|  |  |
| --- | --- |
|  |  |
| TRƯƠNG ĐỨC NHẬT | **BỘ GIÁO DỤC VÀ ĐÀO TẠO**  **TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI**  **---------------------------------------** |
|  |
|  |
|  |
|  |
| **Trương Đức Nhật** |
|  |
|  |
|  |
|  |
|  |
| CÔNG NGHỆ THÔNG TIN |  |
|  |
| **CỔNG THÔNG TIN TRI THỨC THÔNG MINH - ỨNG DỤNG VÀO NÔNG SẢN NÔNG NGHIỆP VIỆT NAM** |
|  |
|  |
|  |
|  |
|  |
|  |
|  |  |
|  | LUẬN VĂN THẠC SĨ KỸ THUẬT  CÔNG NGHỆ THÔNG TIN |
|  |  |
|  |  |
| 2013B |  |
|  |
|  |
|  |
| Hà Nội – Năm 2016 |
|  | |
| **BỘ GIÁO DỤC VÀ ĐÀO TẠO**  **TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI**  **---------------------------------------** | |
|  | |
| **Trương Đức Nhật** | |
|  | |
|  | |
|  | |
|  | |
|  | |
| **CỔNG THÔNG TIN TRI THỨC THÔNG MINH - ỨNG DỤNG VÀO NÔNG SẢN NÔNG NGHIỆP VIỆT NAM** | |
|  | |
|  | |
|  | |
|  | |
| Chuyên ngành : Công nghệ thông tin | |
|  | |
|  | |
|  | |
|  | |
|  | |
| LUẬN VĂN THẠC SĨ KỸ THUẬT  CÔNG NGHỆ THÔNG TIN. | |
|  | |
|  | |
|  | |
|  | |
|  | |
| NGƯỜI HƯỚNG DẪN KHOA HỌC : | |
| 1. TS Phạm Văn Hải | |
|  | |
|  | |
|  | |
|  | |
|  | |
| Hà Nội – Năm 2016 | |

**LỜI CAM ĐOAN**

Những kiến thức trình bày trong luận văn là do tôi tìm hiểu, nghiên cứu và trình bày theo những kiến thức tổng hợp của cá nhân. Kết quả nghiên cứu trong luận văn này chưa từng được công bố tại bất kỳ công trình nào khác. Trong quá trình làm luận văn, tôi có tham khảo các tài liệu có liên quan và đã ghi rõ nguồn tài liệu tham khảo. Tôi xin cam đoan đây là công trình nghiên cứu của tôi và không sao chép của bất kỳ ai.

Tôi xin chịu hoàn toàn trách nhiệm, nếu sai, tôi xin chịu mọi hình thức kỷ luật theo quy định.

Hà Nội, ngày 17 tháng 4 năm 2016

Học viên

Trương Đức Nhật

**LỜI CẢM ƠN**

            Trước tiên, tôi xin bày tỏ lòng biết ơn sâu sắc tới TS. Phạm Văn Hải và các thầy cô Viện CNTT-TT, Trường Đại học Bách Khoa Hà Nội và TS đã nhiệt tình hướng dẫn và đào tọa cho tôi để tạo điều kiện thuận lợi cho tôi nghiên cứu khoa học, và giúp tôi có thể hoàn thành luận văn một cách tốt nhất.

            Cuối cùng tôi xin gửi lời cám ơn đến gia đình, bạn bè, những người đã luôn bên tôi, động viên và khuyến khích tôi trong quá trình thực hiện đề tài nghiên cứu của mình.

Học viên

Trương Đức Nhật

BOOKS

MỤC LỤC

[MỤC LỤC 5](#_Toc450810933)

[DANH MỤC HÌNH VẼ VÀ SƠ ĐỒ 7](#_Toc450810934)

[DANH MỤC BẢNG 7](#_Toc450810935)

[MỞ ĐẦU 8](#_Toc450810936)

[1. Lý do chọn đề tài 8](#_Toc450810937)

[2. Mục tiêu và nhiệm vụ nghiên cứu 9](#_Toc450810938)

[3. Đối tượng và phạm vi nghiên cứu 9](#_Toc450810939)

[4. Phương pháp nghiên cứu 10](#_Toc450810940)

[5. Kết quả dự kiến 10](#_Toc450810941)

[6. Ý nghĩa khoa học và thực tiễn 10](#_Toc450810942)

[7. Bố cục luận văn 11](#_Toc450810943)

[CHƯƠNG 1 NGHIÊN CỨU TỔNG QUAN 12](#_Toc450810944)

[1. Tổng quan về cổng thông tin điện tử về nông nghiệp 12](#_Toc450810945)

[2. Cơ sở lý thuyết tri thức về hệ cơ sở tri thức và suy diễn 13](#_Toc450810946)

[2.1. Hệ cơ sở tri thức 13](#_Toc450810947)

[2.2. Biểu diễn tri thức và các phương pháp suy diễn. 16](#_Toc450810948)

[3. Nghiên cứu hiện trạng các cổng thông tin nông sản 21](#_Toc450810949)

[4. Tổng kết chương 24](#_Toc450810950)

[CHƯƠNG 2: MÔ HÌNH CỔNG THÔNG TIN TRI THỨC NÔNG SẢN 25](#_Toc450810951)

[1. Kiến trúc hệ thống cổng thông tin tri thức 25](#_Toc450810952)

[2. Diễn giải 26](#_Toc450810953)

[2.1. Module Q&A 26](#_Toc450810954)

[2.2. Module Reliability 30](#_Toc450810955)

[3. Ứng dụng suy diễn tri thức trong cổng thông tin 33](#_Toc450810956)

[4. Tổng kết chương 34](#_Toc450810957)

[CHƯƠNG 3: PHÂN TÍCH VÀ THIẾT KẾ HỆ THỐNG 34](#_Toc450810958)

[1. Giới thiệu nghiệp vụ và công tác điều hành cổng thông tin 34](#_Toc450810959)

[1.1. Yêu cầu chức năng 34](#_Toc450810960)

[1.2. Yêu cầu phi chức năng 35](#_Toc450810961)

[1.3. Biểu đồ phân cấp chức năng 35](#_Toc450810962)

[1.4. Các tác nhân tương tác với hệ thống 36](#_Toc450810963)

[2. Biểu đồ ca sử dụng (use case) tổng quan 37](#_Toc450810964)

[3. Biểu đồ hoạt động tư vấn và đánh giá 37](#_Toc450810965)

[4. Biểu đồ trình tự tư vấn và đánh giá 39](#_Toc450810966)

[5. Mô hình thực thể liên kết 42](#_Toc450810967)

[6. Thiết kế cơ sở dữ liệu 43](#_Toc450810968)

[7. Thiết kế cơ sở tri thức và cơ chế suy diễn 45](#_Toc450810969)

[8. Tổng kết chương 47](#_Toc450810970)

[CHƯƠNG 4: CÀI ĐẶT CHƯƠNG TRÌNH VÀ KẾT QUẢ 48](#_Toc450810971)

[1. Cài đặt chương trình 48](#_Toc450810972)

[1.1. Môi trường cài đặt và phát triển 48](#_Toc450810973)

[1.2. Đăng nhập vào quản trị hệ thống 48](#_Toc450810974)

[1.3. Quản trị cơ sở tri thức 49](#_Toc450810975)

[2. Kết quả chương trình 52](#_Toc450810976)

[3. Tổng kết chương 55](#_Toc450810977)

[CHƯƠNG 5: KẾT LUẬN 56](#_Toc450810978)

[1. Kết luận 56](#_Toc450810979)

[2. Định hướng phát triển 56](#_Toc450810980)

[TÀI LIỆU THAM KHẢO 57](#_Toc450810981)

**DANH MỤC HÌNH VẼ VÀ SƠ ĐỒ**

[Hình 1: Kiến Trúc Hệ Cơ Sở Tri Thức 14](#_Toc447790290)

[Hình 2: Sơ Đồ Kiến Trúc Hệ Thống 25](#_Toc447790291)

[Hình 3: Biểu Đồ Phân Cấp Chức Năng 36](#_Toc447790292)

[Hình 4: Biểu Đồ Ca Sử Dụng Tổng Quan 37](#_Toc447790293)

[Hình 5: Biểu Đồ Hoạt Động Tư Vấn Suy Diễn 38](#_Toc447790294)

[Hình 6: Biểu Đồ Hoạt Động Tư Vấn Đánh Giá Độ Chính Xác 39](#_Toc447790295)

[Hình 7: Biểu Đồ Trình Tự Tư Vấn Suy Diễn Bởi Chuyên Gia 40](#_Toc447790296)

[Hình 8: Biểu Đồ Hoạt Động Tư Vấn Đánh Giá Độ Chính Xác Bởi Chuyên Gia 41](#_Toc447790297)

[Hình 9: Mô Hình Thực Thể Liên Kết 42](#_Toc447790298)

[Hình 10: Form Đăng Nhập 48](#_Toc447790299)

[Hình 11: Menu Quản Trị 49](#_Toc447790300)

[Hình 12: Menu Quản Trị Cơ Sở Tri Thức 49](#_Toc447790301)

[Hình 13: Quản Trị Sự Kiện 50](#_Toc447790302)

[Hình 14: Quản Lý Luật Suy Diễn – Chọn Giả Thiết 50](#_Toc447790303)

[Hình 15: Quản Lý Luật Suy Diễn – Chọn Kêt Luận 51](#_Toc447790304)

[Hình 16: Quản Lý Chuyên Mục Cần Tư Vấn 51](#_Toc447790305)

[Hình 17: Quản Lý Câu Hỏi Hỗ Trợ Tư Vấn 52](#_Toc447790306)

[Hình 18: Menu Tư Vấn 52](#_Toc447790307)

[Hình 19: Form Đặt Câu Hỏi Tư Vấn 53](#_Toc447790308)

[Hình 20: Kết Quả Tư Vấn Của Module Q&A 54](#_Toc447790309)

[Hình 21: Kết Quả Tư Vấn Của Module Reliability (1) 54](#_Toc447790310)

[Hình 22: Kết Quả Tư Vấn Của Module Reliability (2) 55](#_Toc447790311)

# DANH MỤC BẢNG

[Bảng 1: Bảng Agr\_User 43](#_Toc447788195)

[Bảng 2: Bảng Ex\_Course 44](#_Toc447788196)

[Bảng 3: Bảng Ex\_Even 44](#_Toc447788197)

[Bảng 4: Bảng Ex\_Rule 44](#_Toc447788198)

[Bảng 5: Bảng Ex\_Question 44](#_Toc447788199)

**MỞ ĐẦU**

1. **Lý do chọn đề tài**

Sản xuất nông nghiệp tăng mạnh cung cấp nguồn hàng nhiều và đều đặn cho thị trường trong nước và giúp cho người tiêu dùng trong nước có thể tiếp cận thực phẩm đa dạng, nhiều và rẻ. Thành công vượt bậc trong ngành nông nghiệp không chỉ giúp Việt Nam tăng GDP trên 8% mà còn giúp giảm đói nghèo.

Sự bùng nổ của Internet cùng sự thiếu kiến thức về nông sản nông nghiệp của người nông dân cần được cải thiện. Đối với nông nghiệp, cổng thông tin tri thức thông minh sẽ đóng vai trò quan trọng trong việc cung cấp điểm truy cập duy nhất tới thông tin và các dịch vụ liên quan tới nông nghiệp, đồng thời giải quyết được nhu cầu tìm kiếm và cần tư vấn hỗ trợ từ người nông dân với vai trò tương đương một chuyên gia trong lĩnh vực nông nghiệp mà hệ thống được tích hợp. thiếu phương thức tiếp cận và giao tiếp với các kiến thức và chuyên gia nông nghiệp [1,2, 7-11].

Hiện nay, tại nhiều nước châu Á như: Thái Lan, Campuchia, Ấn Độ... đã áp dụng mô hình nông nghiệp điện tử, người nông dân được cập nhật và cảnh báo các thông tin nông nghiệp nhanh chóng, kịp thời, từ đó nâng cao hiệu quả và năng suất lao động [1-6]. Tuy nhiên tại Việt Nam, bà con nông dân vẫn gặp phải nhiều khó khăn do thiếu thông tin cập nhật và kiến thức liên quan đến nông nghiệp; dẫn đến nhiều rủi ro như thương lái ép giá, phân bón giả, thuốc trừ sâu giả, dịch bệnh….

Ở Việt Nam, chúng ta vẫn chưa có được một hệ thống cổng thông tin đầy đủ và các cổng thông tin hiện nay còn tồn tại nhiều vấn đề như:

* Các thông tin về nông sản, nông nghiệp còn ít và chưa phổ biến rộng rãi trong cộng đồng.
* Liên kết giữa người nông dân – doanh nghiệp – và chính sách nông nghiệp còn hạn chế bởi thiếu công cụ công nghệ thông tin để quản lý.
* Việc tìm kiếm thông tin vẫn còn khó khăn và chưa thân thiện, mặc dù Internet của chúng ta rất phát triển.
* Khả năng tư vấn tri thức trực tiếp từ các chuyên gia nông nghiệp vẫn chưa được hỗ trợ.

1. **Mục tiêu và nhiệm vụ nghiên cứu**

Nghiên cứu và xây dựng cổng thông tin tri thức tri thức điện tử cho nông sản nông nghiệp với tư vấn thông minh và làm tăng sự tương tác giữa người dùng và cổng thông tin với các tiêu chí cụ thể như sau:

* Tạo các cộng đồng liên kết người nông dân – doanh nghiệp – các chính sách nông nghiệp với nhau.
* Tích hợp kỹ thuật suy diễn tri thức để suy hỗ trợ dịch vụ tư vấn thông minh trong cổng thông tin điện tử.
* Cổng thông tin cài đặt trên Internet thử nghiệm suy diễn với tư vấn dựa trên cơ sở tri thức.

Về mặt ứng dụng thực tiễn, tác giả xây dựng một cổng thông tin chính thức có tư vấn suy diễn bán tự động trên nền Web và thử nghiệm cơ chế suy diễn từ cổng thông tin điện tử nông sản, nông nghiệp.

1. **Đối tượng và phạm vi nghiên cứu**

Nghiên cứu và xây dựng cổng thông tin tri thức thông minh đưa thông tin về nông sản tới người nông dân trong phạm vi:

* Đối tượng nghiên cứu: sản phẩm nông sản cây lúa tại Việt Nam.
* Khai thác thông tin từ nông sản lúa trong Bộ nông nghiệp và phát triển nông thôn

1. **Phương pháp nghiên cứu**

Tiếp cận các phương pháp biểu diễn tri thức logic mệnh đề, suy diễn tiến, tác giả đề xuất các cách suy diễn mô tả được cơ chế suy diễn trong cổng thông tin nông sản- nông nghiệp.

Phân tích thiết kế hệ thống cổng thông tin nông sản và tích hợp các kỹ thuật suy diễn, ứng dụng thực tiễn vào miền nghiên cứu nông sản – nông nghiệp Việt Nam.

1. **Kết quả dự kiến**

Kết quả của nghiên cứu là sản phẩm phần mềm có tư vấn suy diễn trên nền Web, đóng góp mới trong việc tích hợp các cơ chế suy diễn trong cổng thông tin nông sản nông nghiệp. Các biểu diễn tri thức trong hệ cơ sở tri thức đối với cổng thông tin nông sản nông nghiệp đòi hỏi phải am hiểu rõ về đặc thù nghiệp vụ, cơ chế suy diễn theo hình thức bán tự động, kết hợp với tri thức chuyên gia.

Nghiên cứu cổng thông tin tri thức thông minh bằng cách xây dựng cổng thông tin kết hợp với các kỹ thuật suy diễn mang lại hiệu quả cho việc tư vấn của các chuyên gia nông sản – nông nghiệp. Kết quả thực nghiệm, giúp cho các chuyên gia tư vấn trực tuyến, đồng thời cập nhật các tri thức chuyên gia, nhằm hỗ trợ tương tác trực tuyến giữa các chuyên gia, nhà chính sách và cộng đồng nông dân.

1. **Ý nghĩa khoa học và thực tiễn**

Hầu hết các cổng thông tin nông nghiệp ra đời hiện nay đáp ứng sự tương tác giữa các tác nhân trong hệ thống. Khi quản lý cổng thông tin, dịch vụ tích hợp dịch vụ tư vấn dựa vào tri thức suy diễn tự động còn là nghiên cứu mới. Chính vì thế, cổng thông tin tri thức hiện đang được các nhà khoa học, kỹ thuật quan tâm.

Để giảm thiểu chi phí, tăng thời gian tương tác với tư vấn duy diễn dựa vào hệ cơ sở tri thức và cổng thông tin, nghiên cứu đưa ra giải pháp tư vấn suy diễn áp dụng vào cổng thông tin nông sản nông nghiệp.

1. **Bố cục luận văn**

MỞ ĐẦU

* Tính cấp thiết của đề tài
* Đối tượng và phạm vi nghiên cứu
* Mục tiêu nghiên cứu
* Phương pháp nghiên cứu

CHƯƠNG 1: NGHIÊN CỨU TỔNG QUAN

* Tổng quan về cổng thông tin điện tử về nông nghiệp
* Cơ sở lý thuyết tri thức về hệ cơ sở tri thức và suy diễn
* Nghiên cứu hiện trạng các cổng thông tin nông sản
* Tổng kết chương

CHƯƠNG 2: MÔ HÌNH CỔNG THÔNG TIN TRI THỨC NÔNG SẢN

* Kiến trúc hệ thống cổng thông tin tri thức
* Diễn giải
* Ứng dụng suy diễn tri thức trong cổng thông tin
* Tổng kết chương

CHƯƠNG 3: PHÂN TÍCH THIẾT KẾ HỆ THỐNG

* Giới thiệu nghiệp vụ và công tác điều hành cổng thông tin
* Biểu đồ ca sử dụng (use case) tổng quan và các biểu đồ hoạt động
* Mô hình liên kết thực thể
* Tổng kết chương

CHƯƠNG 4: CÀI ĐẶT VÀ KẾT QUẢ

* Cài đặt chương trình
* Kết quả chương trình
* Tổng kết chương

CHƯƠNG 5: KẾT LUẬN

* Kết luận
* Định hướng phát triển

**CHƯƠNG 1 NGHIÊN CỨU TỔNG QUAN**

1. **Tổng quan về cổng thông tin điện tử về nông nghiệp**

Ở Việt Nam, sản xuất nông nghiệp là một lĩnh vực hết sức quan trọng đối với nền kinh tế Việt Nam. Hiện nay, dân số nông thôn chiếm đến 60,7% dân số quốc gia, ngành nông lâm thủy sản giải quyết công ăn việc làm cho 56,8% người trong độ tuổi lao động và đóng góp đến 20,9% GDP quốc gia [13], thu lại lợi nhuận thông qua kim ngạch xuất khẩu và phát triển nền sản xuất hàng hóa bền vững. Vì vậy, bất kỳ tác động nào của việc sản xuất nông nghiệp và nông thôn đều gây ra những ảnh hưởng lớn lao và lâu dài đến toàn bộ nền kinh tế.

Trong những năm qua, sản xuất nông nghiệp của Việt Nam đã phát triển toàn diện và đạt được nhiều thành tựu quan trọng cả về năng suất và sản lượng. Nhiều loại nông sản, thủy sản cũng đã phát triển nhanh và trở thành những mặt hàng xuất khẩu như: lúa gạo, thủy sản, cà phê, hồ tiêu, hạt điều, chè, cao su, rau quả [12, 13]…

Sản xuất nông nghiệp tăng mạnh cung cấp nguồn hàng nhiều và đều đặn cho thị trường trong nước và giúp cho người tiêu dùng trong nước có thể tiếp cận thực phẩm đa dạng, nhiều và rẻ. Thành công vượt bậc trong ngành nông nghiệp không chỉ giúp Việt Nam tăng GDP trên 8% mà còn giúp giảm đói nghèo [12].

Việc gia nhập tổ chức Thương mại thế giới (WTO) đã tạo cơ hội cho Việt Nam phát triển thị trường sản phẩm mới như cà phê, hạt điều, rau, chè, trái cây nhiệt đới và sản phẩm thịt lợn. Những mặt hàng này được sản xuất ở Việt Nam rẻ hơn nhiều so với những nước láng giềng ASEAN. Nhưng, giá thành vận chuyển tương đối cao và chi phí xử lý sau thu hoạch đã làm tăng giá thành sản phẩm, làm giảm tính cạnh tranh của nhiều mặt hàng nông sản của Việt Nam trong thị trường khu vực.

Trước những cơ hội đó, Việt Nam cần phải giải quyết vấn đề về giá thành hay năng suất của nông sản. Điều đó đòi hỏi cần phải có những hệ thống quản lý nông sản và những hệ thống hỗ trợ tương tác về nông sản, giúp phổ biến kiến thức nông sản cho mọi người, đặc biệt là những người nông dân.

Cùng với sự bùng nổ của internet hiện nay, việc xây dựng một hệ thống portal có thể nói là tối ưu cho việc hỗ trợ quản lý và phổ biến thông tin nông nghiệp tới mọi người qua internet. Tuy nhiên, một hệ thống portal là vẫn chưa đủ để đáp ứng nhu cầu tìm kiếm thông tin của người nông dân, bởi họ cần một hệ thống thông minh hơn, có thể trả lời mọi câu hỏi, tư vấn cho họ các kiến thức về nông nghiệp, nông sản.

Khái niệm này tuy không mới trên thế giới nhưng khá mới ở Việt Nam, hệ thống cổng thông tin tri thức thông minh, có thể hiểu là sự tích hợp việc quản lý tri thức về một miền kiến thức nào đó vào hệ thống portal [13, 16-17].

Đối với nông nghiệp, cổng thông tin tri thức thông minh sẽ đóng vai trò quan trọng trong việc cung cấp điểm truy cập duy nhất tới thông tin và các dịch vụ liên quan tới nông nghiệp, đồng thời giải quyết được nhu cầu tìm kiếm và cần tư vấn hỗ trợ từ người nông dân với vai trò tương đương một chuyên gia trong lĩnh vực nông nghiệp.

1. **Cơ sở lý thuyết tri thức về hệ cơ sở tri thức và suy diễn**
   1. **Hệ cơ sở tri thức**
      1. **Khái niệm**

Hệ cơ sở tri thức là hệ thống dựa trên tri thức cho phép mô hình hóa các tri thức của chuyên gia, dùng tri thức này nhằm giải quyết vấn đề phức tạp thuộc cùng lĩnh vực [17]. Đặc biệt hệ chuyên gia là một loại hệ cơ sở tri thức được thiết kế cho một lĩnh vực ứng dụng cụ thể. Ngày nay, ứng dụng của các hệ cơ sở tri thức ngày càng trở lên phong phú và phổ biến trong nhiều lĩnh vực khác nhau. Ví dụ: hệ chuyên gia chẩn đoán bệnh trong y học, hệ chuyên gia tư vấn học tập và giảng dạy, hệ chẩn đoán hỏng hóc máy tính…

* + 1. **Kiến trúc của một hệ cơ sở tri thức**

Một hệ cơ sở tri thức bao gồm các thành phần chính sau [17]:

Động cơ suy diễn

Bộ xử lý ngôn ngữ tự nhiên

Tìm kiếm

ễn

Điều khiển

Giải thích

Cơ sở tri thức

Vùng nhớ làm việc

Sự kiện

Chuyên gia

Tiếp nhận tri thức

Luật

**Hình 1: Kiến trúc hệ cơ sở tri thức**

Trong đó, 2 thành phần chính quan trọng nhất là ***Cơ sở tri thức*** *và* ***động cơ suy diễn****.*

* ***Cơ sở tri thức*** là nơi lưu trữ biểu diễn các tri thức mà hệ đảm nhận, làm cơ sở cho các hoạt động của hệ. Cơ sở tri thức bao gồm các sự kiện và các luật.
* ***Động cơ suy diễn:*** là quá trình trong hệ chuyên gia cho phép khớp các sự kiện trong vùng nhớ làm việc với các tri thức về các lĩnh vực trong cơ sở tri thức, để rút ra các kết luận về các vấn đề đang giải quyết.

Ngoài ra còn có các bộ xử lý ngôn ngữ tự nhiên, bộ giải thích WHY-HOW, vùng nhớ làm việc, bộ tiếp nhận tri thức và người chuyên gia (kĩ sư tri thức).

Tri thức không có cấu trúc mà vẫn hiểu được là tri thức ngầm ý, các tri thức không rõ. Hệ cơ sở tri thức là một tập hợp các lập luận, các luật, các quy trình thủ tục được tổ chức thành các lược đồ , giản đồ. Đó là một tập hợp của tất cả các thông tin cũng như các kiến thức về một lĩnh vực cụ thể nào đó **.**

Quá trình thu thập tri thức là quá trình rút trích tri thức và định dạng tri thức được thực hiện từ nhiều nguồn khác nhau, đặc biệt là các chuyên gia trong một lĩnh vực cụ thể nào đó. Công việc này là một trong những bước quan trọng và thường mất rất nhiều thời gian và công sức trong quá trình xây dựng một hệ cơ sở tri thức trong một hệ cơ sở tri thức. Trong quá trình phát triển một hệ cơ sở tri thức thì các kĩ sư tri thức, những nhà chuyên môn trong lĩnh vực trí tuệ nhân tạo có nhiệm vụ thu thập tri thức từ các chuyên gia thuộc lĩnh vực chuyên ngành sau đó “sao chép” các tri thức đó vào cơ sở tri thức và diễn đạt tri thức đó dưới dạng có thể dùng được trong các hệ cơ sở tri thức.

* + 1. **Ứng dụng của hệ cơ sở tri thức trong thực tế.**

Trong thực tế, rất nhiều lĩnh vực khác nhau có áp dụng hệ cơ sở tri thức nói chung và hệ chuyên gia nói riêng như: nông nghiệp, công nghiệp, y học, quân sự…

Các dạng bài toán hay áp dụng hệ chuyên gia bao gồm:

* Diễn giải (Interpretation): Đưa ra mô tả tình huống các dữ liệu thu thập được.
* Dự báo (Hediction): Đưa ra hậu quả của một tình huống nào đó, như là dự báo thời tiết, dự báo giá cả thị trường.
* Chuẩn đoán (Diagnosis): Xác định các lỗi , các bộ phận hỏng hóc của hệ thống dựa trên các dữ liệu quan sát được.
* Gỡ rối (Debugging): Mô tả các phương pháp khắc phục hệ thống khi gặp sự cố.
* Thiết kế (Design) : Lựa chọn cấu hình các đói tượng nhằm thoả mãn một số ràng buộc nào đó.
* Giảng dạy (Instruction): Phần mềm dạy học, có thể chuẩn đoán và sửa lỗi của học sinh trong quá trình học tập.
  1. **Biểu diễn tri thức và các phương pháp suy diễn.**
     1. **Tri thức và suy diễn**

***Tri thức*** hay kiến thức có nhiều ý nghĩa tùy theo văn cảnh, nhưng lúc nào cũng có liên quan với những khái niệm như hiểu biết, ý nghĩa, thông tin, giảng dạy, giáo dục (quá trình giáo dục), giao tiếp, diễn tả, học hỏi, suy luận, nhận thức và kích thích trí nhớ. Bởi thế ta có thể định nghĩa tri thức là [17]:

* Các thông tin, các tài liệu, các cơ sở lý luận, các kỹ năng khác nhau, đạt được bởi một tổ chức hay một cá nhân thông qua các trải nghiệm thực tế hay thông qua sự giáo dục đào tạo, các hiểu biết về lý thuyết hay thực tế về một đối tượng, một vấn đề, có thể lý giải được về nó.
* Những gì đã biết, đã được hiểu biết trong một lĩnh vực cụ thể hay toàn bộ, trong tổng thể.
* Các cơ sở, các thông tin, tài liệu, các hiểu biết hoặc những thứ tương tự có được bằng kinh nghiệm thực tế hoặc do những tình huống, hoàn cảnh cụ thể. Những tranh cãi về mặt triết học nhìn chung bắt đầu với phát biểu của Plato: tri thức như là "justified true belief ". Tuy nhiên không có một định nghĩa chính xác nào về tri thức hiện nay được mọi người chấp nhận, có thể bao quát được toàn bộ, vẫn còn nhiều học thuyết, các lý luận khác nhau về tri thức.

Tri thức giành được thông qua các quá trình nhận thức phức tạp: quá trình tri giác, quá trình học tập, tiếp thu, quá trình giao tiếp, quá trình tranh luận, quá trình lý luận, hay kết hợp các quá trình này.

***Tri thức*** được phân chia thành các loại sau:

* Tri thức sự kiện: Tri thức sự kiện là một khẳng định về một sự kiện, hiện tượng hay một khái niệm nào đó trong một hoàn cảnh không gian hoặc thời gian nhất định. Khẳng định về hiện tượng: ”Mặt trời lặn ở phương Tây”. Khái niệm về: “tam giác đều : là tam giác có ba góc bằng nhau”.
* Tri thức thủ tục**:** Tri thức thủ tục là tri thức mô tả cách giải quyết một vấn đề, quy trình xử lý các công việc, lịch trình tiến hành các thao tác … Các dạng của tri thức thủ tục thường dùng là các luật, chiến lược, lịch trình ví dụ như : IF muốn thi khối A THEN phải học tốt các môn tự nhiên.
* Tri thức heuristic:là tri thức nông cạn do không đảm bảo hoàn toàn chính xác hoặc tối ưu theo một nghĩa nào đó về cách giải quyết vấn đề. Tri thức heuristicthường được coi là một mẹo nhằm dẫn dắt tiến trình lập luận. Ví dụ như một số giải thuật tìm đường đi ngắn nhất, giải thuật A\* có thể được coi là lời giả của một vấn đề tốt nhưng chưa hẳn tối ưu.

Ngoài ra, người ta phân chia ra tri thức mêta: tri thức tham chiếu đến các tri thức khác; tri thức có cấu trúc: tri thức về các quan hệ giữa các khái niệm, quan hệ giữa các đối tượng…

* + 1. **Biểu diễn tri thức**

Tri thức của một hệ cơ sở tri thức có thể được biểu diễn theo nhiều cách khác nhau. Thông thường người ta sử dụng các cách sau đây:

* Biểu diễn tri thức bởi các luật sản xuất.
* Biểu diễn tri thức bởi các mệnh đề logic.
* Biểu diễn tri thức nhờ mạng ngữ nghĩa.
* Biểu diễn tri thức nhờ ngôn ngữ nhân tạo.

Ngoài ra, người ta còn sử dụng cách biểu diễn tri thức nhờ các sự kiện không chắc chắn, nhờ bộ ba : đối tượng, thuộc tính và giá trị (O-A-V: Object-Attribute-Value), nhờ khung (frame), v.v..Tuỳ theo từng hệ chuyên gia, người ta có thể sử dụng một cách hoặc đồng thời cả nhiều cách.

1. **Biểu diễn tri thức bởi các luật sản xuất.**

Luật là một kiểu sản xuất được nghiên cứu từ những năm 1940. Trong một hệ thống dựa trên luật, công cụ suy luận sẽ xác định những luật nào là tiên đề thỏa mãn các sự việc. Các luật sản xuất thường được viết dưới dạng IF THEN. Có hai dạng : IF < điều kiện > THEN < hành động > hoặc IF < điều kiện > THEN < kết luận > DO < hành động > . Tuỳ theo hệ chuyên gia cụ thể mà mỗi luật có thể được đặt tên. Chẳng hạn mỗi luật có dạng Rule : tên. Sau phần tên là phần IF của luật. Phần giữa IF và THEN là phần trái luật (LHS: Left - Hand -Sid ), có nội dung được gọi theo nhiều tên khác nhau, như tiền đề (antecedent), điều kiện (conditional part), mẫu so khớp (pattern part), phần sau THEN là kết luận hay hậu quả (consequent). Một số hệ cơ sở tri thức có thêm phần hành động (action) được gọi là phần phải luật (RHS: Right - Hand -Side).

**Ví dụ** :

Rule : tụ huyết trùng IF lợn sốt cao AND nước tiểu vàng AND da xuất huyết AND cổ sưng THEN bị tụ huyết trùng lợn.

1. **Biểu diễn tri thức nhờ mệnh đề logic.**

Người ta sử dụng các ký hiệu để thể hiện tri thức và các phép toán lôgic tác động lên các ký hiệu để thể hiện suy luận lôgic. Kỹ thuật chủ yếu thường được sử dụng là lôgic vị từ (predicate logic). Các vị từ thường có chứa hạng, biến hay hàm. Người ta gọi các vị từ không chứa biến (có thể chứa hạng) là các mệnh đề (preposition). Mỗi vị từ có thể là một sự kiện ( fact) hay một luật. Luật là vị từ gồm hai vế trái và phải được nối nhau bởi một dấu mũi tên (→). Các vị từ còn lại (không chứa mũi tên) được gọi là các sự kiện. Người ta gọi tập hợp các sự kiện và các luật là một cơ sở tri thức.

1. **Biểu diễn tri thức nhờ mạng ngữ nghĩa.**

Trong phương pháp này, người ta sử dụng một đồ thị gồm các nút (node) và các cung (arc) nối các nút để biểu diễn tri thức. Nút dùng để thể hiện các đối tượng, thuộc tính của đối tượng và giá trị của thuộc tính. Còn cung dùng để thể hiện các quan hệ giữa các đối tượng. Các nút và các cung đều được gắn nhãn. Bằng cách thêm vào đồ thị các nút mới và các cung mới, người ta có thể mở rộng một mạng ngữ nghĩa. Các nút mới được thêm thể hiện các đối tượng tương tự (với các nút đã có trong đồ thị), hoặc tổng quát hơn. Một trong những tính chất quan trọng của mạng ngữ nghĩa là tính thừa kế. Khi sử dụng mạng ngữ nghĩa để biểu diễn tri thức, người ta phải xây dựng các phép toán tương ứng.

1. **Biểu diễn tri thức nhờ ngôn ngữ nhân tạo.**

Nói chung, theo quan điểm của người sử dụng, ngôn ngữ tự nhiên sẽ là phương cách thuận tiện nhất để giao tiếp với một hệ cơ sở tri thức, không những đối với người quản trị hệ thống (tư cách chuyên gia), mà còn đối với người sử dụng cuối. Hiện nay đã có những hệ cơ sở tri thức có khả năng đối thoại trên ngôn ngữ tự nhiên (thông thường là tiếng Anh) nhưng chỉ hạn chế trong lĩnh vực ứng dụng chuyên môn của hệ cơ sở tri thức. Hình dưới đây thể hiện một đơn vị tri thức (luật) trong hệ cơ sở tri thức MYCIN dùng để chẩn đoán các bệnh virut. Cột bên trái là một luật được viết bằng tiếng Anh, cột bên phải là mã hoá nhân tạo của luật đó.

* + 1. **Các phương pháp suy diễn**

1. **Suy diễn tiến.**

Đi từ giả thiết ta sẽ suy diễn dẫn tới kết quả mong muốn. Là quá trình suy diễn bắt đầu từ tập sự kiện đã biết, rút ra những sự kiện mới và cứ như vậy cho đến khi có được sự kiện cần chứng minh hoặc không có. Quá trình suy diễn tiến là quá trình xem xét các luật, với mỗi luật ta xét phần điều kiện (ở vế trái) tới phần kết luận (ở vế phải) và khi mà tất cả các điều kiện của luật đều thỏa mãn thì ta suy ra sự kiện trong phần kết luận. Chính vì lẽ đó mà có tên là suy diễn tiến. Trong mỗi bước của thủ tục, người ta xét một luật trong tập luật. So sánh mỗi điều kiện (ở vế trái) của tập luật với các sự kiện trong cơ sở sự kiện, nếu tất cả các điều kiện của luật được tho mãn thì sự kiện trong phần kết luận được xem là sự kiện được suy ra. Nếu sự kiện này là sự kiện mới (không có trong bộ nhớ làm việc) thì nó được đưa vào bộ nhớ làm việc. Quá trình trên cứ lặp lại cho đến khi nào không có luật nào sinh ra sự kiện mới. Quá trình suy diễn tiến không định hướng tới giải quyết một vấn đề nào cả, không hướng tới tìm ra câu trả lời cho một câu hỏi nào cả. Suy diễn tiến chỉ là quá trình suy ra các sự kiện mới từ các sự kiện có trong bộ nhớ làm việc.

**Ưu điểm**

* Ưu điểm chính của suy diễn tiến là làm việc tốt khi bài toán về bản chất đi thu thập thông tin rồi thấy điều mình cần suy diễn.
* Suy diễn tiến cho ra số lượng lớn các thông tin từ các thông tin ban đầu. Nó sinh ra nhiều thông tin mới.
* Suy diễn tiến là tiếp cận lý tưởng với các loại bài toán cần giải quyết các nhiệm vụ như lập kế hoạch, điều hành, điều khiển và diễn dịch.

**Nhược điểm**

* Hệ thống suy diễn không cảm nhận được một vài thông tin quan trọng. Hệ thống hỏi các câu hỏi có thể hỏi mà không biết rằng chỉ cần một ít câu thôi để đi đến kết luận được.
* Hệ thống có thể hỏi cả câu hỏi không liên quan. Có thể các câu trả lời cũng quan trọng, nhưng làm người dùng lúng túng khi phải trả lời các câu chẳng dính dáng đến chủ đề.

1. **Suy diễn lùi.**

Ý tưởng của suy diễn lùi là sẽ xuất phát từ kết luận ta sẽ suy diễn dẫn tới một điều gì đó đúng trong giả thiết đã cho. Là quá trình xuất phát từ sự kiện cần chứng minh và thay vào đó là những sự kiện ở vế trái của một luật có vế phải là sự kiện cần chứng minh. Quá trình này được thực hiện cho đến khi đưa về các sự kiện là tập sự kiện con của tập sự kiện giả thiết. (Nghĩa là: để đưa ra kết luận b, ta thử tìm tất cả các luật có dạng: a1 ⋀ ....⋀ an ⟹ b, để có b, phải đưa ra các kết luận a1,...,an. Quá trình xác định ai cũng tương tự như đối với b, nếu đến một lúc nào đó phát hiện được rằng có một ai nào đó không dẫn xuất được từ các giả thiết thì quay lui sang các luật sản xuất khác sinh ra b có dạng b1⋀....⋀bm ⟹ b. Ngược lại, nếu mọi ai đều dẫn xuất được giả thiết thì quá trình dẫn xuất ra b là đúng).

**Ưu điểm**

* Một trong các ưu điểm chính của suy diễn lùi là phù hợp với bài toán đưa ra giả thuyết rồi xem hiệu quả của giả thuyết đó có đúng không.
* Suy diễn lùi tập trung vào đích đã cho. Nó tạo ra một loạt câu hỏi chỉ liên quan đến vấn đề đang xét, đến hoàn cảnh thuận tiện đối với người dùng.
* Khi suy diễn lùi muốn suy diễn cái gì đó từ các thông tin đã biết, nó chỉ tìm trên một phần của cơ sơ tri thức thích đáng đối với bài toán đang xét.

**Nhược điểm**

* Nhược điểm cơ bản của loại suy diễn này là nó theo đuổi một dòng suy luận, thay vì đúng ra phải dừng ở đó mà sang chỗ khác. Tuy nhiên, người ta có thể dùng nhân tố tin cậy và các luật meta để khắc phục hiện tượng này.
* So với suy diễn lùi, suy diễn tiến đơn giản hơn. Song quá trình suy diễn tiến phải huy động mọi luật có thể áp dụng được, mà không lưu ý đến liệu luật có liên quan đến kết luận mong muốn hay không. Do vậy, đối với các cơ sở tri thức lớn với số luật ngày càng tăng, thuật giải này dẫn tới bùng nổ tổ hợp. Ngược lại, suy diễn lùi phức tạp hơn, nhưng có ưu điểm là chỉ chọn những luật hướng tới đích đặt ta. Về thực chất cơ chế suy diễn lùi được cài đặt ở đây tương ứng với tìm kiếm sâu trên đồ thị Và/Hoặc biểu diễn tập luật.

1. **Nghiên cứu hiện trạng các cổng thông tin nông sản**

* **Giới thiệu cổng thông tin điện tử**

Thuật ngữ tin học này xuất hiện từ năm 1998, tuy nhiên đến nay vẫn chưa thể có một định nghĩa hoàn chỉnh. Tuy nhiên ta có thể hiểu:

“Cổng thông tin (*portal*) là một hoặc một nhóm trang web mà từ đó người truy cập có thể dễ dàng truy xuất các trang web và các dịch vụ thông tin khác trên mạng máy tính.”

Ngày nay, cổng thông tin được ứng dụng khá rộng rãi trong nhiều lĩnh vực và công nghệ khác nhau. Tuy nhiên chúng đều phải có chung một số tính năng cơ bản. Các tính năng này như là một tiêu chuẩn để phân biệt giữa các công thông tin với một website tổng hợp tin tức, hay một hệ thống quản lý nội dung web CMS.

* **Các tính năng cơ bản**

**Khả năng cá nhân hoá** (*Customization hay Personalization*): cho phép thiết đặt các thông tin khác nhau cho các loại đối tượng sử dụng khác nhau theo yêu cầu. Tính năng này dựa trên hoạt động thu thập thông tin về người dùng và cộng đồng người dùng, từ đó cung cấp các thông tin chính xác tại thời điểm được yêu cầu.

**Tích hợp nhiều loại thông tin** (*Content aggregation*): cho phép xây dựng nội dung thông tin từ nhiều nguồn khác nhau cho nhiều đối tượng sử dụng. Sự khác biệt giữa các nội dung thông tin sẽ được xác định qua các ngữ cảnh hoạt động của người dùng (user-specific context), ví dụ như đối với từng đối tượng sử dụng sau khi thông qua quá trình xác thực thì sẽ được cung cấp các thông tin khác nhau, hoặc nội dung thông tin sẽ được cung cấp khác nhau trong quá trình cá nhân hoá thông tin.

**Xuất bản thông tin** (*Content syndication*): thu thập thông tin từ nhiều nguồn khác nhau, cung cấp cho người dùng thông qua các phương pháp hoặc giao thức (protocol) một cách thích hợp. Một hệ thống xuất bản thông tin chuyên nghiệp phải có khả năng xuất bản thông tin với các định dạng đã được quy chuẩn, ví dụ như RDF (Resource Description Format), RSS (Realy Simple Syndication), NITF (News Industry Text Format) và NewsXML. Ngoài ra, các tiêu chuẩn dựa trên XML cũng phải được áp dụng để quản trị và hiển thị nội dung một cách thống nhất, xuyên suốt trong quá trình xuất bản thông tin. Các tiêu chuẩn dựa trên XML này cho phép đưa ra giải pháp nhanh nhất để khai thác và sử dụng thông tin trên các Web site khác nhau thông qua quá trình thu thập và bóc tách thông tin với các định dạng đã được quy chuẩn.

**Hỗ trợ nhiều môi trường hiển thị thông tin** (*Multidevice support*): cho phép hiển thị cùng một nội dung thông tin trên nhiều loại thiết bị khác nhau như: màn hình máy tính (PC), thiết bị di động (Mobile phone, Wireless phone, PDA), sử dụng để in hay cho bản fax…. một cách tự động bằng cách xác định thiết bị hiển thị thông qua các thuộc tính khác nhau. Ví dụ: cùng một nội dung đó, khi hiển thị trên màn hình máy tính thì sử dụng HTML, nhưng khi hệ thống xác định được thiết hiển thị là PDA hay mobile phone, hệ thống sẽ loại bỏ các ảnh có trong nội dung và tự động chuyển nội dung đó sang định dạng WML (Wireless Markup Language) để phù hợp cho việc hiển thị trên màn hình của thiết bị di động.

**Khả năng đăng nhập một lần** (*Single Sign On - SSO*): cho phép dịch vụ xuất bản thông tin hoặc các dịch vụ khác của portal lấy thông tin về người dùng khi hoạt động mà không phải yêu cầu người dùng phải đăng nhập lại mỗi khi có yêu cầu. Đây là một tính năng rất quan trọng vì các ứng dụng và dịch vụ trong portal sẽ phát triển một cách nhanh chóng khi xuất hiện nhu cầu, mà các ứng dụng và dịch vụ này tất yếu sẽ có các nhu cầu về xác thực hoặc truy xuất thông tin người dùng.

**Quản trị portal** (*Portal administration*): xác định cách thức hiển thị thông tin cho người dùng cuối. Tính năng này không chỉ đơn giản là thiết lập các giao diện người dùng với các chi tiết đồ hoạ (look-and-feel), với tính năng này, người quản trị phải định nghĩa được các thành phần thông tin, các kênh tương tác với người sử dụng cuối, định nghĩa nhóm người dùng cùng với các quyền truy cập và sử dụng thông tin khác nhau.

**Quản trị người dùng** (*Portal user management*): cung cấp các khả năng quản trị người dùng cuối, tuỳ thuộc vào đối tượng sử dụng của portal. Tại đây, người sử dụng có thể tự đăng ký trở thành thành viên tại một công thông tin công cộng (như Yahoo, MSN…) hoặc được người quản trị tạo lập và gán quyền sử dụng tương ứng đối với các công thông tin doanh nghiệp. Mặt khác, tuỳ vào từng kiểu portal mà số lượng thành viên có thể từ vài nghìn tới hàng triệu.

Cổng thông tin tri thức là cổng thông tin tích hợp với các kỹ thuật suy diễn để hệ thống có thể hỗ trợ cho các hoạt động tư vấn. Các hoạt động này cần thiết cho cổng thông tin tri thức nông sản.

* **Hiện trạng các cổng thông tin nông sản**

Cổng thông tin tri thức được áp dụng tại rất nhiều nơi trên thế giới và cho rất nhiều các lĩnh vực, ngành nghề nghiên cứu khác nhau như: cổng thông tin tri thức khoa học, cổng thông tin tri thức nghiên cứu các thông tin trong trường đại học, hay như cổng thông tin tri thức y tế cho một thư viện y tế của Đức [1-3].

Riêng về lĩnh vực nông nghiệp, một số nước trên thế giới quan tâm tới nông nghiệp cũng đã áp dụng các cổng thông tin nói chung hay cổng thông tin tri thức nói riêng để hỗ trợ và đã đạt được những thành tựu rất rõ rệt. Các vấn đề trên hầu hết đều đã được giải quyết bởi hệ thống cổng thông tin tri thức ở các nước phát triển trên thế giới như Mỹ, Châu Âu và Ấn Độ...

Cổng thông tin **http://www.indiaagristat.com**/ là nguồn số liệu thông kê cho nông nghiệp Ấn Độ [12]. Hệ thống cung cấp thông tin các số thống kê trên các lĩnh vực như giáo dục nông nghiệp, xuất khẩu nông nghiệp, điều tra dân số nông nghiệp, giá nông nghiệp, bảo hiểm nông nghiệp, chăn nuôi, tiếp thị nông nghiệp, và tất cá các số liệu thông kê khác liên quan đến nông nghiệp của Ấn Độ.

**http://www.isapindia.org**/ trang web xã hội về kinh doanh nông nghiệp chuyên nghiệp (ISAP - Indian society of agribusiness professional), cộng đồng này giúp đỡ người nông dân và giải quyết sự phân chia thu nhập nông thôn đô thị. Đây là một mạng lưới các đồng minh khu vực nông nghiệp và các chuyên gia ở Ấn Độ và các nước đang phát triển [3,4].

**http://agricoop.nic.in**/ - trang web này sẽ chú trọng vào các ngành nông nghiệp, việc làm, cơ hội, khu vực nông nghiệp và cơ sở hạ tầng [5,6]. Nó cũng hiển thị mạng lưới nhà máy thông tin bảo vệ và nhiệm vụ an ninh lương thực quốc gia.

Nhận xét về các cổng thông tin không được liên kết với nhau, dẫn tới các thông tin nhiều nhưng thực sự việc tìm kiếm trên đó còn gặp nhiều khó khăn, không thân thiện với người nông dân. Hơn nữa, việc liên kết giữa người nông dân – doanh nghiệp – các chính sách nông nghiệp vẫn chưa thực sự rõ ràng.

Đối với khu vực Đông Nam Á cũng tập trung nhiều nước phát triển nông nghiệp như Thái Lan, Indonesia, Việt Nam...Vì vậy, Nông nghiệp tại các nước này khá phát triển nhưng việc quản lý thông tin nông sản, nông nghiệp vẫn chưa được tập trung. Đặc biệt, các cổng thông tin về nông nghiệp hầu như rất ít và không được phổ biến [7,8].

Tại Việt Nam, một nước đang phát triển, nông nghiệp chiếm tỷ trọng đa số với gần 80% dân số làm nông nghiệp và trên 50% diện tích đất nông nghiệp. Do đó, việc phát triển nông nghiệp đặc biệt cần được chú trọng [13]. Với việc phát triển mạnh mẽ của công nghệ thông tin tại Việt Nam, rất nhiều các cổng thông tin đã được xây dựng như: cổng thông tin điện tử chính phủ, cổng thông tin đào tạo... Hầu như mỗi bộ ngành, lĩnh vực đều có một cổng thông tin, và cổng thông tin nông nghiệp. Cổng thông tin cung cấp các chính sách nông nghiệp, tin tức về nông sản nói chung **http://www.agroviet.gov.vn/,** nhưng các thông tin này lại thiếu tính liên kết, chỉ mang tính chất đưa tin thông thường.

Ngoài ra, còn có cổng thông tin hỗ trợ việc tra cứu, tìm kiếm thông tin nông sản nông nghiệp **http://mard.gov.vn/** qua các thời kỳ từ 1995 – 2015... nhưng khả năng tìm kiếm không mạnh mẽ và khó tìm được thông tin chính xác.

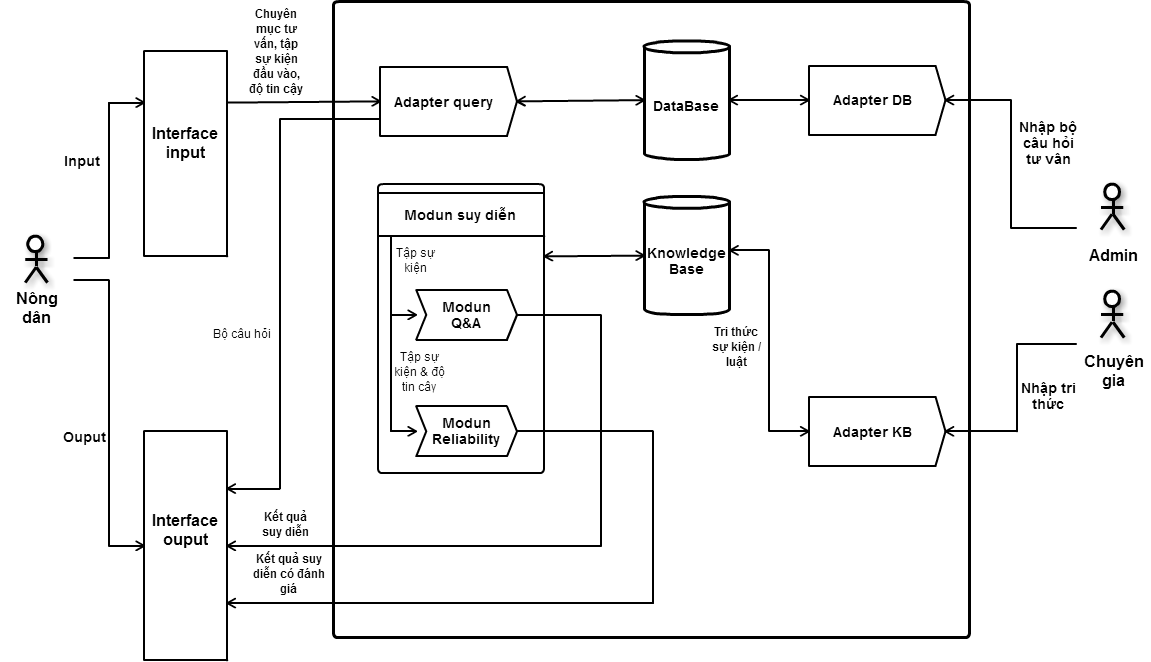
1. **Tổng kết chương**

Chương giới thiệu tổng quan về cổng thông tin, cổng thông tin tri thức và kiến thức cơ bản về hệ cơ sở tri thức. Chương này cũng nêu khái quát các cổng thông tin nông sản trên thế giới để hình thành cơ sở lý thuyết cho chương 2 về đề xuất mô hình cổng thông tin tri thức nông sản.

**CHƯƠNG 2: MÔ HÌNH CỔNG THÔNG TIN TRI THỨC NÔNG SẢN**

1. **Kiến trúc hệ thống cổng thông tin tri thức**

Kiến trúc hệ thống cổng thông tin tri thức mô tả như hình.



**Hình 2: Sơ đồ kiến trúc hệ thống**

Mô hình kiến trúc hệ thống cổng thông tintri thức tư vấn nông sản nông nghiệp cho người nông dân, nhà chính sách và các doanh nghiệp.

**Đầu vào**: Các tiêu chí cho một vấn đề nông nghiệp cần hỏi từ người nông dân, hoặc doanh nghiệp.

**Đầu ra**: Các hành động được tư vấn bởi chuyên gia.

Bộ máy suy diễn này gồm hai module chính, mỗi module sẽ giải quyết một tập các vấn đề từ đơn giản tới phức tạp của người dùng.

* Module Q&A: sẽ giải quyết vấn đề suy diễn với tập sự kiện cần tư vấn đầu vào đơn giản, từ đó đưa ra tập kết quả suy diễn tương ứng..
* Module Reliability: sẽ giải quyết vấn đề suy diễn khi tập sự kiện cần tư vấn đầu vào chưa chắc chắn của người dùng, từ đó đưa ra được tập kết quả suy diễn tương ứng.

1. **Diễn giải**
   1. **Module Q&A**

***Bài toán đặt ra:*** Khi người dùng đưa ra các câu hỏi, hệ thống cần đưa ra được kết quả phù hợp dựa trên những tri thức chuyên gia.

*Đầu vào*: câu hỏi của người dùng sẽ được phân tách thành những tiêu chí nhỏ hơn.

*Đầu ra*: đưa ra các kết luận từ những luật phù hợp trong bộ dữ liệu Knowledge Base với các tiêu chí đó.

***Phương pháp đánh giá:***

* Với mỗi câu hỏi của người dùng, hệ thống sẽ có một bộ các tiêu chí cho trước. Từ đó người dùng sẽ chọn các tiêu chí đánh giá và chọn câu trả lời cho mỗi tiêu chí.
* Mỗi tiêu chí sau khi được người dùng đánh giá sẽ trở thành một tiêu chí đầu vào **tc**k.
* Chúng ta sẽ xây dựng một tập cơ sở các luật **Rj** (1<= j <= n)này với các tiêu chí điều kiện tương ứng với các kết luận đầu ra:
  + **R1 : dk1 ^ dk2 ^....^ dkm  kq1**
  + **R2 : dk1 ^ dk2 ^....^ dkm  kq2**
  + **.....**
  + **Rn : dk1 ^ dk2 ^....^ dkm  kqn**

***Phương pháp suy diễn:***

Có thể thấy cách suy diễn đơn giản nhất là chúng ta trực tiếp so sánh từng điều kiện **dki** trong luật **Rj** với những tiêu chí **tck** đầu vào mà người dùng đưa vào. Nếu tất cả các điều kiện **dki** của luật **Rj** đều thỏa mãn tiêu chí **tck** thì kết quả của luật **Rj** sẽ được đưa ra làm kết quả đầu ra.

**Ví dụ 1:** người nông dân muốn tư vấn về giống lúa, hệ thống sẽ phân tích câu hỏi này cần các tiêu chí như sau:

* Tiêu chí năng suất – tc1 > 5tấn/ha
* Tiêu chí giá thành bán – tc2 > 10.000đ/kg
* Tiêu chí thời gian thu hoạch – tc3 < 6 tháng
* Tiêu chí sâu bệnh – tc4 > 5 loại sâu bệnh
* Module **Q&A** sẽ đưa ra danh sách tất cả các giống lúa đáp ứng được các tiêu chí mà người nông dân muốn tìm kiếm.

***Thuật toán 1:***

**B1:** Gọi tập tiêu chí đầu vào là **TC** = {tc1, tc2, ..., tck}; **R** là tập luật cần kiểm tra; **KQ** là tập các kết quả đầu ra = Null;

**B2:** Kiểm tra từng luật **Rj**, so sánh các **dkj** có thuộc tập **TC** không, nếu tất cả **dki** (tức mọi điều kiện ở vế trái của luật Rj)thỏa mãn thì đưa **kqj**vào tập kết quả đầu ra **KQ.**

**B3:** Đưa ra tập các kết quả **kqj.**

|  |
| --- |
| ***Giải thuật:*** |
| Procedure **modun1**;  Begin  TC={tc1, tc2, ..., tck};  R = luật của KB;  KQ = {};  while R<>∅ do  begin  r := get(R);  //nếu mọi giá trị vế trái của luật r thuộc TC  if ∀left(r) ∈ TC then  KQ = KQ + r;  R:=R \ {r};  end  end;  If KQ<>∅ then write (“Các kết quả là:” + KQ)  Else write (“Không tìm thấy đáp án”);  End |

**Đánh giá thuật toán 1:**

Phương pháp này khá nhanh và đơn giản, tuy nhiên lại đưa ra quá nhiều kết quả đầu ra thừa khiến người dùng không thể xác định được kết quả chính xác.

Bởi, nếu những luật R có tập điều kiện ít hơn số tiêu chí đầu vào thì kết quả đưa ra có thể không chính xác do bị sót tiêu chí đánh giá.

**Ví dụ 2:** nếu cũng với câu hỏi như ở Ví dụ 1 nhưng trong tập luật cho sẵn như sau:

* R1: năng suất tc1 > 5tấn/ha + giá thành bán tc2 > 10.000đ/kg  giống A
* R2: năng suất tc1 > 5tấn/ha + giá thành bán tc2 > 10.000đ/kg  giống B
* R3: thời gian thu hoạch tc3 < 6 tháng + tiêu chí giá thành bán tc2 > 10.000đ/kg  giống C

thì thuật toán sẽ đưa ra cả 3 giống A,B,C; nhưng thực tế có thể giống A không đáp ứng được tất cả các tiêu chí của người nông dân đưa ra vì giống A không đáp ứng được tiêu chí sâu bệnh.

Một phương pháp suy diễn tốt cần đưa ra ít kết quả nhất nhưng cũng đảm bảo độ chính xác của nó.

Để cải tiến Module này, chúng tôi đã áp dụng thêm luật suy diễn **Modus Pones[18]** tổng quát. Trong mỗi bước của thủ tục lập luận, ta xét một luật trong cơ sở luật. So sánh mỗi điều kiện **dki** của luật **Rj** với các sự kiện, tiêu chí **tc**k nếu tất cả điều kiện của luật đều thỏa mãn thì sự kiện trong phần kết luận của luật chính là sự kiện được suy ra. Nếu sự kiện này là sự kiện mới, thì nó được đặt vào bộ nhớ làm việc.

***Thuật toán Modus Pones:***

**B1:** Gọi tập tiêu chí đầu vào là **TC** = {tc1, tc2, ..., tck}; **R** là tập luật cần kiểm tra; **KQ** là tập các kết quả đầu ra = Null;

**B2:** Xét từng luật rj có dạng: dk1 ^ dk2 ^....^ dkm  kqj và dki ∈ T ∀i =1, m (tập luật có được từ **modun1**).

* Nếu kqi không nằm trong TC thì cho kqi vào tập TC và gọi lại **modun1** với tập TC mới, để tìm tập luật R thỏa mãn mới.

**B3:** Quá trình lặp lại đến khi duyệt hết tập luật R mà không sinh thêm sự kiện mới (tức kqi  không nằm trong TC).

**B4:** Tập kết quả kq trong tập luật R cuối cùng chính là kết quả cần đưa ra.

|  |
| --- |
| ***Giải thuật:*** |
| Procedure **modunPones**;  Begin  TC={tc1, tc2, ..., tck};  R = luật của KB;  KQ = {};  S:= modun1(R, TC);{ S: là tập luật có dạng rj : dk1 ^ dk2 ^....^ dkm  kqj và dki ∈ TC ∀i =1, m }  while S<>∅ do  begin  r := get(S);  if right(r) not ∈ TC  TC:= TC + right(r);  S:= modun1(R,TC);  if S<>∅ then KQ:=S;  else break;  else  S:= S \ {r}  end  if KQ<>∅ then write (“Các kết quả là:” + KQ)  else write (“Không tìm thấy đáp án”);  End; |

Có thể nhận thấy, việc sử dụng luật suy diễn **Modus Pones[18]** sẽ giúp chúng ta thu hẹp được số các kết quả thu được sau mỗi lần bổ sung sự kiện đầu vào và tính toán lại tập luật. Tuy nhiên, phương pháp này vẫn chưa đánh giá được đâu là kết quả đúng nhất. Chúng ta có thể dựa vào tần suất xuất hiện của các luật sau mỗi lần tính toán lại tập luật. Luật chính xác nhất là luật xuất hiện lặp lại nhiều lần nhất trong tập luật cuối cùng.

Để đánh giá được tần suất của các luật, chúng ta sẽ cải tiến thuật toán của **modun1** - modun tính toán tập luật thỏa mãn bộ tiêu chí đầu vào.

* 1. **Module Reliability**

***Bài toán đặt ra:*** Trong quá trình tư vấn, với những câu hỏi của người dùng, các tư vấn của chuyên gia không có tính chính xác hoàn toàn. Tương tự, các câu hỏi mà hệ thống đưa ra từ module Q&A, người dùng không thể chắc chắn về các tiêu chí đã chọn. Vì thế bài toán đặt ra, hệ thống phải cho phép nhập liệu với các luật và tiêu chí với độ tin cậy nhất định mà vẫn đưa ra được tư vấn cho người dùng tương ứng với câu hỏi và độ tin cậy.

*Đầu vào*: câu hỏi của người dùng cũng sẽ được phân tách thành những tiêu chí nhỏ hơn, và người dùng sẽ phải chọn thêm độ tin cậy cho những tiêu chí mà mình lựa chọn nếu họ không chắn chắn về tiêu chí đó.

*Đầu ra*: đưa ra các kết luận từ những luật phù hợp với các tiêu chí đó cùng với độ tin cậy của kết quả.

**Ví dụ 3:** chúng ta tiếp tục với Ví dụ 1, trường hợp này người dùng sẽ phải chọn thêm tính chắc chắn cho các tiêu chí đã chọn như sau:

* Tiêu chí năng suất – tc1 > 5tấn/ha : chắn chắn phải đạt 100%
* Tiêu chí giá thành bán – tc2 > 10.000đ/kg: không cần thiết lắm 70%
* Tiêu chí thời gian thu hoạch – tc3 < 6 tháng: có thể nhiều hơn 1 chút 90%
* Tiêu chí sâu bệnh – tc4 > 5 loại sâu bệnh: phải chống chịu được nhiều sâu bệnh 100%
* Module **Reliability** sẽ đưa ra danh sách tất cả các giống lúa đáp ứng được các tiêu chí mà người nông dân muốn tìm kiếm cùng với tỉ lệ tin cậy cho từng kết quả

Trường hợp này, chúng ta có thể sử dụng cơ chế suy diễn kết hợp với lý thuyết chắc chắn Stanford để giải quyết giúp các chuyên gia khi tư vấn đưa ra được tính chính xác của kết luận.

***Phương pháp đánh giá:***

* Với mỗi câu hỏi của người dùng, hệ thống cũng sẽ có một bộ các tiêu chí cho trước. Từ đó người dùng sẽ chọn các tiêu chí (P) và chọn mức độ chắc chắn cho tiêu chí đó (CF).
* Từ đó, ta có bộ các tiêu chí và độ tin cậy cho tiêu chí đó p = CF[P].
* Tương ứng như thế, ta cũng phải đưa ra độ tin cậy cho mỗi một luật R mà chuyên gia đưa ra t = CF[R].
* MB: độ đo tin cậy, giá trị [0; 0.25; 0.5; 0.75; 1].
* MD: độ đo thiếu tin cậy, giá trị [0; 0.25; 0.5; 0.75; 1].
* CF: hệ số tin cậy, giá trị [-1;1]; CF = MB - MD
* Từ lý thuyết chắc chắn Stanford, chúng ta có thể đưa ra được độ chắc chắn cho những kết luận đưa ra.

VD với một luật R1: if P then Q

Thì CF[Q] = CF[R1] \* CF[P]: là độ chắc chắn của kết luận (1)

***Sử dụng luật của Standford:***

|  |
| --- |
| Luật ghép:  CF(e1 and e2) = MIN(CF(e1), CF(e2))  CF(e1 or e2) = MAX(CF(e1), CF(e2)) |

***Phương pháp suy diễn tiến:***

Tiếp tục phương pháp suy diễn ở Modun Q&A, modun này chúng ta bổ sung và đưa vào độ chắc chắn cho từng luật R trong cơ sở dữ liệu luật.

Dựa vào những tiêu chí và mức độ chính xác tương ứng mà người dùng đưa vào, modun sẽ tìm ra những luật có điều kiện, tiêu chí phù hợp được tính toán từ modun Q&A. Sau đó, ta dựa vào lý thuyết chắc chắn Stanford để tính ra độ chắc chắn cho từng kết luận, hoặc có thể chỉ đưa ra kết luận có độ chắc chắn cao nhất.

***Thuật toán:***

**B1:** Gọi tập tiêu chí đầu vào là **TC** = {(tc1, p1), (tc2, p2) , ..., (tck, pk)}; **R** là tập luật cần kiểm tra; **KQ** là tập các kết quả đầu ra = Null;

**B2:** Đưa bộ tiêu chí TC vào modun Q&A cải tiến để tìm ra bộ luật phù hợp với tiêu chí đầu vào nhất.

**B3:** Tính độ chắc chắn của kết luận trong từng luật theo công thức Stanford, đưa ra kết luận có độ chắc chắn cao nhất.

|  |
| --- |
| ***Giải thuật đánh giá độ tin cậy*** |
| Procedure **modun2**;  Begin  TC={(tc1,p1), (tc2,p2) ,..., (tck,pk)};//tiêu chí đầu vào có trọng số tin cậy p  R = {(r1,t1), (r2,t2), ..., (rm, tm)}; //luật của KB có trọng số tin cậy t  KQ = **modunPonesQA**(TC); //tập luật thỏa mãn tiêu chí  MaxCF = 0;  while KQ<>∅ do  begin  kq := get(KQ);  cf = tính độ chắc chắn của right(kq)  if cf > MaxKQ then MaxCF = cf; maxKQ = kq;  R:=R \ {r};  end  end;  write (“Các kết quả là:” + maxKQ);  End |

**Đánh giá thuật toán**

Đánh giá độ tin cậy sử dụng lý thuyết chắc chắn Stanford cơ bản đã giải quyết được vấn đề của bài toán đưa ra là tính được độ tin cậy cho từng kết quả.

**Ví dụ 4:** từ dữ liệu trong Ví dụ 3, ta giả sử hệ thống trả ra kết quả:

* Giống A đáp ứng được các tiêu chí của bạn với độ chính xác 90%.
* Giống B đáp ứng được các tiêu chí của bạn với độ chính xác 80%.
* Giống C đáp ứng được các tiêu chí của bạn với độ chính xác 70%.

Dễ dàng thấy rằng với những kết quả đưa ra, người sử dụng sẽ chọn ngay giống A.

Tuy nhiên, chúng ta có thể nhận thấy hệ thống chỉ đưa ra được những lời tư vấn theo dạng liệt kê sắp xếp theo độ giảm dần của độ tin cậy mà không có một lời khuyên ở dạng phân tích chuyên sâu hơn.

**Ví dụ 5**: với 3 giống A,B,C mà hệ thống đưa ra, chúng ta có thể phân tích thêm và đưa ra 1 lời khuyên ở dạng tổng quát như sau: “Giống A rất tốt về năng suất tuy nhiên nếu bạn băn khoăn về giá thành thì giống C sẽ cho giá thành cao hơn, còn giống B có khả năng chống chịu sâu bệnh rất tốt nếu phải canh tác ở vùng nhiều sâu bệnh”.

Chúng ta hoàn toàn có thể giải quyết được vấn đề này bằng cách đưa các kết luận trả về từ **B3** vào đầu vào của suy diễn cấp hai và đưa ra kết luận cuối cùng như sau.

1. **Ứng dụng suy diễn tri thức trong cổng thông tin**

Cổng thông tin tri thức nông sản là nguồn tri thức của các chuyên gia. Đối với chuyên gia giỏi, họ thường có ít thời gian trong tư vấn nông nghiệp cho cộng đồng. Cổng thông tin nông sản là một cổng thông tin trực tuyến cung cấp các nguồn thông tin, chuỗi nông sản cung ứng cho nông sản Việt Nam nói chung và các tỉnh nói riêng. Ứng dụng các kỹ thuật suy diễn trong cổng thông tin tri thức nhằm tái tạo các tri thức của chuyên gia, sử dụng lại giảm giá thành tư vấn và tăng khả năng tương tác của chuyên gia với cộng đồng.

Nghiên cứu tích hợp các kỹ thuật suy diễn trong cổng thông tin tri thức nông sản là sự phát huy khả năng đột phá kỹ thuật dựa trên cơ sở tri thức. Để nâng cao chất lượng tư vấn được tốt, mô hình suy diễn dựa vào tri thức chuyên gia sẽ đem lại những ứng dụng tư vấn tối ưu cho các mô hình tư vấn dịch vụ nông sản Việt Nam hiện nay.

1. **Tổng kết chương**

Trong chương này, ta đã xây dựng được kiến trúc hệ thống cơ bản của cổng thông tin tri thức và đặt ra những bài toán suy diễn và giải quyết những bài toán đó.

Chương 2 đã giải quyết hai bài toán suy diễn cần giải quyết tương ứng với hai module trong hệ thống tri thức, đó là:

* Module Q&A: giải quyết bài toán suy diễn tư vấn từ tập câu hỏi sự kiện của người dùng và những luật tri thức của chuyên gia.
* Module Reliability: dựa trên những đánh giá không chắc chắn của chuyên gia cho từng tập luật mà đánh giá được kết quả tư vấn cho người dùng.

Hai bài toán suy diễn được giải quyết ở chương này chính là cơ sở khoa học để có thể ứng dụng vào xây dựng hệ thống Web Portal tri thức thông minh. Chương 3 sẽ mô tả rõ chức năng và thiết kế cụ thể của hệ thống Portal tri thức thông minh.

**CHƯƠNG 3: PHÂN TÍCH VÀ THIẾT KẾ HỆ THỐNG**

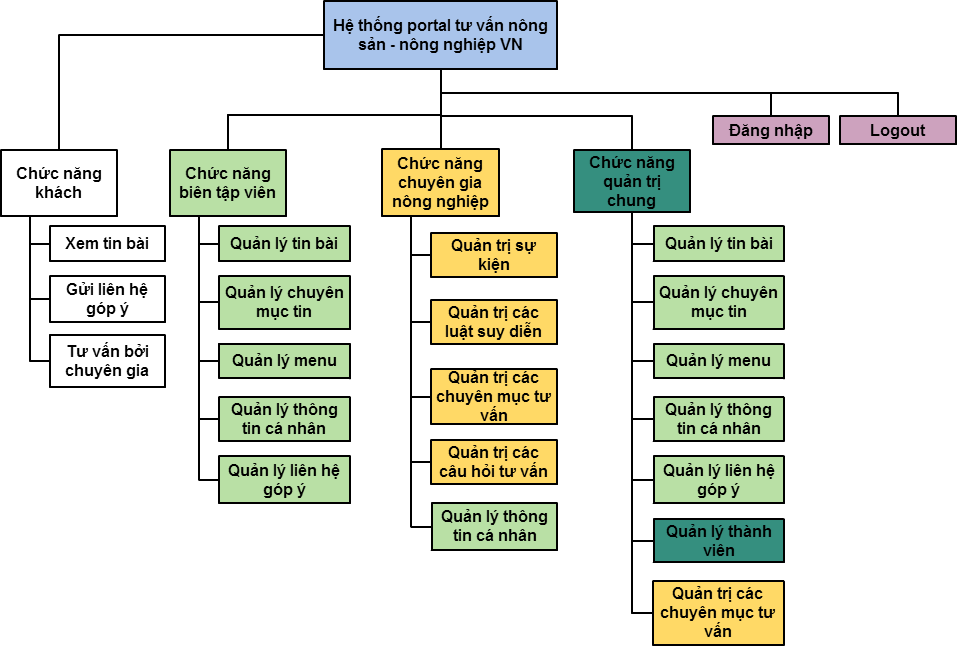
1. **Giới thiệu nghiệp vụ và công tác điều hành cổng thông tin**
   1. **Yêu cầu chức năng**

Hệ thống cần đáp ứng được yêu cầu của một website portal cơ bản với những chức năng:

* Đăng nhập và quản trị tài khoản cá nhân.
* Đưa nội dung trực tuyến lên trên internet nhanh chóng và quản lý được nội dung đó.
* Quản lý được các tài khoản user sử dụng hệ thống.
* Thay đổi bố cục tin bài, cách sắp xếp nội dung.
* Tiếp nhận các ý kiến phản hồi của người dùng.

Ngoài ra, hệ thống cần giải quyết được mục tiêu tư vấn được kiến thức nông nghiệp cho người nông dân:

* Đưa ra câu trả lời cho những vấn đề của người nông dân.
* Cập nhật và quản lý được những chuyên mục hay vấn đề của người nông dân.
* Những kiến thức tư vấn cần có sự tương tác và cập nhật từ chuyên gia
  1. **Yêu cầu phi chức năng**
* Giao diện hệ thống phải đơn giản, dễ dùng.
* Những tư vấn phải có độ chính xác cao và có thể ứng dụng được.
  1. **Biểu đồ phân cấp chức năng**



**Hình 3: Biểu đồ phân cấp chức năng**

* 1. **Các tác nhân tương tác với hệ thống**

Hệ thống portal tư vấn nông sản nông nghiệp gồm các 4 tác nhân chính:

* Khách: hệ thống nhằm phục vụ chủ yếu cho các đối tượng khách là những người nông dân, những người quan tâm muốn tìm hiểu thông tin về nông nghiệp và chính sách nông nghiệp, vì thế những đối tượng này không cần đăng nhập vẫn có thể truy nhập vào hệ thống và xem các thông tin được hiển thị trên giao diện của ứng dụng và sử dụng chức năng tư vấn chuyên gia một cách đầy đủ.
* Biên tập viên: là những đối tượng cần truy cập hệ thống để quản lý thông tin trên website. Đối tượng này cần đăng nhập vào hệ thống để có thể sử dụng chức năng của biên tập viên. Các user biên tập viên sẽ do Quản trị hệ thống tạo mới.
* Chuyên gia nông nghiệp: là những đối tượng chuyên gia, có hiểu biết và kiến thức về lĩnh vực nông nghiệp. Đối tượng này cần đăng nhập vào hệ thống để sử dụng chức năng chuyên gia. Các user chuyên gia sẽ do Quản trị hệ thống tạo mới.
* Quản trị hệ thống: là đối tượng có quyền cao nhất trong hệ thống, chịu trách nhiệm quản trị thông tin trên hệ thống, trừ các thông tin liên quan tới tri thức chuyên gia.

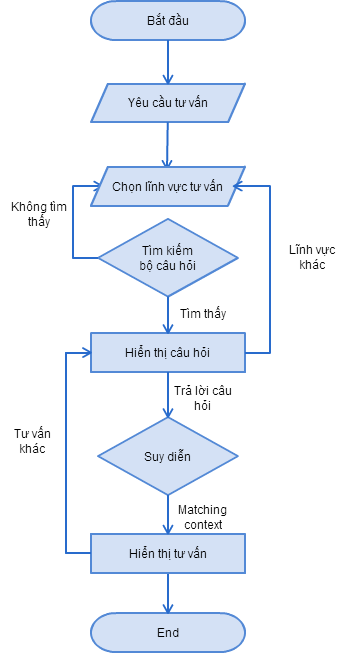
1. **Biểu đồ ca sử dụng (use case) tổng quan**



**Hình 4: Biểu đồ ca sử dụng tổng quan**

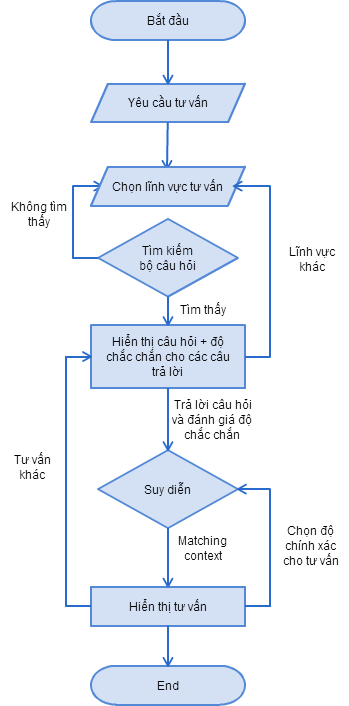
1. **Biểu đồ hoạt động tư vấn và đánh giá**

Hoạt động tư vấn suy diễn bởi chuyên gia:



**Hình 5: Biểu đồ hoạt động tư vấn suy diễn**

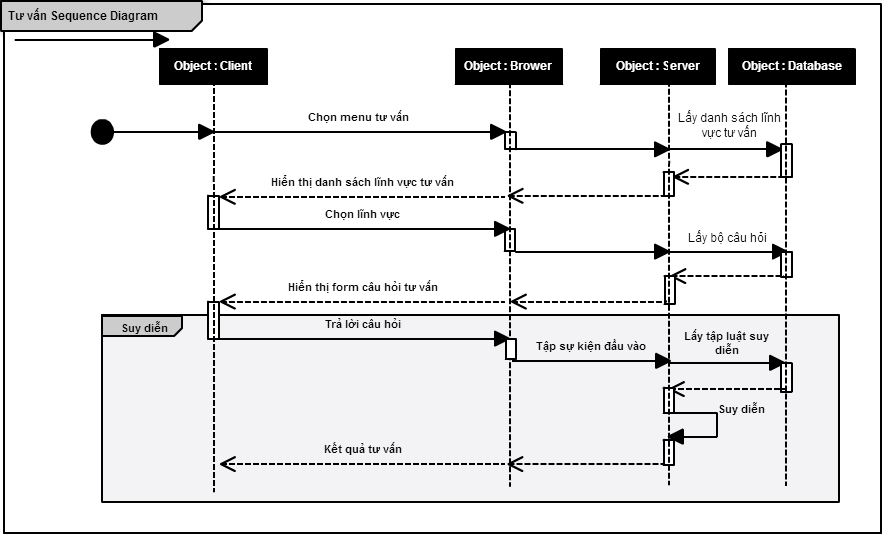
Hoạt động tư vấn đánh giá độ chính xác bởi chuyên gia:



**Hình 6: Biểu đồ hoạt động tư vấn đánh giá độ chính xác**

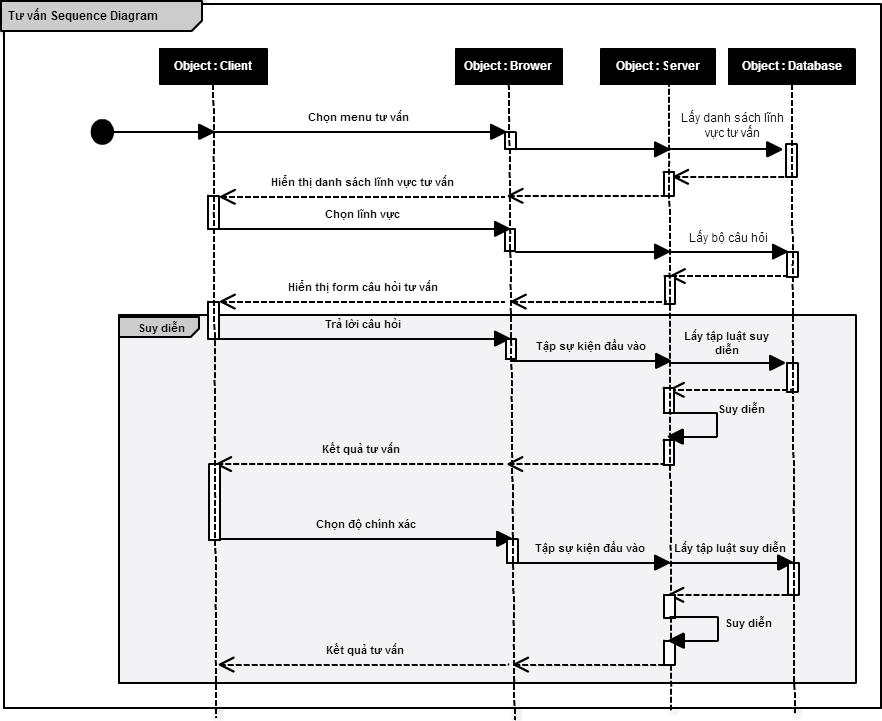
1. **Biểu đồ trình tự tư vấn và đánh giá**

Trình tự tư vấn suy diễn bởi chuyên gia:



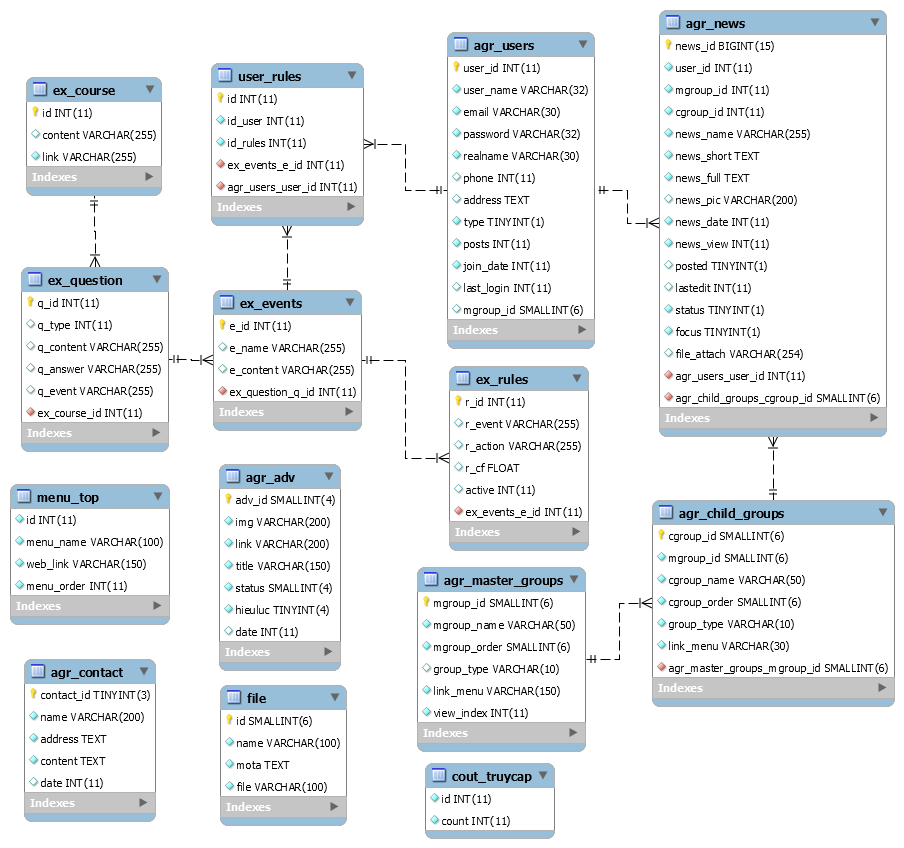
**Hình 7: Biểu đồ Trình tự tư vấn suy diễn bởi chuyên gia**

Hoạt động tư vấn đánh giá độ chính xác bởi chuyên gia:



**Hình 8: Biểu đồ Hoạt động tư vấn đánh giá độ chính xác bởi chuyên gia**

1. **Mô hình thực thể liên kết**

****

**Hình 9: Mô hình thực thể liên kết**

1. **Thiết kế cơ sở dữ liệu**

Một số bảng quan trọng cho hệ thống tư vấn nông nghiệp:

**Một số bảng quan trọng trong cơ sở dữ liệu**

|  |  |  |
| --- | --- | --- |
| STT | Tên bảng | Mô tả |
| 1 | Bảng agr\_user | Lưu trữ thông tin người dùng và vai trò |
| 2 | Bảng ex\_course | Lưu trữ thông tin các chuyên mục tư vấn |
| 3 | Bảng ex\_event | Lưu trữ các sự kiện dùng cho suy diễn |
| 4 | Bảng ex\_rule | Lưu trữ các luật của hệ thống sử dụng cho suy diễn |
| 5 | Bảng ex\_question | Lưu trữ câu hỏi sử dụng cho việc tư vấn theo từng chuyên mục |

**Bảng agr\_user** : lưu trữ thông tin người dùng và vai trò

**Bảng 1: Bảng agr\_user**

|  |  |  |  |
| --- | --- | --- | --- |
| Tên trường | Kiểu dữ liệu | Ghi chú | Giải thích |
| user\_id | int(11) | Primary key | ID tài khoản người sử dụng |
| user\_name | varchar(32) |  | Tên tài khoản người sử dụng |
| email | varchar(32) |  | Tên email người sử dụng |
| password | varchar(32) |  | Mật khẩu tài khoản người sử dụng |
| realname | varchar(32) |  | Tên người sử dụng |
| phone | int(11) |  | Số điện thoại người sử dụng |
| address | text |  | Địa chỉ người sử dụng |
| type | tinyint(1) |  | Kiểu tài khoản người sử dụng |

**Bảng ex\_course** : lưu trữ thông tin chuyên mục cần tư vấn

**Bảng 2: Bảng ex\_course**

|  |  |  |  |
| --- | --- | --- | --- |
| Tên trường | Kiểu dữ liệu | Ghi chú | Giải thích |
| id | int(11) | Primary key | ID chuyên mục |
| content | varchar(255) |  | Tên chuyên mục |

**Bảng ex\_ event** : lưu trữ các sự kiện dùng cho suy diễn

**Bảng 3: Bảng ex\_event**

|  |  |  |  |
| --- | --- | --- | --- |
| Tên trường | Kiểu dữ liệu | Ghi chú | Giải thích |
| e\_id | int(11) | Primary key | ID của sự kiện |
| e\_name | varchar(255) |  | Tên của sự kiện |
| e\_content | varchar(255) |  | Nội dung sự kiện |

**Bảng ex\_rule** : lưu trữ các luật của hệ thống sử dụng cho suy diễn

**Bảng 4: Bảng ex\_rule**

|  |  |  |  |
| --- | --- | --- | --- |
| Tên trường | Kiểu dữ liệu | Ghi chú | Giải thích |
| r\_id | int(11) | primary key | ID của luật |
| r\_event | varchar (255) |  | Giả thiết của luật, tập giả thiết được lưu cách nhau bởi dấu phẩy `,` |
| r\_action | varchar (255) |  | Kết luận của luật |
| r\_cf | float |  | Độ tin tưởng của luật |
| active | int(11) |  | Trạng thái của luật |

**Bảng ex\_question**: lưu trữ câu hỏi sử dụng cho tư vấn nông nghiệp theo từng chuyên mục

**Bảng 5: Bảng ex\_question**

|  |  |  |  |
| --- | --- | --- | --- |
| Tên trường | Kiểu dữ liệu | Ghi chú | Giải thích |
| q\_id | int(11) | Primary key | ID của câu hỏi |
| q\_type | int(11) |  | Loại chuyên mục tư vấn |
| q\_content | varchar(255) |  | Nội dung câu hỏi |
| q\_answer | varchar(255) |  | Nội dung câu trả lời, danh sách câu trả lời được lưu dưới dạng string, cách nhau bởi ký tự @ |
| q\_event | varchar(255) |  | Sự kiện của câu trả lời, danh sách sự kiện tương ứng được lưu dưới dạng string, cách nhau bởi ký tự @ |

1. **Thiết kế cơ sở tri thức và cơ chế suy diễn**

Cơ sở tri thức hệ thống sẽ được lưu trữ vào 2 bảng chính là:

* **Bảng ex\_ event:** sẽ lưu trữ tất cả các sự kiện, tiêu chí. Cấu trúc của một event như sau:

|  |  |
| --- | --- |
| **name** | Các sự kiện sẽ được đánh dấu 1 key name khi sử dụng với luật tri thức |
| **content** | Giải thích chi tiết sự kiện |

**VD:**

|  |  |
| --- | --- |
| **Name** | **Content** |
| a1 | Cây lúa bị lá vàng |
| a2 | Cây lúa có biểu hiện lá xanh và bị xoắn |

* **Bảng ex\_rule:** sẽ lưu trữ tất cả các luật tri thức. Cấu trúc của một rule như sau:

|  |  |
| --- | --- |
| events | Tập các sự kiện, tiêu chí |
| action | Kết quả của luật |
| cf | Độ tin cậy của luật (nhỏ hơn 1) |

**VD:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Rule** | **Events** | **Action** | **CF** |
| R1 | a1 | b1 | 1 |
| R2 | a2 | b2 | 1 |
| R3 | b1 | c1 | 0.7 |
| R4 | b2 | c2 | 0.8 |
| R5 | a1,b1 | d1 | 0.6 |
| R6 | b1,c2 | d2 | 0.9 |
| R7 | c1,d1 | e1 | 1 |

**Cơ chế suy diễn:**

Giả sử người dùng đưa vào 2 tiêu chí đầu vào [(a1, 0.6); (a2, 0.7)], hệ thống sẽ suy diễn trong tập luật tri thức có sẵn từ VD trên.

|  |
| --- |
| Tập tiêu chí TC = [(a1, 0.6); (a2, 0.7)]  Tập luật suy diễn từ tập TC: R = [R1; R2]  Bổ sung tiêu chí TC từ action của 2 luật tương ứng: [(b1, 1 \* 0.6); (b2, 1 \* 0.7)]   * TC = [(a1, 0.6); (a2, 0.7); [(b1, 0.6); (b2, 0.7)] |
| Tập tiêu chí TC = [(a1, 0.6); (a2, 0.7); [(b1, 0.6); (b2, 0.7)]  Tập luật suy diễn từ tập TC: R = [R1; R2; R3; R4; R5]  Bổ sung tiêu chí TC từ action của 2 luật tương ứng: [(c1, 0.7 \* 0.6); (c2, 0.8 \* 0.7); (d1, 0.6 \* MIN(0.6, 0.6))]   * TC = [(a1, 0.6); (a2, 0.7); (b1, 0.6); (b2, 0.7), (c1, 0.42); (c2, 0.56); (d1, 0.36)] |
| Tập tiêu chí TC = [(a1, 0.6); (a2, 0.7); (b1, 0.6); (b2, 0.7), (c1, 0.42); (c2, 0.56); (d1, 0.36)]  Tập luật suy diễn từ tập TC: R = [R1; R2; R3; R4; R5; R6; R7]  Bổ sung tiêu chí TC từ action của 2 luật tương ứng: [(d2, 0.9 \* MIN(0.6, 0.56)); (e1, 1 \* MIN(0.42, 0.36))]   * TC = [(a1, 0.6); (a2, 0.7); (b1, 0.6); (b2, 0.7), (c1, 0.42); (c2, 0.56); (d1, 0.36), (d2, 0.504); (e1, 0.36)] |
| Tập tiêu chí TC = [(a1, 0.6); (a2, 0.7); (b1, 0.6); (b2, 0.7), (c1, 0.42); (c2, 0.56); (d1, 0.36), (d2, 0.504); (e1, 0.36)]  Không tìm thấy TC mới => tập kết luận chính là action của tập luật R   * KQ = [(b1, 0.6); (b2, 0.7), (c1, 0.42); (c2, 0.56); (d1, 0.36), (d2, 0.504); (e1, 0.36)] |

1. **Tổng kết chương**

Chương này đã đưa ra và phân tích cụ thể và chi tiết về những chức năng chính của hệ thống, cùng cách thức, trình tự hoạt động của từng module suy diễn được nêu ở chương 2.

Với những chức năng và yêu cầu cụ thể từ hệ thống, tôi đã xây dựng một hệ thống web portal tri thức thông minh cơ bản với hai module được cài đặt sẵn với miền tri thức là nông sản nông nghiệp Việt Nam, cụ thể là tri thức về cây lúa Việt Nam.

Chương tiếp theo sẽ mô tả cài đặt và kết quả của hệ thống này.

**CHƯƠNG 4: CÀI ĐẶT CHƯƠNG TRÌNH VÀ KẾT QUẢ**

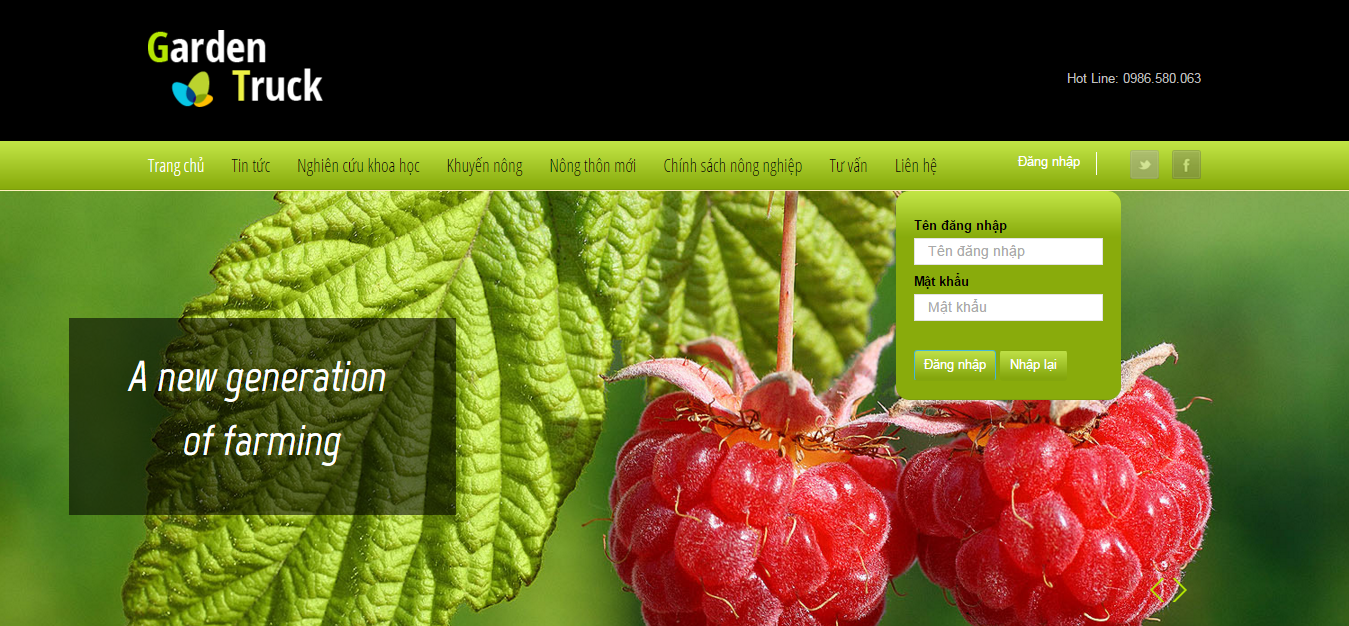
1. **Cài đặt chương trình**
   1. **Môi trường cài đặt và phát triển**

Hệ thống tư vấn được cài đặt trên môi trường lập trình và server web:

* Ngôn ngữ lập trình: PHP 5.5
* Hệ cơ sở dữ liệu: MySQL 5
* Web Server: Apache 2.4.9

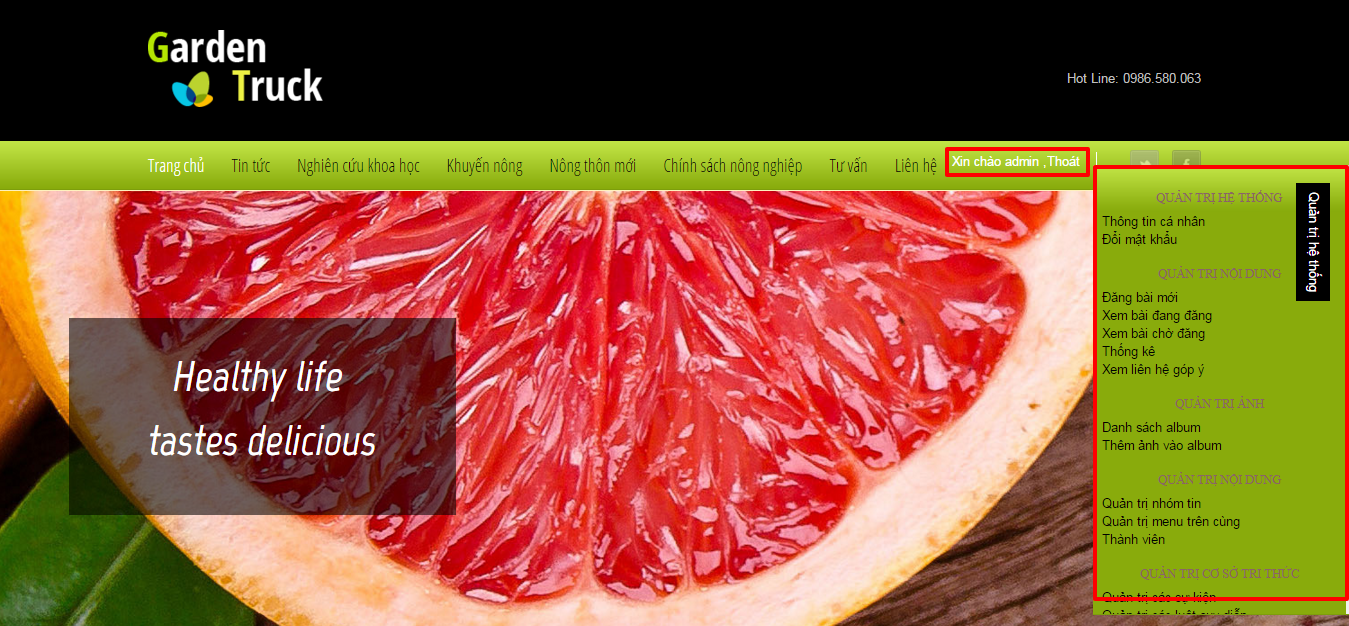
Chương trình Demo:

* Đường dẫn: <http://ducnhat.esy.es/index.php>
* Tài khoản Admin: admin
  1. **Đăng nhập vào quản trị hệ thống**



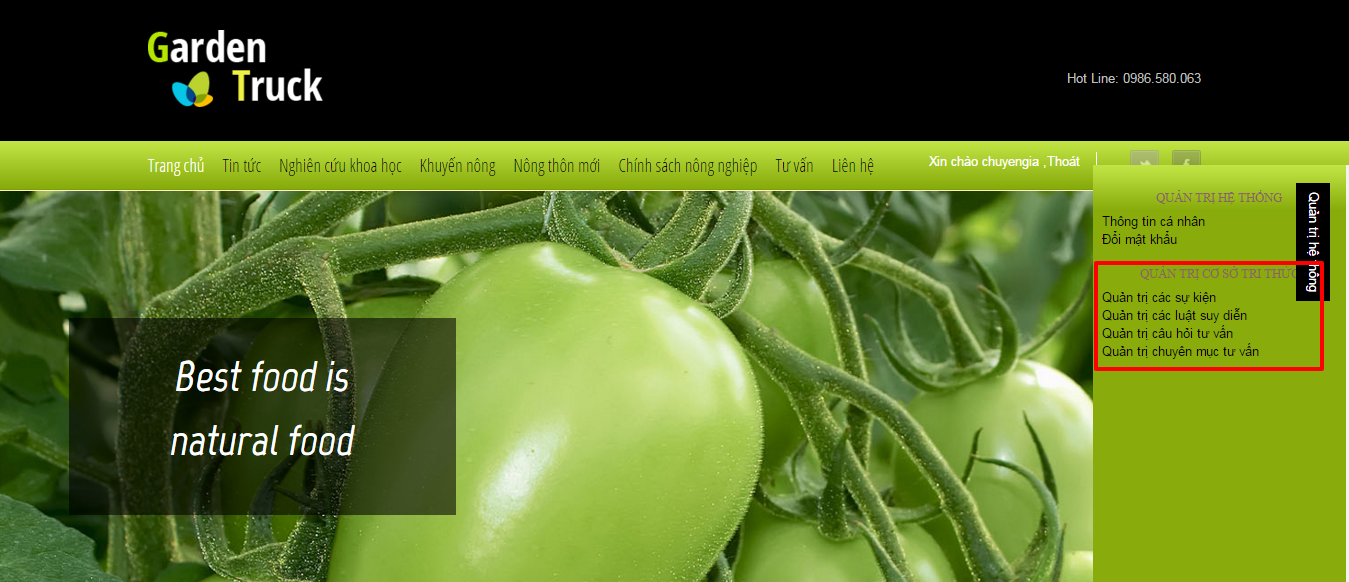
**Hình 10: Form đăng nhập**

* Truy cập menu quản trị



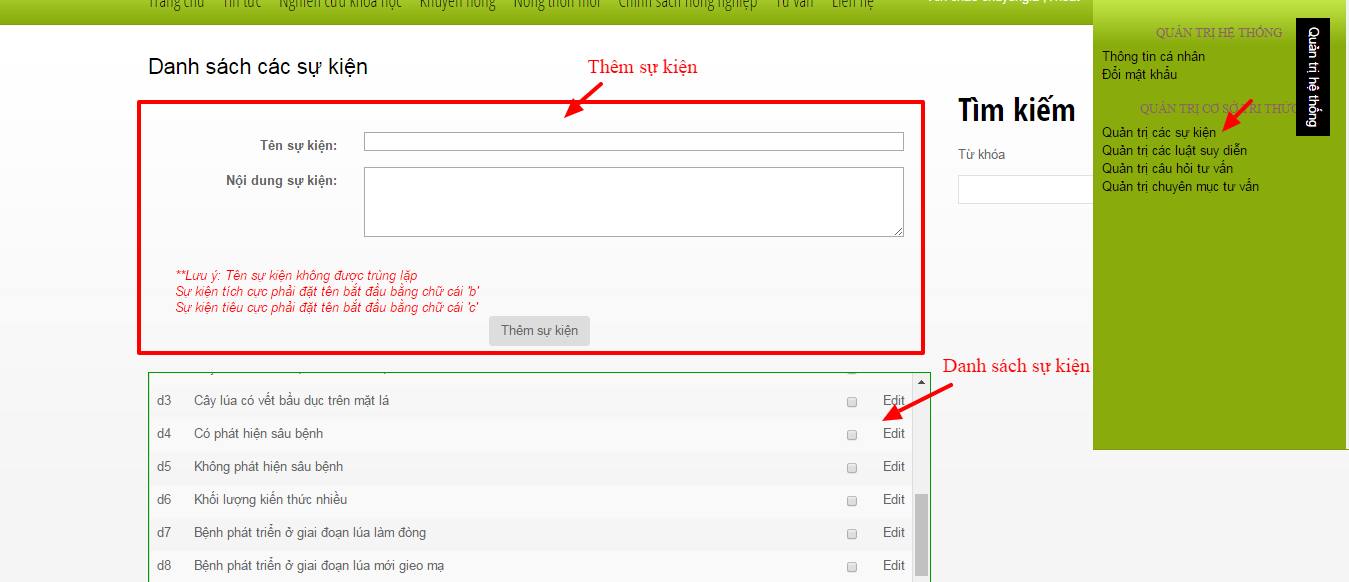
**Hình 11: Menu quản trị**

* 1. **Quản trị cơ sở tri thức**
* Menu quản trị cở sở tri thức:



**Hình 12: Menu quản trị cơ sở tri thức**

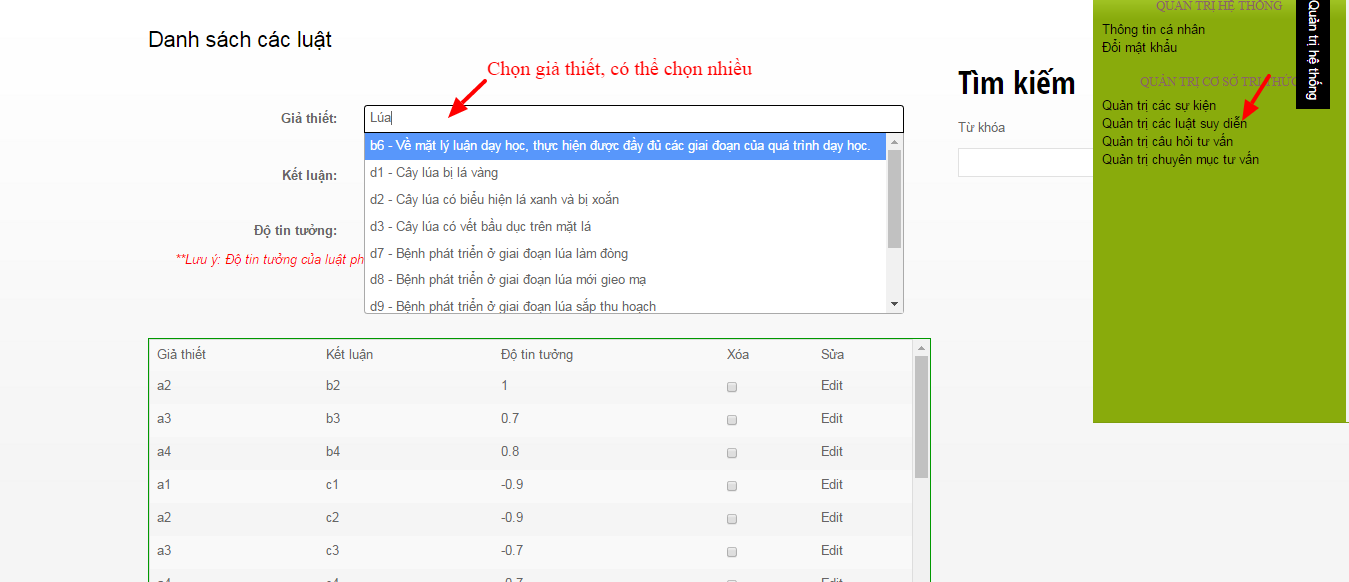
* Truy cập quản lý các sự kiện



**Hình 13: Quản trị sự kiện**

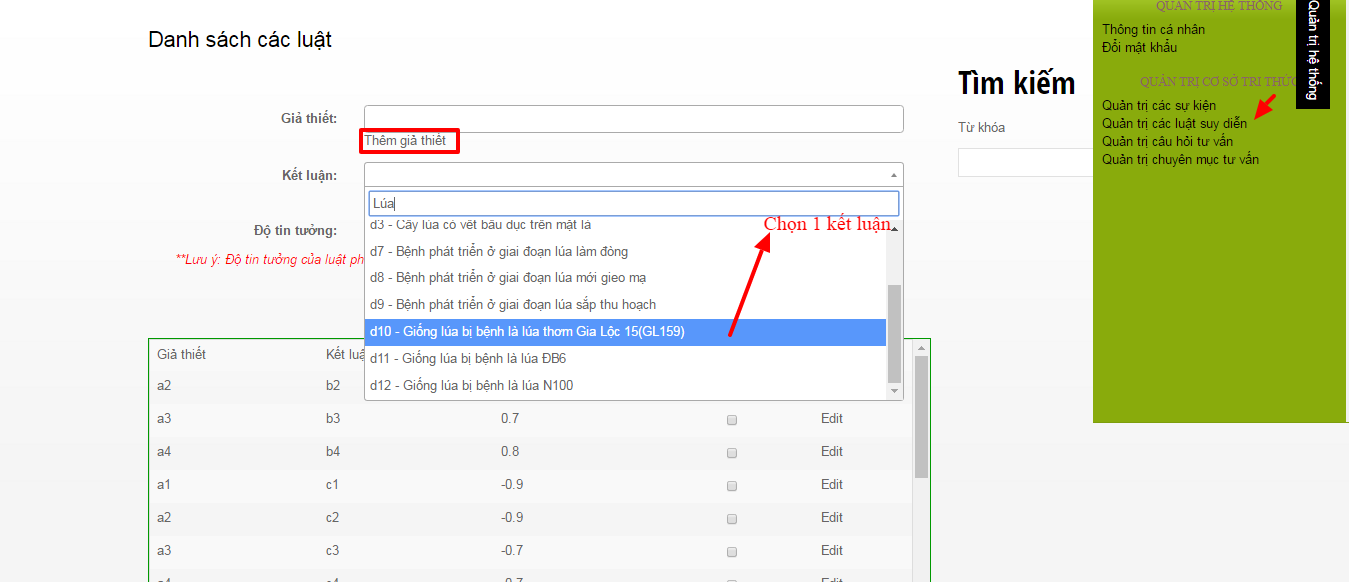
* Truy cập quản lý các luật suy diễn

Chọn giả thiết, có thể chọn nhiều giả thiết



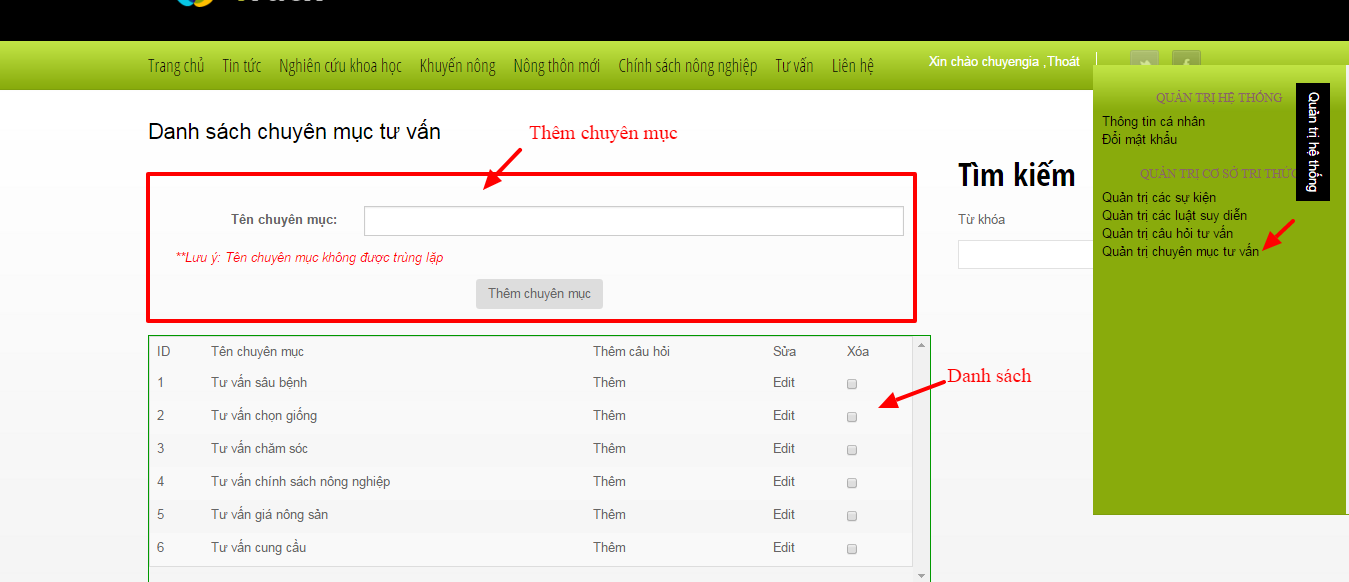
**Hình 14: Quản lý luật suy diễn – chọn giả thiết**

Thêm giả thiết mới (chính là các sự kiện) hoặc chọn kết luận cho giả thiết, chỉ được chọn 1 kết luận, nếu không thấy có thể thêm kết luận mới.



**Hình 15: Quản lý luật suy diễn – chọn kêt luận**

Truy cập quản lý chuyên mục cần tư vấn:



**Hình 16: Quản lý chuyên mục cần tư vấn**

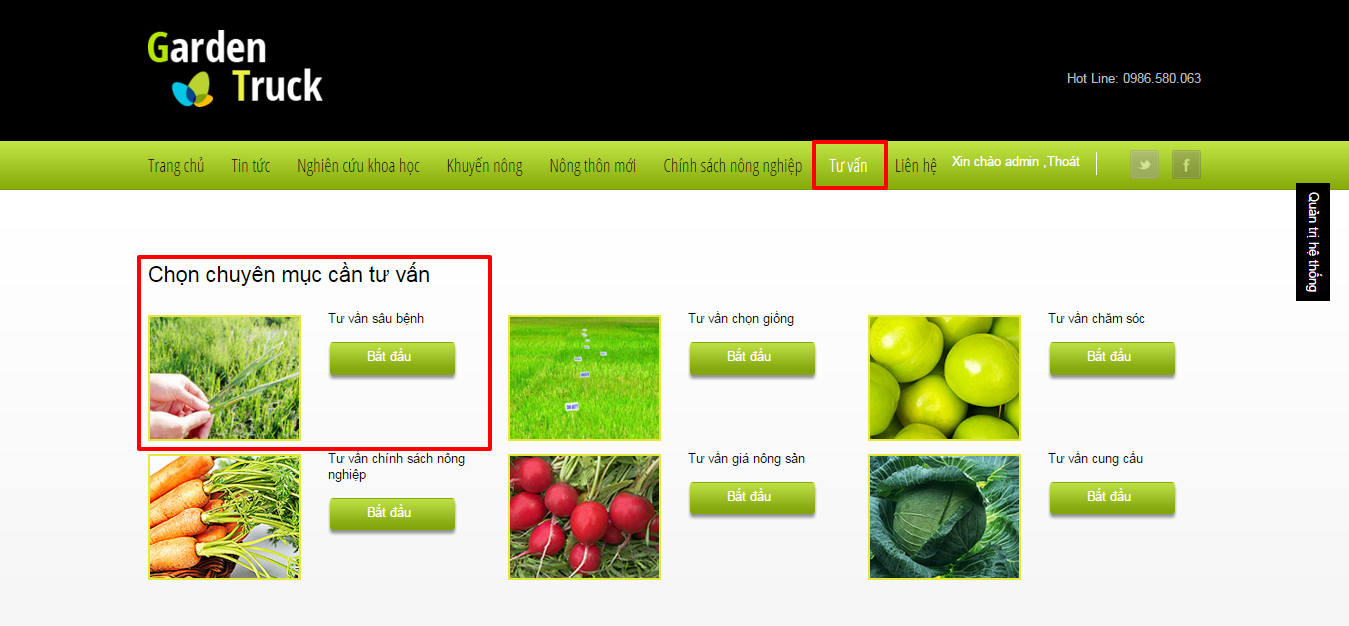
Truy cập quản lý các câu hỏi hỗ trợ tư vấn



**Hình 17: Quản lý câu hỏi hỗ trợ tư vấn**

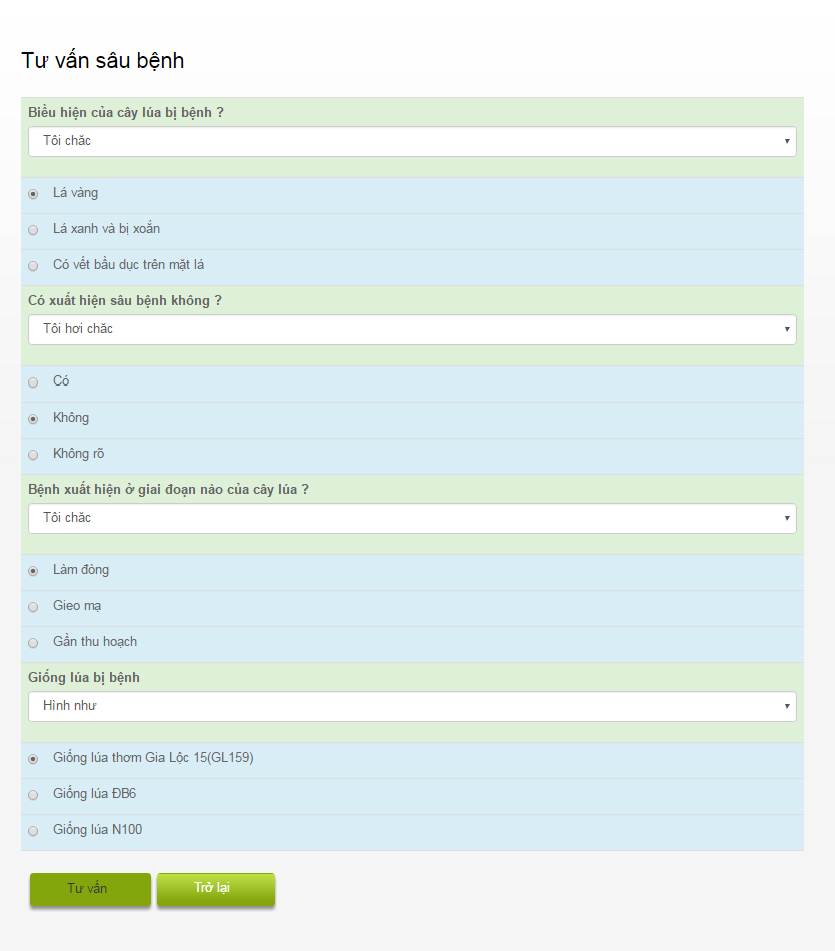
1. **Kết quả chương trình**

Trên thanh menu, truy cập module tư vấn



**Hình 18: Menu tư vấn**

Trả lời các câu hỏi để nhận được sự tư vấn



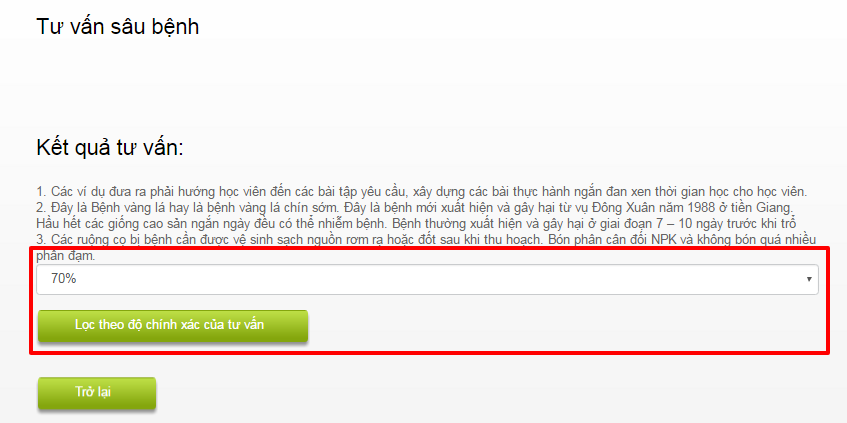
**Hình 19: Form đặt câu hỏi tư vấn**

Kết quả tư vấn (kết quả này được suy diễn từ Module Q&A ở chương 2)



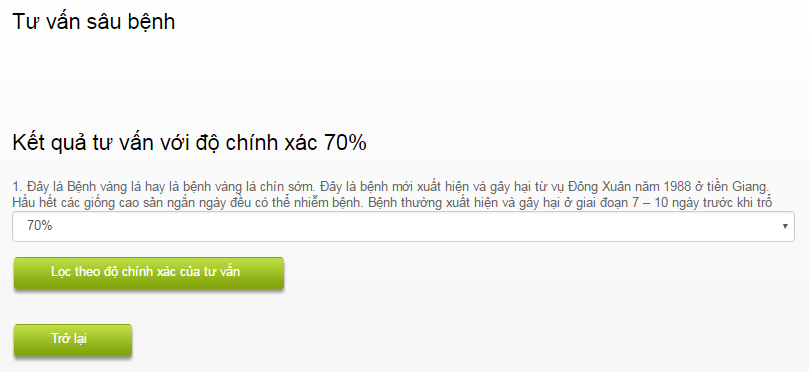
**Hình 20: Kết quả tư vấn của Module Q&A**

Nếu người dùng muốn lọc theo độ chính xác của các tư vấn, thì cần chọn độ chính xác và bấm nút “Lọc theo độ chính xác của tư vấn”



**Hình 21: Kết quả tư vấn của Module Reliability (1)**

Kết quả tư vấn (kết quả này được suy diễn từ Module Reliability ở chương 2)

****

**Hình 22: Kết quả tư vấn của Module Reliability (2)**

1. **Tổng kết chương**

Với những kết quả được thực nghiệm từ chương trình Demo, tuy vẫn còn nhiều hạn chế do bộ cơ sở tri thức chưa đầy đủ, nhưng hệ thống đã ứng dụng được thành công tri thức thông minh lên một hệ thống web portal và đáp ứng được những yêu cầu chức năng đã nêu ở chương 3.

**CHƯƠNG 5: KẾT LUẬN**

1. **Kết luận**

Luận văn đã xây dựng cổng thông tin tri thức tri thức điện tử cho nông sản nông nghiệp với tư vấn thông minh và làm tăng sự tương tác giữa người dùng và cổng thông tin. Kết quả thử nghiệm tương tác giữa các tác nhân trong hệ thông đã tạo ra cộng đồng liên kết người nông dân – doanh nghiệp – các chính sách nông nghiệp với nhau.

Nghiên cứu đưa ra các kỹ thuật suy diễn tri thức để suy hỗ trợ dịch vụ tư vấn thông minh trong cổng thông tin điện tử. Cổng thông tin cài đặt trên Internet thử nghiệm suy diễn với tư vấn dựa trên cơ sở tri thức, đánh giá thử nghiệm bước đầu kết hợp suy diễn và cổng thông tin.

Kết quả sản phẩm phần mềm trên Web cài đặt trực tuyến trên Internet và thử nghiệm cơ chế suy diễn từ cổng thông tin điện tử nông sản, nông nghiệp.

1. **Định hướng phát triển**

Trong thời gian tới, nghiên cứu tiếp tục phát triển các mô hình tri thức suy diễn cộng tác nhóm chuyên gia. Để thực hiện việc này, cần xây dựng mô hình trợ giúp suy diễn nhóm có cơ chế ưu tiên các luật, giải quyết sự xung đột giữa các luật. Đồng thời tiếp cận các cơ chế suy diễn lùi để giải quyết các bài toán tư vấn có kết quả đầu ra của tư vấn nông nghiệp, cần các hiện tượng, thông tin đầu vào.

Cổng thông tin tri thức sẽ được phát triển cài đặt trên nền dịch vụ điện toán đám mây để đáp ứng số lượng người truy cập lớn cho cổng thông tin nông nghiệp.

**TÀI LIỆU THAM KHẢO**

[1] Jongsoo Yoon, Tom E. Yoon, Joey F. George, Anticipating information needs for senior portal contents, Computers in Human Behavior, Vol. 27, No. 2, 2011, pp. 1012-1020

[2] Ahmad, F. ; Zakaria, Nur Haryani ; Osman, W.R.S., Transforming Information-Based Agricultural Portal to Knowledge-Based Agricultural Hub, 3rd International Conference on Information and Communication Technologies: From Theory to Applications, 2008 , pp. 1-

[3] Zakaria, S. ; Aziz, A.A. , Online Information Satisfaction: A Rasch Model Measurement of Information Management Experience within the Malaysian Agricultural Extension Services, Fifth International Joint Conference on INC, IMS and IDC, 2009 , pp. 842- 848

[4] Hilda M. Munyua, Christine Stilwell, Three ways of knowing: Agricultural knowledge systems of small-scale farmers in Africa with reference to Kenya, Library & Information Science Research, Vol.35, No. 4, 2013, pp. 326-337

[5] Athanasios Babalis, Ioannis Ntintakis, Dimitrios Chaidas, Athanasios Makris, Design and Development of Innovative Packaging for Agricultural Products, Procedia Technology, Vol. 8, 2013, pp. 575-579

[6] M. Debolini, E. Marraccini, D. Rizzo, M. Galli, E. Bonari, Mapping local spatial knowledge in the assessment of agricultural systems: A case study on the provision of agricultural services, Applied Geography, Vol. 42, 2013, pp. 23-33

[7] E. Antonopoulou, S.T. Karetsos, M. Maliappis, A.B. Sideridis, Web and mobile technologies in a prototype DSS for major field crops, Computers and Electronics in Agriculture, Vol.70, No. 2, 2010, pp. 292-301

[8] Vittorio Rossi, Francesca Salinari, Stefano Poni, Tito Caffi, Tiziano Bettati, Addressing the implementation problem in agricultural decision support systems, Computers and Electronics in Agriculture, Vol. 100, 2014, pp. 88-99

[9] Ye-lu ZHENG, Qi-yun HE, Ping QIAN, Ze LI, Construction of the Ontology-Based Agricultural Knowledge Management System, Journal of Integrative Agriculture, Vol. 11, No. 5, 2012, pp. 700-709

[10] Ye-lu ZHENG, Qi-yun HE, Ping QIAN, Ze LI, Construction of the Ontology-Based Agricultural Knowledge Management System, Journal of Integrative Agriculture, Vol. 11, No. 5, 2012, pp. 700-709

[11] Asanee Kawtrakul, Ontology Engineering and Knowledge Services for Agriculture Domain, Journal of Integrative Agriculture, Vol.11, No. 5, 2012, pp. 741-751.

[12] Cổng thông tin Viettel http://www.viettelpost.com.vn/Default.aspx?tabid=188&catid=87&itemid=8997

[13] Cổng thông tin điện tử - Bộ Nông nghiệp <http://www.mard.gov.vn/>

[14] Forward Chaining, <https://en.wikipedia.org/wiki/Forward_chaining>.

[15] Agri One – Cổng thông tin nông nghiệp Viettel <http://agrione.org.vn>

[16] Philip T Moore, Hai V. Pham, Personalization and Rule Strategies in HumanCentric Data Intensive Intelligent Context-Aware Systems In The Knowledge Engineering Review, Vol. 00:0, pp.1-24, United Kingdom, 2011.

[17] Nguyễn Thanh Thủy, Hệ cơ sở tri thức, 2009 NXB Giáo dục.

[18] Modus Pone, <https://en.wikipedia.org/wiki/Modus_ponens>