

NGHIÊN CỨU XÂY DỰNG PHẦN MỀM QUẢN LÝ CƠ SỞ DỮ LIỆU NGUỒN GEN CÂY TRỒNG NÔNG NGHIỆP

Nguyễn Tiến Hưng¹, Đới Hồng Hạnh¹, Lê Đình Sơn²

TÓM TẮT

Đến năm 2020, hệ thống bảo tồn quỹ gen cây trồng nông nghiệp đang lưu giữ khoảng 34.500 nguồn gen được thu thập trên cả nước và nhập nội. Hoạt động bảo tồn quỹ gen cây trồng nông nghiệp sinh ra một lượng dữ liệu tương đối lớn và phân tán, đa thuộc tính. Vì vậy, việc quản lý, cập nhật, khai thác sử dụng dữ liệu này gặp nhiều khó khăn nếu không có phần mềm chuyên dụng. Triển khai nội dung nghiên cứu thuộc đề tài KHCN cấp Quốc gia, Trung tâm Tài nguyên thực vật đã xây dựng phần mềm “Phần mềm quản lý cơ sở dữ liệu (CSDL) nguồn gen cây trồng nông nghiệp” trên nền tảng web sử dụng công nghệ .NET và Angular framework đáp ứng nhu cầu quản lý, chia sẻ CSDL nguồn gen cây trồng nông nghiệp và có thể liên kết được với CSDL khác. Phần mềm có thể hoạt động tốt trên tất cả các thiết bị có cài đặt trình duyệt web nên rất thuận lợi cho việc tra cứu, khai thác sử dụng cũng như quản lý CSDL nguồn gen cây trồng nông nghiệp tại Việt Nam.

Từ khóa: Cơ sở dữ liệu, phần mềm, cây trồng nông nghiệp.

1. TỔNG QUAN VỀ NỘI DUNG NGHIÊN CỨU

Bảo tồn quỹ gen sinh vật cũng đồng nghĩa với việc bảo vệ sự sống trên trái đất trước áp lực gia tăng dân số và biến đổi khí hậu. Vì vậy, bảo tồn quỹ gen sinh vật là nhiệm vụ ưu tiên và đòi hỏi sự phối hợp của toàn cầu. Xây dựng CSDL về bảo tồn và khai thác sử dụng bền vững nguồn gen cây trồng nông nghiệp thuộc quỹ gen sinh vật để hoạch định, giám sát và phát triển nhiệm vụ này là công việc tất yếu và không thể thiếu trong thời điểm hiện nay. CSDL bảo tồn quỹ gen cây trồng nông nghiệp lớn và phân tán, đa thuộc tính có khả năng chia sẻ qua mạng internet ở cả cấp quốc gia và quốc tế. Tuy vậy, CSDL quỹ gen cây trồng nông nghiệp của Việt Nam vẫn chưa được xây dựng cấu trúc chung, phần mềm chuyên dụng để quản lý, khai thác cho tất cả các cơ quan tham gia bảo tồn, chưa có khả năng liên kết đến lớp dữ liệu chi tiết do các cơ quan tham gia bảo tồn quản lý.

Quỹ gen sinh vật là nguồn vật liệu nền tảng không thể thiếu đối với công tác chọn tạo giống, công nghệ sinh học, công nghiệp thực phẩm, y dược... đặc biệt là khoa học sự sống. Bảo tồn và khai thác sử dụng bền vững nguồn gen sinh vật (bảo tồn nguồn gen sinh vật) tạo vị thế cạnh tranh cho mỗi quốc gia trong việc phát triển kinh tế, xã hội, an ninh quốc phòng cùng các giá trị về sinh thái, môi trường.

Công tác này đòi hỏi tổ chức thực hiện ở cấp quốc gia và mang tính toàn cầu với sự tham gia của nhiều loại hình cơ quan, tổ chức, nhiều ngành nghề khoa học kỹ thuật. Vì vậy nó đòi hỏi và sinh ra một lượng dữ liệu thông tin khổng lồ. Bởi thế đối với công tác bảo tồn nguồn gen sinh vật nói chung, xây dựng Hệ thống CSDL nguồn gen cây trồng nông nghiệp nói riêng là nhiệm vụ được thực hiện trước và liên tục được nâng cấp theo thời gian và không gian.

Các quốc gia phát triển có hạ tầng công nghệ thông tin tiên tiến đã xây dựng được CSDL quỹ gen quốc gia có khả năng liên kết trực tuyến đến các CSDL của đơn vị nghiên cứu chuyên sâu cho từng đối tượng bảo tồn. Trong khi đó, ở Việt Nam, CSDL quỹ gen quốc gia chưa được hình thành và mới chỉ có các phần mềm quản lý dữ liệu riêng lẻ, rời rạc hỗ trợ nghiệp vụ các cơ sở bảo tồn và nghiên cứu triển khai.

Ở Việt Nam, hệ thống các tổ chức bảo tồn tài nguyên di truyền thực vật gồm 23 cơ quan mạng lưới, đầu mối là Trung tâm Tài nguyên thực vật đang bảo tồn khoảng 34.500 nguồn gen của trên 300 loại cây trồng. Ở mỗi đơn vị bảo tồn đều thiết lập CSDL ở các cấp độ khác nhau. Đáp ứng những yêu cầu cơ bản cho các đơn vị bảo tồn và khai thác sử dụng.

2. VẬT LIỆU VÀ PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Vật liệu

Cơ sở dữ liệu được sinh ra từ hoạt động bảo tồn nguồn gen cây trồng nông nghiệp tại hệ thống bảo tồn tài nguyên di truyền thực vật Quốc gia, bao gồm:

¹ Trung tâm Tài nguyên thực vật, Viện Khoa học Nông nghiệp Việt Nam

² Học viện Kỹ thuật Quân sự

*Email: doihonghanh@gmail.com

Dữ liệu điều tra thu thập, dữ liệu mô tả đánh giá, dữ liệu lưu giữ và cấp phát nguồn gen.

2.2. Phương pháp nghiên cứu

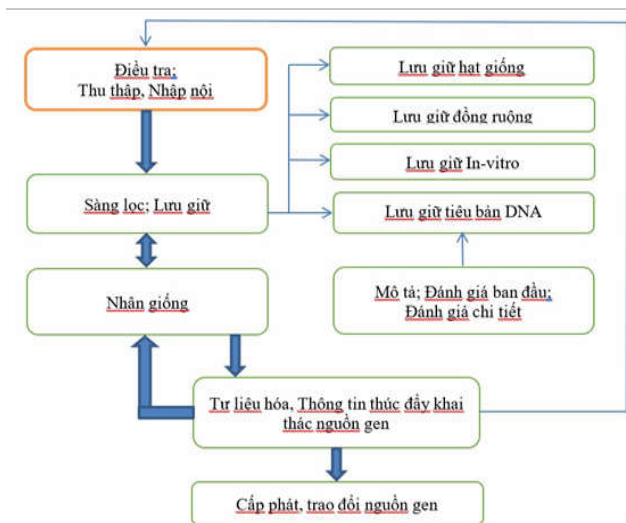
Về lý thuyết	+ Tìm hiểu về mô hình, CSDL lớn trong việc lưu trữ nguồn gen, + Tìm hiểu về mô hình, thiết kế CSDL phân tán, + Tìm hiểu về mô hình, bài toán tìm kiếm, đối sánh nguồn gen trong kho dữ liệu lớn, + Phát triển hệ thống thông tin tích hợp hỗ trợ nghiệp vụ quản lý.
Về thực nghiệm	+ Lấy mẫu dữ liệu về các nguồn gen, số hóa và cập nhật vào hệ thống, + Cài đặt, thử nghiệm hệ thống phần mềm, hỗ trợ công tác quản lý, tìm kiếm, đối sánh nguồn gen, + Thử nghiệm, đánh giá hiệu quả hệ thống phần mềm trên bộ CSDL.

3. KẾT QUẢ NGHIÊN CỨU VÀ THẢO LUẬN

3.1. Quy trình nghiệp vụ bảo tồn lưu giữ cây trồng nông nghiệp tại hệ thống bảo tồn

Quá trình khảo sát và phân tích đã xác định được quy trình nghiệp vụ của hoạt động bảo tồn nguồn gen cây trồng nông nghiệp.

Công tác bảo tồn nguồn gen cây trồng nông nghiệp được thực hiện gồm các bước: Điều tra thu thập (hoặc nhập nội) đưa vào lưu giữ tại các loại hình ngân hàng gen sau đó nhân giống, mô tả đánh giá sau đó tiến hành tư liệu hóa thông tin thúc đẩy khai thác nguồn gen và cuối cùng cấp phát, trao đổi nguồn gen. Căn cứ vào sơ đồ nghiệp vụ để xây dựng mô hình hóa chi tiết quy trình nghiệp vụ, xây dựng biểu đồ các trường hợp sử dụng, xây dựng kiến trúc phần mềm quản lý và thiết kế cơ sở dữ liệu cho phần mềm.



Hình 1. Sơ đồ nghiệp vụ bảo tồn và lưu giữ cây trồng nông nghiệp

3.2. Thiết kế cấu trúc CSDL nguồn gen cây trồng nông nghiệp

CSDL nguồn gen cây trồng nông nghiệp là một trong 6 nhóm CSDL về bảo tồn quỹ gen sinh vật thường được tổ chức thành hai lớp:

- Lớp cơ bản: Bao gồm những thông tin chính về loại đối tượng bảo tồn được cụ thể hóa đến đơn vị bảo tồn là:

- + Mẫu giống (accession): CÁI GÌ.
- + Số lượng mẫu giống bảo tồn: BAO NHIÊU.
- + Nơi thu mẫu và cơ quan lưu giữ: Ở ĐÂU.
- + Mức độ tồn tại trong sản xuất và tự nhiên, mức độ nghiên cứu mô tả đánh giá: NHƯ THẾ NÀO.



Hình 2. Mô hình kiến trúc phần mềm

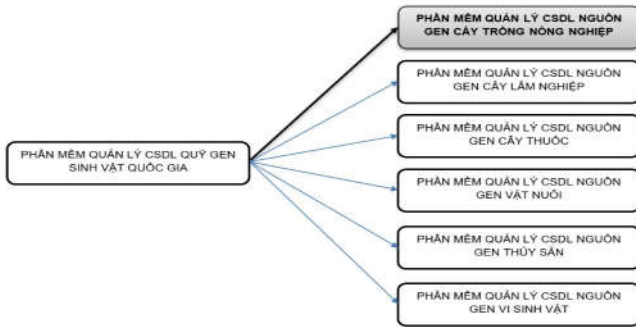
- Lớp chi tiết: Lớp dữ liệu này chứa dữ liệu thông tin lai lịch đầy đủ (passport data) bao hàm những thông tin gắn liền với mẫu thu kể cả dữ liệu không gian, tri thức truyền thống về bảo tồn, canh tác và sử dụng nguồn gen. Dữ liệu mô tả đánh giá đặc điểm hình thái, đánh giá sâu về chất lượng, chống chịu điều kiện bất thuận sinh học, phi sinh học, đánh giá bản chất di truyền bằng chỉ thị phân tử, hình ảnh đặc tả nguồn gen... Các báo cáo sử dụng, khai thác nguồn gen (cấp phát vật liệu chọn tạo giống, phục tráng, chọn tạo...). Lớp CSDL này thường do các tổ chức bảo tồn chuyên sâu theo nhóm cây trồng thiết lập.

3.3. Thiết kế hệ thống phần mềm

Hệ thống phần mềm được phát triển trên nền tảng website sử dụng công nghệ .NET và Angular

framework, là một trong khối các hệ thống phần mềm quản lý CSDL nguồn gen quốc gia (khoanh đậm ở hình 3).

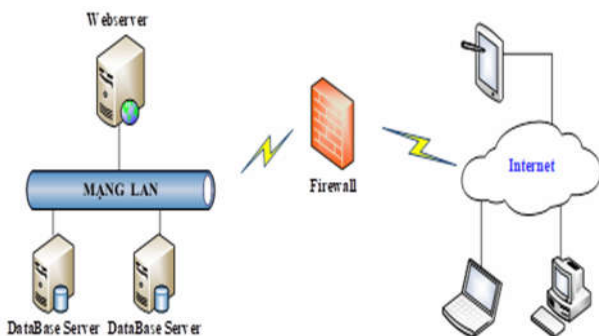
Mỗi phần mềm con quản lý CSDL của từng đối tượng và một phần dữ liệu được đồng bộ sang phần mềm quản lý CSDL quỹ gen sinh vật Quốc gia phục vụ công tác thống kê, khai thác cơ bản để người sử dụng có cái nhìn tổng thể về công tác bảo tồn tài nguyên sinh vật của Việt Nam.



Hình 3. Sơ đồ tổng quan các hệ thống trong phần mềm quản lý CSDL quỹ gen sinh vật Quốc gia

Phần mềm quản lý CSDL nguồn gen cây trồng nông nghiệp được xây dựng trên nền tảng website sử dụng công nghệ .NET và Angular framework đáp ứng nhu cầu quản lý, chia sẻ CSDL nguồn gen cây trồng nông nghiệp và liên kết được với CSDL quỹ gen sinh vật. Phần mềm này có thể hoạt động tốt trên tất cả các thiết bị cài đặt trình duyệt web nên rất thuận lợi cho việc tra cứu, khai thác sử dụng cũng như quản lý.

Hệ thống sau khi phát triển đã được triển khai trên hệ thống máy chủ của Trung tâm Tài nguyên thực vật. Cụm máy chủ được thiết kế theo kiến trúc Cluster bao gồm 2 server website và 1 server cài đặt hệ quản trị cơ sở dữ liệu Microsoft SQL Server nhằm đảm bảo nâng cao khả năng sẵn sàng cho các hệ thống khi có lượng truy cập lớn.



Hình 4. Kiến trúc triển khai phần mềm quản lý CSDL nguồn gen cây trồng nông nghiệp

3.4. Kết quả

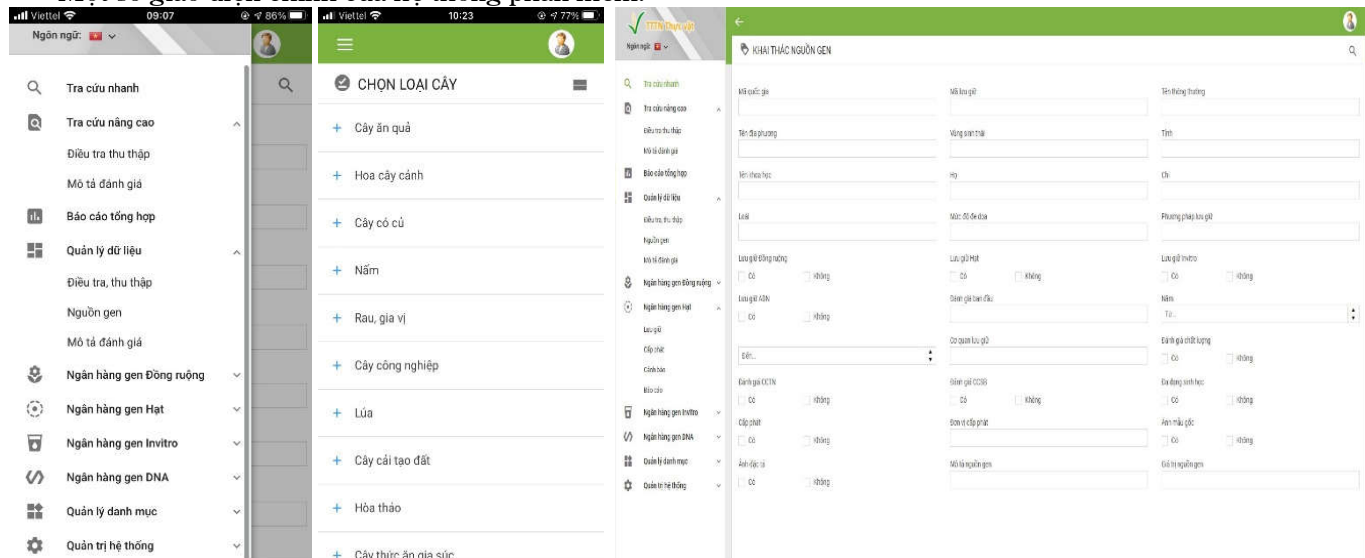
Thực hiện Chương trình bảo tồn quỹ gen cây trồng nông nghiệp, Việt Nam đã thành lập hệ thống bảo tồn tài nguyên di truyền thực vật Quốc gia hoạt động theo nguyên tắc mạng lưới bảo tồn với Trung tâm Tài nguyên thực vật là cơ quan điều phối mạng lưới. Trung tâm đảm nhận nhiệm vụ nghiên cứu, điều tra, thu thập, bảo tồn, đánh giá, tư liệu hóa và khai thác, sử dụng tài nguyên di truyền thực vật thông qua xây dựng, quản lý Ngân hàng gen hạt, Ngân hàng gen đồng ruộng cây thường niên, Ngân hàng gen In-vitro và xây dựng, quản lý CSDL quỹ gen cây trồng của toàn hệ thống. Đến năm 2020 đã xây dựng được phần mềm quản trị CSDL nguồn gen cây trồng nông nghiệp với 9 mô-đun chức năng chính: Tra cứu, Báo cáo tổng hợp, Quản lý thông tin nguồn gen, Ngân hàng gen hạt, Ngân hàng gen đồng ruộng, Ngân hàng gen Invitro, Ngân hàng gen DNA, Quản trị hệ thống, Quản lý danh mục.



Hình 5. Giao diện đăng nhập

Với phần mềm này, Trung tâm sẽ lưu trữ được toàn bộ dữ liệu sinh ra từ hoạt động bảo tồn của Hệ thống bảo tồn tài nguyên di truyền thực vật quốc gia với khoảng 34.500 mẫu nguồn gen lưu giữ bao gồm: Dữ liệu điều tra thu thập, dữ liệu mô tả đánh giá ban đầu, đánh giá chi tiết, dữ liệu lưu giữ và cấp phát.

Một số giao diện chính của hệ thống phần mềm:



Giao diện chức năng chính trên mobile

Giao diện Quản lý nhóm cây trên mobile

Giao diện chính của phần mềm trên máy tính

4. KẾT LUẬN

Đã đánh giá được hiện trạng CSDL nguồn gen cây trồng nông nghiệp, xác định nhu cầu quản lý và sử dụng dữ liệu này của 3 nhóm cơ quan tổ chức: quản lý nhà nước; trực tiếp làm nhiệm vụ bảo tồn; khai thác sử dụng nguồn gen.

Đã thiết lập được cấu trúc CSDL nguồn gen cây trồng nông nghiệp đảm bảo nghiệp vụ và phù hợp với quốc tế.

Đã thiết lập được phần mềm để lưu giữ, quản lý, sử dụng và chia sẻ dữ liệu về bảo tồn, sử dụng hiệu quả và bền vững nguồn gen cây trồng nông nghiệp.

TÀI LIỆU THAM KHẢO

- Đặng Thị Việt Đức, Nguyễn Thanh Tuyên (2011). *Vai trò của công nghệ thông tin và truyền thông trong nền kinh tế tri thức và trường học của Việt Nam*. Tạp chí Công nghệ và Thông tin truyền thông (kỳ 2, tháng 2/2011), tr. 19-26.
- Hồng Minh (2011). *Giải pháp chiến lược công nghệ thông tin hiện đại*. Tạp chí Công nghệ và Thông tin truyền thông (kỳ 2, tháng 1/2011), tr. 48-50.

- Nguyễn Thanh Minh, Nguyễn Bội Ngọc (2011). *Xây dựng Chính phủ điện tử: Những thách thức căn bản*. Tạp chí Công nghệ và Thông tin truyền thông (kỳ 2, tháng 1/2011), tr. 14-16.

- Lê Văn Hưng, Nguyễn Bá Tú (2014). *Thực trạng công tác xây dựng cơ sở dữ liệu và trao đổi thông tin về nguồn gen và TTTT về nguồn gen ở Việt Nam*. Tạp chí Nông nghiệp và PTNT, tháng 10, tr. 27-34.

- Nguyễn Tiến Hưng, Lã Tuấn Nghĩa, Lưu Ngọc Trinh, Nguyễn Thị Sen, Hoàng Thị Hải, Nguyễn Thị Quyên, Nguyễn Thị Huyền (2012). *Kết quả xây dựng hệ thống thông tin tài nguyên thực vật giai đoạn 2006-2011 và những định hướng tiếp theo*. Tạp chí Nông nghiệp và PTNT, chuyên đề Tài nguyên thực vật, tháng 12/2012, tr.172-180.

- Nguyễn Đăng Khôi (1995). *Báo cáo của Ban chủ nhiệm Dự án, trang 10-26 trong Tài nguyên di truyền thực vật ở Việt Nam*. Các báo cáo tại Hội thảo Quốc gia về tăng cường chương trình tài nguyên thực vật ở Việt Nam, Hà Nội, 28-30/3/1995, NXB Nông nghiệp.

- CBD, Business, 2020. *A Magazine on Business and Biodiversity*, Nov. 2015. Vol. 10 Issue 1.

**RESEARCH AND DEVELOPMENT OF SOFTWARE FOR AGRICULTURAL PLANT
DATABASE MANAGEMENT**

Nguyen Tien Hung, Doi Hong Hanh, Le Dinh Son

Summary

By 2020, the National Plant Genetic Resources Conservation Network is storing about 34,500 accessions collected across the country and imported. Conservation of agricultural plant genetic resources a relatively large amount of data, which is scattered and multi-attributed. So, it is very difficult to management, update and exploit this data if have not specialized software. Deploying the research content of the national science and technology topic, the Plant Resources Center had built the software "Software to manage the database of agricultural plant genetic resources" on the web platform using .NET technology and Angular framework meet the needs of managing and sharing agricultural crop genetic resources database and can be linked with other databases. The software can work well on all devices with a web browser installed, so it is very convenient for searching, exploiting and using as well as managing the database of genetic resources of agricultural crops in Vietnam.

Keywords: *Data, software, agricultural plant.*

Người phản biện: PGS.TS. Khuất Hữu Trung

Ngày nhận bài: 10/5/2021

Ngày thông qua phản biện: 11/6/2021

Ngày duyệt đăng: 18/6/2021