trợ bởi W3C. OWL biểu diễn ý nghĩa của các thuật ngữ trong các từ vựng và mối liên hệ giữa các thuật ngữ này để đảm bảo phù hợp với quá trình xử lý bởi các phần mềm. OWL được xem như là một kỹ thuật trọng yếu để cài đặt cho Semantic Web trong tương lai. OWL được thiết kế đặc biệt để cung cấp một cách thức thông dụng trong việc xử lý nội dung thông tin của

Web. Ngôn ngữ này được kỳ vọng rằng sẽ cho phép các hệ thống máy tính có thể đọc được thay thế cho con người. Vì OWL được viết bởi XML, các thông tin OWL có thể dễ dàng trao đổi giữa các kiểu hệ thống máy tính khác nhau, sử dụng các hệ điều hành và các ngôn ngữ ứng dụng khác nhau. Mục đích chính của OWL là sẽ cung cấp các chuẩn để tạo ra một nền tảng để quản lý tài sản, tích hợp mức doanh nghiệp và để chia sẻ cũng như tái sử dụng dữ liệu trên Web. OWL được phát triển bởi nó có nhiều tiện lợi để biểu diễn ý nghĩa và ngữ nghĩa hơn so với XML, RDF và RDFS, và vì OWL ra đời sau các ngôn ngữ này, nó có khả năng biểu diễn các nội dung mà máy có thể biểu diễn được trên Web.

#### **4. Các phiên bản của OWL**

Hiện nay có ba loại OWL : OWL Lite, OWL DL (description logic), và OWL Full.

OWL Lite: hỗ trợ cho những người dùng chủ yếu cần sự phân lớp theo thứ bậc và các ràng buộc đơn giản. Ví dụ: Trong khi nó hỗ trợ các ràng buộc về tập hợp, nó chỉ cho phép tập hợp giá trị của 0 hay 1. Điều này cho phép cung cấp các công cụ hỗ trợ OWL Lite dễ dàng hơn so với các bản khác.

OWL DL (OWL Description Logic): hỗ trợ cho những người dùng cần sự diễn cảm tối đa trong khi cần duy trì tính toán toàn vẹn (tất cả các kết luận phải được đảm bảo để tính toán) và tính quyết định (tất cả các tính toán sẽ kết thúc trong khoảng thời gian hạn chế). OWL DL bao gồm tất cả các cấu trúc của ngôn ngữ OWL, nhưng chúng chỉ có thể được sử dụng với những hạn chế nào đó (Ví dụ: Trong khi một lớp có thể là một lớp con của rất nhiều lớp, một lớp không thể là một thể hiện của một lớp khác). OWL DL cũng được chỉ định theo sự tương ứng với logic mô tả, một lĩnh vực nghiên cứu trong logic đã tạo nên sự thiết lập chính thức của OWL.

OWL Full muốn đề cập tới những người dùng cần sự diễn cảm tối đa và sự tự do của RDF mà không cần đảm bảo sự tính toán của các biểu thức. Ví dụ, trong OWL Full, một lớp có thể được xem xét đồng thời như là một tập của các cá thể và như là một cá thể trong chính bản thân nó. OWL Full cho phép một ontology gia cố thêm

ý nghĩa của các từ vựng được định nghĩa trước (RDF hoặc OWL).

Các phiên bản này tách biệt về các tiện ích khác nhau, OWL Lite là phiên bản dễ hiểu nhất và phức tạp nhất là OWL Full.

Mối liên hệ giữa các ngôn ngữ con của OWL:

- Mọi ontology hợp lệ dựa trên OWL Lite đều là ontology hợp lệ trên OWL DL
- Mọi ontology hợp lệ dựa trên OWL DL đều là ontology hợp lệ trên OWL Full
- Mọi kết luận hợp lệ dựa trên OWL Lite đều là kết luận hợp lệ trên OWL DL
- Mọi kết luận hợp lệ dựa trên OWL DL đều là kết luận hợp lệ trên OWL Full

Công cụ để xây dựng các Ontology là Protégé. Công cụ này được sử dụng để tạo ra file OWL

## **II. Ứng dụng của ontology trong công nghệ tri thức**

**Hệ Thống Tìm Kiếm Thông Tin Mạng:** Ontology được sử dụng để tạo ra các hệ thống tìm kiếm thông tin mạng thông minh. Chúng giúp cải thiện khả năng tìm kiếm, hiểu ngữ cảnh và kết nối thông tin từ nhiều nguồn khác nhau. Ví dụ, các công cụ tìm kiếm thông tin y tế có thể sử dụng ontology để hiểu và phân loại thông tin về bệnh tật, thuốc, và triệu chứng.

**Hệ Thống Hỗ Trợ Ra Quyết Định:** Ontology được sử dụng để xây dựng hệ thống hỗ trợ ra quyết định (Decision Support Systems - DSS) trong nhiều lĩnh vực, bao gồm y tế, tài chính, và quản lý. Chúng giúp tổ chức và tổng hợp thông tin, làm cho dữ liệu dễ dàng hiểu và áp dụng trong việc đưa ra quyết định chính xác.

**Trí Tuệ Nhân Tạo (AI) và Học Máy:** Ontology đóng vai trò quan trọng trong việc đào tạo và phát triển mô hình máy học. Chúng giúp cải thiện khả năng hiểu dữ liệu, phân loại và xử lý thông tin. Hệ thống học máy có thể sử dụng ontology để hiểu ngữ cảnh và quy tắc của các dự đoán.

**Chatbot và Giao Tiếp Máy Người (HCI):** Ontology được sử dụng để cung cấp cơ sở tri thức cho chatbot và hệ thống giao tiếp máy người. Chúng giúp chatbot hiểu và đưa ra câu trả lời thông minh hơn dựa trên ngữ cảnh của cuộc trò chuyện và kiến thức về lĩnh vực cụ thể.

**Quản Lý Tri Thức Doanh Nghiệp:** Ontology được sử dụng trong quản lý tri thức doanh nghiệp (Enterprise Knowledge Management) để tổ chức và chia sẻ thông tin và kiến thức trong tổ chức. Chúng giúp cải thiện quy trình làm việc, giúp nhân viên dễ dàng truy cập thông tin cần thiết.

**Lĩnh Vực Trí Tuệ Nhân Tạo Mềm (Soft AI):** Ontology đóng vai trò quan trọng trong xây dựng các hệ thống trí tuệ nhân tạo mềm như hệ thống khuyến nghị sản phẩm, hệ thống phân loại nội dung, và phân tích dữ liệu lớn. Chúng giúp tạo ra mô hình và hệ thống có khả năng hiểu thông tin và tương tác với người dùng.

**Phát Triển Ứng Dụng IoT (Internet of Things):** Ontology được sử dụng trong phát triển các ứng dụng IoT để mô tả và quản lý dữ liệu từ các thiết bị và cảm biến khác nhau. Chúng giúp tạo ra một lớp trừu tượng để biểu diễn các dịch vụ và tương tác trong môi trường IoT.

**Tìm Kiếm Tri Thức Toàn Cầu:** Ontology có thể được sử dụng để tạo ra các hệ thống tìm kiếm tri thức toàn cầu, giúp người dùng truy cập và chia sẻ kiến thức từ các nguồn trên toàn thế giới.

**Quản Lý Tri Thức Về Dự Án và Quản Lý Tri Thức Cộng Đồng:** Ontology được sử dụng trong quản lý tri thức về dự án (Project Knowledge Management) và quản lý tri thức cộng đồng (Community Knowledge Management) để tổ chức, lưu trữ và truy cập kiến thức và thông tin cộng đồng.

Những ứng dụng này chỉ là một phần nhỏ của những cách ontology có thể được áp dụng trong công nghệ tri thức. Ontology là một công cụ mạnh mẽ để biểu diễn và tổ chức kiến thức, giúp cải thiện khả năng tìm kiếm thông tin, ra quyết định, và tạo ra các hệ thống thông minh trong nhiều lĩnh vực.