Journal of Innovative Technology Convergence
Vol. 2, No. 2 December 2020, pp. 63-74

Design and Development of a Comprehensive, Robust and KM-Scalable Database (Backend) for a PILI (Canarium ovatum Engl.) Information System

John Rey S. Lirag[1], Ronnel R. Atole[2], Leo Constantine S. Bello[3], Rowell John B. Artiaga[4], Salvador V. Briones II[5]

Abstract: This paper presents the results of efforts to design and develop a comprehensive repository of knowledge for Pili (Canarium ovatum Engl.), an endemic crop in the Bicol Region known for its various uses. This study, in general, responds to the challenges entailed in the national government's having identified "PILI" as the flagship commodity of the Bicol region. In particular, it seeks to address the dearth of accessible knowledge on the Pili industry. Enabling users around the world to become acquainted with this endemic crop. The database's initial goal is to provide key facts on population dynamics of Pili crop (Canarium ovatum Engl.) varieties and is designed and structured towards knowledge management (KM) buildup.

Many similar information systems applied in agriculture were consulted and visited by the researchers in the literature such as the International Rice Research Institute (IRRI's) comprehensive crop information system [1], and the more recent Mango Resource Information System [2] which was developed for the management of the phenotypic, genetic, molecular to chemical and other information on mango cultivars.

[1] Information and Communication Technology Management Office, Partido State University, Goa, Camarines Sur, Philippines
Email: johnreylirag@parsu.edu.ph
[2] Information and Communication Technology Management Office, Partido State University, Goa, Camarines Sur, Philippines
Email: ronnel.atole@parsu.edu.ph
[3] Information and Communication Technology Management Office, Partido State University, Goa, Camarines Sur, Philippines
Email: lbello@parsu.edu.ph
[4] College of Arts and Sciences, Institute of Information Technology, Partido State University, Goa, Camarines Sur, Philippines
Email: john.artiaga@parsu.edu.ph
[5] College of Arts and Sciences, Institute of Information Technology, Partido State University, Goa, Camarines Sur, Philippines
Email: salvador.briones@parsu.edu.ph

Acknowledgement: We would like to thank DOST-PCAARRD for the funding and giving us the opportunity to become a member of multi-project program known as the "Accelerated R&D Program for Capacity Building of Research and Development Institutions and Industrial Competitiveness: Niche Centers in the Regions for R&D (NICER) Program: Pili Research & Development Center".

Received [October 20, 2020]; Revised [December 4, 2020]; Accepted [December 8, 2020]

Published by InnoCon Publishing
ISSN 2704-4440

The breadth and depth of information in the database are designed to cater to the different needs of diverse groups of stakeholders (farmers, vendors, processors, scientists, researchers, policy-makers, teachers, and students). Its various applications range from basic Pili bio-ecological data management to e-commerce and analytics. This Pili database covers the more than 14 NSIC (National Seed Industry Council) accredited pili varieties, 200 accessions compiled from more than 150 references in partnership with more than 10 collaborators, and more than 500 pictures taken from different cultivars in the Bicol region.

At present, the database has 180 tables, enriched by the collection of various Pili related information ranging from its nomenclature, accession, NSIC registration, distribution, morphological characteristics, genetic resources, geo-mapping and characterization, propagation techniques, emerging diseases, growth and yield response, postharvest processing, to products, e-commerce, data analytics and other related data present in different cultivars. The database is normalized to handle complex queries needed to generate knowledge not only for Philippine Council for Agriculture, Aquatic and Natural Resources Research and Development (PCAARRD) and its member-agencies but for Pili farmers/producers, vendors, enthusiasts, retail outlets, and the market as a whole. It is designed with flexibility in mind to interface and interconnect/link with other PCAARRD systems for a dynamic web of knowledge. It is believed that making information and data on Pili publicly available will enliven and invigorate the stakeholders in this industry and help boost yield/production, encourage more research and development efforts, promote innovation, step up marketability of various pili-related products, and consequently improve the living conditions of marginalized farmers and vendors.

Keywords: Structured Relational Database Model, MySQL, Canarium ovatum Engl., Knowledge Management

1. Introduction

1.1 Pili Industry

The agriculture sector is a unique sector due to its strategic importance around the world. It is crucial for both citizens (consumers) and the economy (regional and global) which, ideally, should make the whole sector a network of interacting organizations [3]. The Philippines, a tropical country, is very rich in edible fruit and nut-bearing trees, of which about 167 are indigenous [4]. Several fruit trees that bear edible nuts are claimed to have their center of diversity in the Philippines. The most important of these is pili, Canarium ovatum Engl. (Figure 1), of which geographic distribution in the country remains limited to areas located relatively closer to its center of origin [5]. Pili is considered to be the most important nut-producing species indigenous to the country and is identified as the “flagship commodity” of the Bicol region. It has a nationwide acceptance and has great potential to develop into a major industry [6] As a nut, the development of this crop is promising as demand for the processed kernel is rising. The pili nut kernel (Figure 2) is the most important part of the tree and has many uses [5]. Pili nuts are mainly used to manufacture candies and confectioneries, while pili nut oil is highly in demand locally and in foreign countries like Guam, Australia, Canada, and the United States [6].

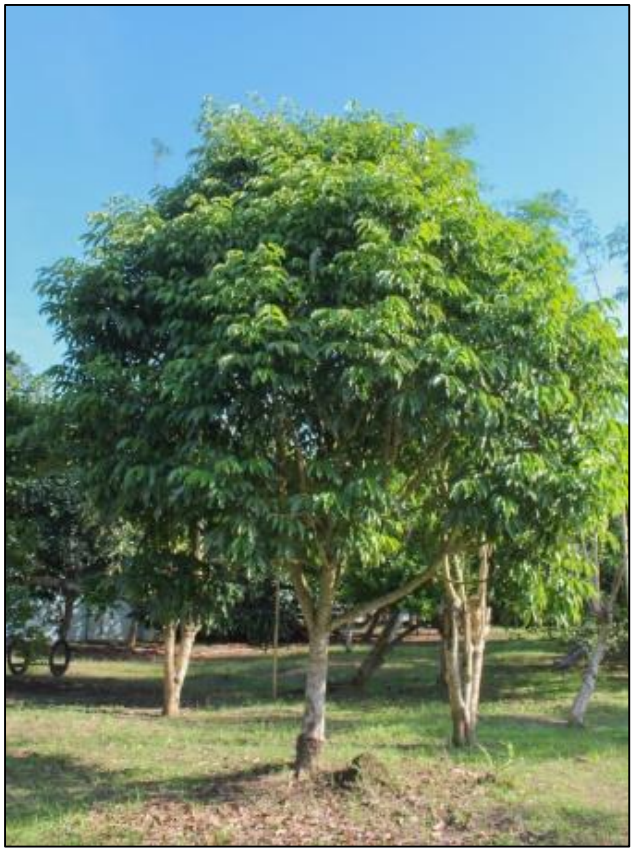

Figure 1. Pili (Canarium ovatum Engl.) Tree

Figure 2. Pili Nut Kernel

1.2 Pili Problem Domain

Despite numerous programs being orchestrated by the Philippine government and other participating agencies, this cannot hide the fact that information about pili potentials is still little known to others.

Based on the related studies consulted, researchers observed that most agricultural studies focused mainly on staples and stretch little attention to a minor or underutilized crop [5]. Similar scenario to

Food Supplying agencies or entities down to Municipal Agriculture Office they have been more attentive on a few staple crops being produced like rice, corn, coconut, and sugarcane industry. Underutilized crops though minors yet can be developed and turns into major industry when proper intervention and support is implemented. Nonetheless, crops considered to be neglected at the global level are staples at a regional level or even the national level.

Just like other agricultural crops, Pili (Canarium ovatum Engl.) crop is considered to be one of the minorities. Research studies about Pili becomes challenging due to inadequate resources. Another reason is that knowledge on this crop is scattered, information like germplasm is not readily accessible and is commonly found in grey literature or written in the little-known language. Knowledge of its genetic composition and potential is also limited. These results in an inefficient approach to its genetic conservation and utilization. With these limitations, there is also grey literature wished that in the near future, Pili will be graded equal with the cashew and macadamia in the world market.

Due to the fact that the availability of information on Pili are lacking, limited, and scattered, the progress and growth of the knowledge and its utilization have been greatly hindered. This reason calls the attention of the researchers to design and develop methods and ways of pooling information on Pili into one readily accessible source. This allows unequivocally linking of all information present in the Pili database. The Pili products, though to some extent has been a success, is yet to be backed by readily accessible data and knowledge. Though Pili products are everywhere these days, there is no comprehensive and easily retrievable repository of knowledge on this particular industry.

The researchers believed that the success of this database will promote awareness to all the stakeholders including farmers and vendors, on the marketability and possibility of investing in the Pili industry and rebound to a vibrant area of this agricultural sector. Moreover, scientists, researchers and students will be given access to a comprehensive source of information about Pili ranging from its classification, production sites, uses/applications, products, production data, plantations, technologies, the profile of the farmer, profile of the vendors and growers, prevailing market condition and any other information.

The Philippine government has also established initiatives in promoting and cultivating the survival of the Pili gene pool. Regular seminars are held on Pili production to interest farmers; provide seedlings to farmers in a reasonable price and high-quality propagation materials; agricultural agencies also conducted regular visits on farmer’s orchard by expert personnel and provide technical assistance to them; and, a promise of Government assistance to Pili farmers in marketing their produce. The success of both will surely be of great help to promote this pili industry, prospects, and potentials.

2. Methodology

2.1 Database Requirement Analysis and Scoping

To support the aim to provide a comprehensive collection and global information with all you ever wanted to know about pili. This stage of the database lifecycle (DBLC) as shown in Figure 3 is considered by the researchers as the most important stage in the creation of the Pili database. This includes characterizing what information is needed by the target clientele, how is the information be gathered and stored, where to pull the information, how to validate the integrity of the information or is the information scientific, what security mechanism to be imposed, and including all the conditions under which information regarding pili needs to be accessed by different types of users.

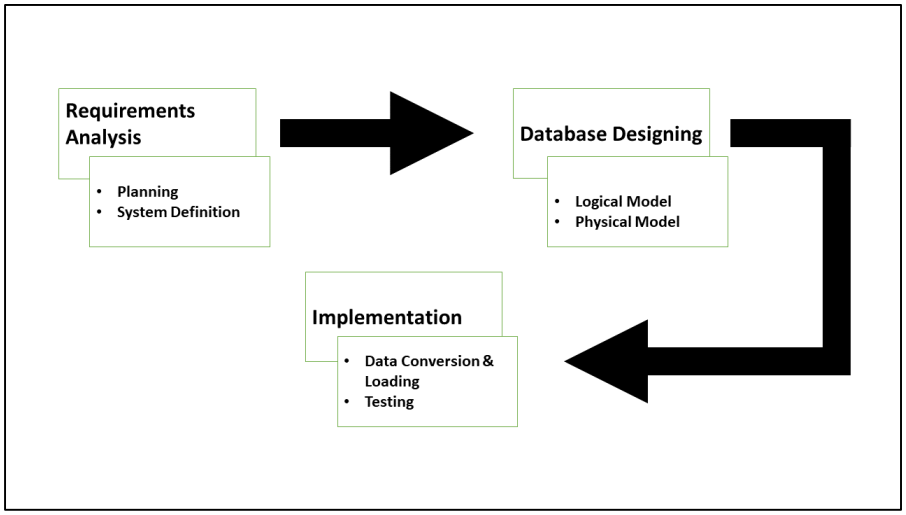

Figure 3. Database Development Lifecycle

2.2 Data Gathering

Secondary data gathering was conducted to compile knowledge on the Pili database ranging from production data, the volume of sales, profile of farmers and vendors, biological characteristics of Pili and diversities, *etc.* Research data and outputs on Pili, as well as records, were also considered.

In the evaluation phase, a survey instrument was developed, respondents of which represent all stakeholders in the Pili industry such as vendors, farmers, manufacturer, researchers, consumers, *etc*. with this scenario, a simple descriptive treatment of data and qualitative approach will suffice.

In other situations, key-informant interviews and focused group discussions were conducted such as in formulating guidelines, policies, and procedures for the Information System operation, or in validating secondary data.

2.3 Database Structure, Design and Implementation

Databases are considered as a knowledge medium, a representation of a combination of knowledge types, formats, and purpose of represented knowledge [7][8]. Rosen and Rimor [9] states that "building a database involves procedural knowledge including strategies for building the database and meta-cognitive knowledge including insights on the knowledge base of the data, and on the structure of the database itself". This section presents and discusses the database structure, design, and implementation.

Pili database was implemented using MySQL, a relational database management system, fast, multi-threaded, multi-user, and robust. This database stores comprehensive collection and global information with all you ever wanted to know about Canarium ovatum Engl. The database schema has been normalized into third normal form to avoid storage of data with logical inconsistencies. The database contains different things for different people. For example, pili processors will dive into the largest existing compilation of nearby pili farms to become the supplier of raw products; marginalized farmers will find numerous information illustrating the facts, prospects, and potentials of the Pili industry; researchers and scientists will make use of journals, articles, utility models, inventions, research papers for their references. This database will be a repository of Pili information ranging from its nomenclature, accession, NSIC registration, distribution, morphological characteristics, genetic resources, geo-mapping and characterization, propagation techniques, emerging diseases, growth and yield response, postharvest processing, to products and other related data present in different cultivars. This information can lead to the generation of new knowledge and can be extracted into meaningful information, figures needed in the visualization of business processes from macro to the micro-level of view. The database

model is built around a variety of data types, including ids, texts, time and date, multimedia as well as links that inherent the attributes of classes of the database.

The interface is programmed in PHP using the Laravel framework. It meets the W3C XHTML 1.0 strict specification drafted by the World Wide Web Consortium. As a result, the web page is accessible and will display correctly in any browser that conforms to the standard. The web interface also makes use of Bootstrap, a free and open-source CSS framework directed at responsive, mobile-first front-end web development. It contains CSS- and JavaScript-based design templates for typography, forms, buttons, navigation, and other interface components. This allows easy knowledge entry and organized knowledge display. The interface includes a normal user interface for viewing only and an administrator functions overlay for logged-in users with adequate privileges.

2.4 Data Model & Vocabularies

A conceptual diagram of the many types of data that can be generated from one Pili is shown in Figure 4.

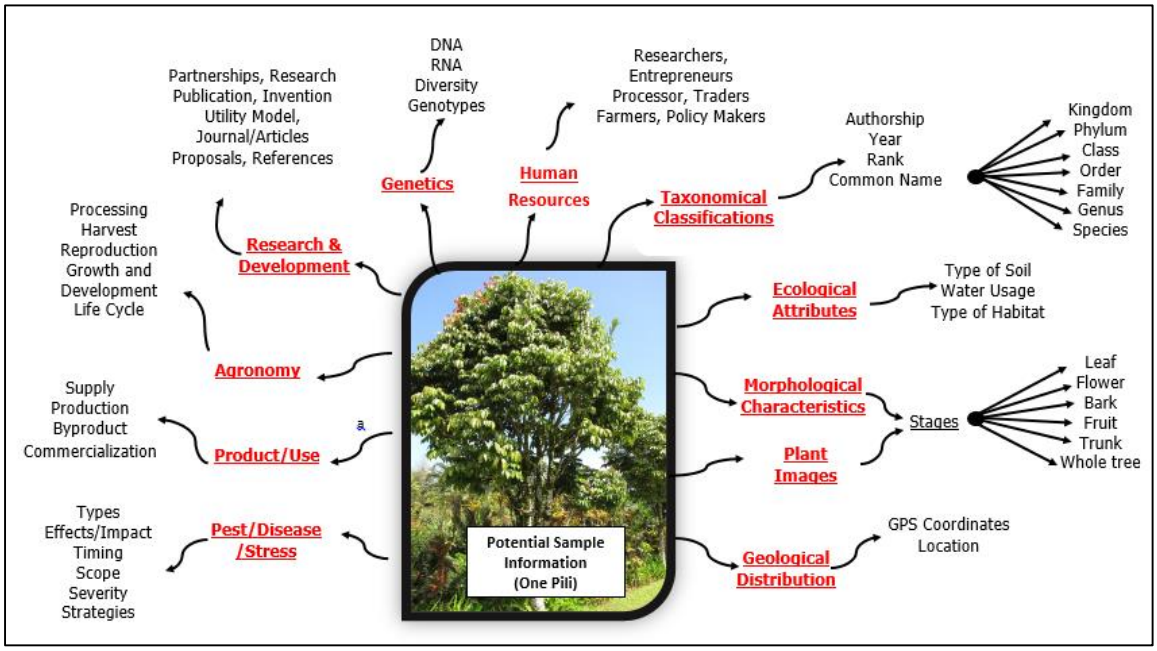

Figure 4. A Mind Map of the Potential Information Associated with a Pili Sample

A sample is described with the taxonomical classification, morphological characteristics, ecological attributes, geological distribution, product/uses, pest/disease/stress, agronomy, Research and Development, genetics, human resource, and plant images.

1. Pest, Disease, Stress. Entities under this section stored information about what particular type of threat/stress is present in a crop(s), “abiotic/biotic”. Information about effects, timing, severity to strategies/treatment are also stored here.
2. Product/Use. This section stored the different types of Pili products, edible or non-edible.
3. Agronomy. This section stored information about growth and development, lifecycle, reproduction, and other details linked to different Pili varieties.
4. Research and Development. This portion of the database stored information about publications, utility models, inventions, articles, and other details that deal with Pili.

5. Genetics. Pili genetic information.
6. Human resources. This stores the different types of users and their role to user's privileges
7. Taxonomy. This section deals with information on Nomenclature, ranks, authorships, and other details.
8. Ecological attributes. This stores the ecological requirements of the Pili crop.
9. Morphological Characteristics. This stores the knowledge on Pili from roots to tip including its properties.
10. Geological Distribution. This stores information about the location to GPS coordinates of individual Pili crop or farm.

2.5 Database Design

To facilitate the organization of Pili knowledge, the database is designed in a Structured Relational Model that can handle massive access and retrieval of data. The database is normalized to handle complex queries needed to generate the requested information. The tables in the database are categorized into four (4), namely:

1. Base information. The base information is comprised of the Taxonomy, Morphology, Botanical Description, Properties, Molecular Biology, Habitat and Ecology, and Cultivation and Uses. This will include all information needed to build and understand Pili knowledge when it comes to its physical properties and uses.

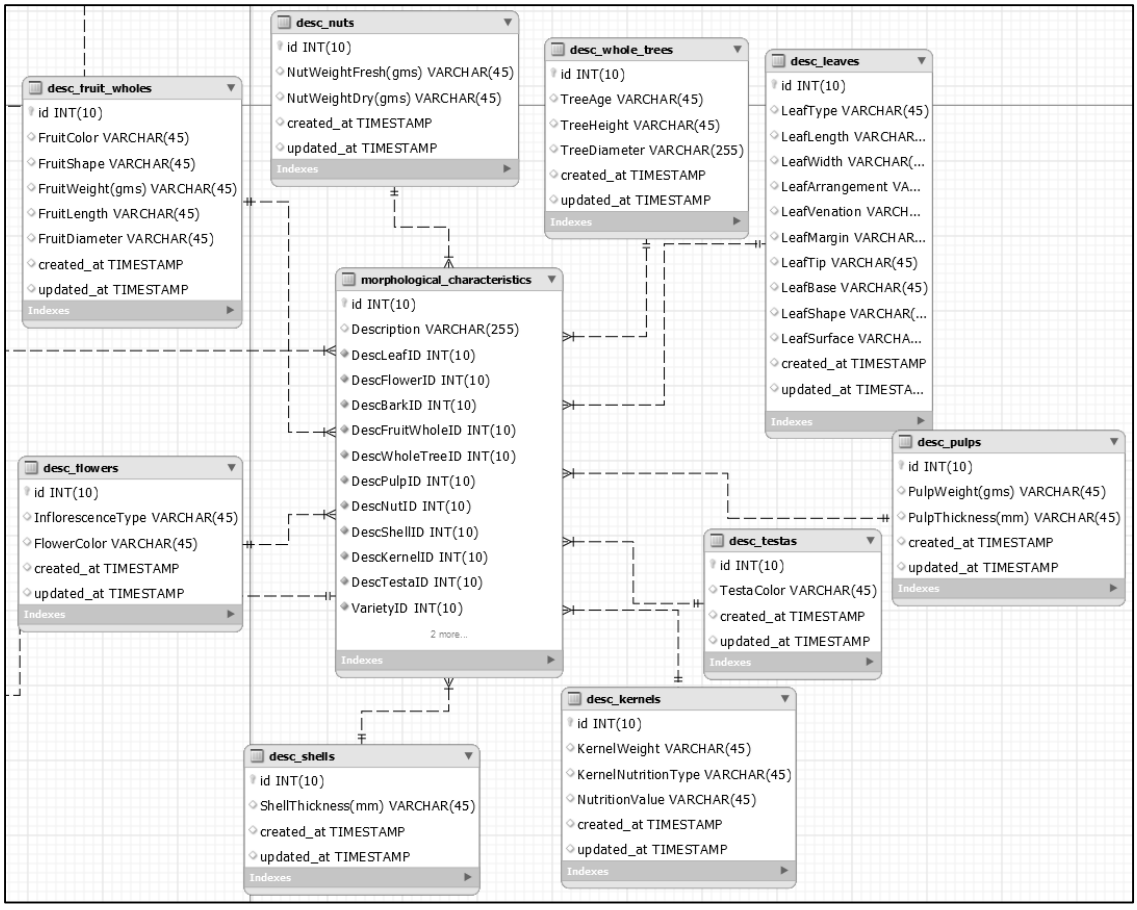

Figure 5. The Database Schema for Pili Morphological Characteristics

2. Promotions. This will include information needed to commercialize and promote Pili products and by-products.
3. Human Resources. Information on entities such as Pili growers, vendors, traders, processors, partnerships, oranizations alongside their location and contact information and their products. Information on agencies is also being sought in this category.
4. Research and Development. This will be comprised of information on Pili related studies and researches. This includes the authors, publications, patents, and research products along with other related data.

The schema depicted in Figure 5 will hold the information about the morphological characteristics of individual Pili NSIC registered variety. The following tables are: Tree characteristics (age, height, growth habit, fruiting season and others); Fruit characteristics (weight, length, width and shape); Pulp characteristics (color, weight, thickness, % based on whole fruit weight); Nut characteristics (weight, length, % based on whole fruit weight); Shell characteristics (weight, thickness, % based on whole nut weight); Kernel characteristics (weight, % based on whole nut weight, color); Leaf characteristics (type, length, width, arrangement, venation, margin, tip, base, shape, surface); Flower characteristics (type, color).

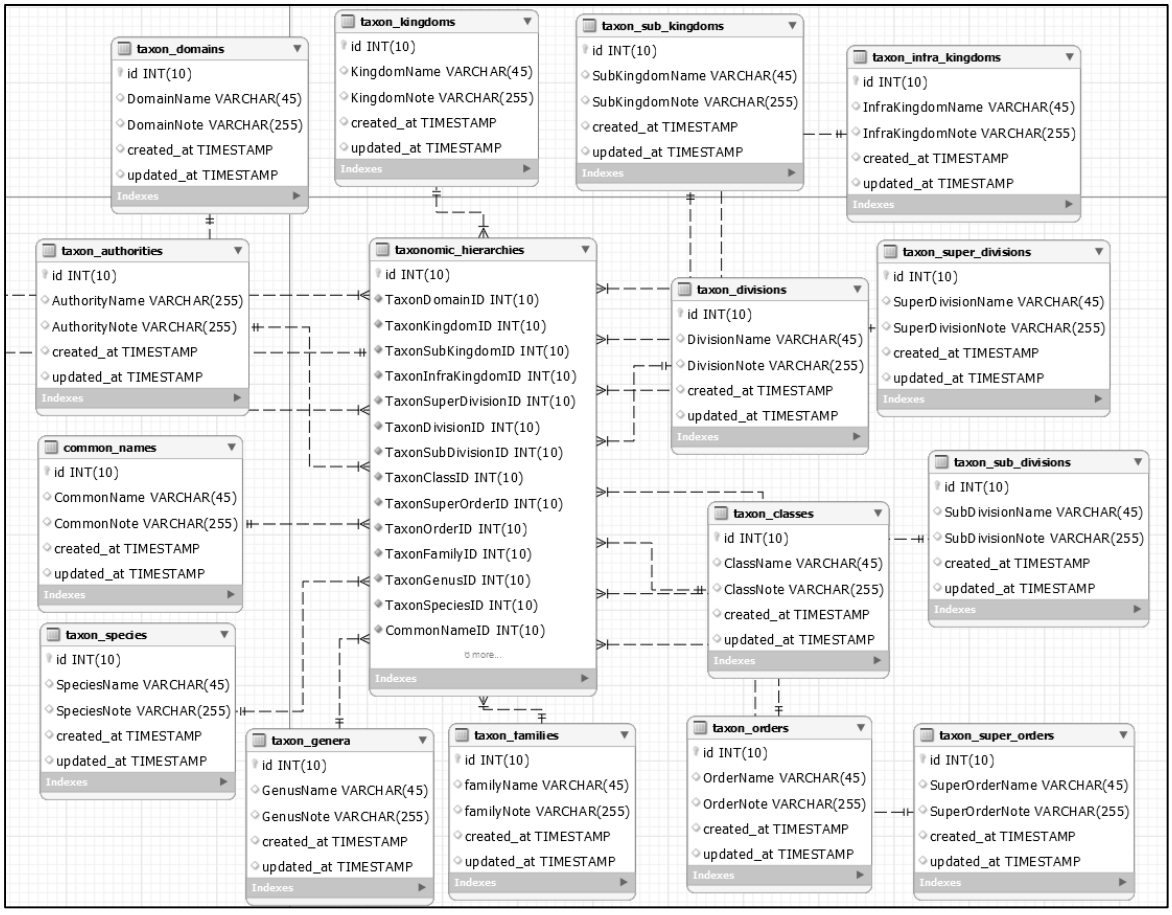

Figure 6. The Database Schema for Pili Taxonomic Hierarchy

The schema depicted in Figure 6 will hold the information about the Taxonomic Hierarchy of Pili (Canarium ovatum Engl.).

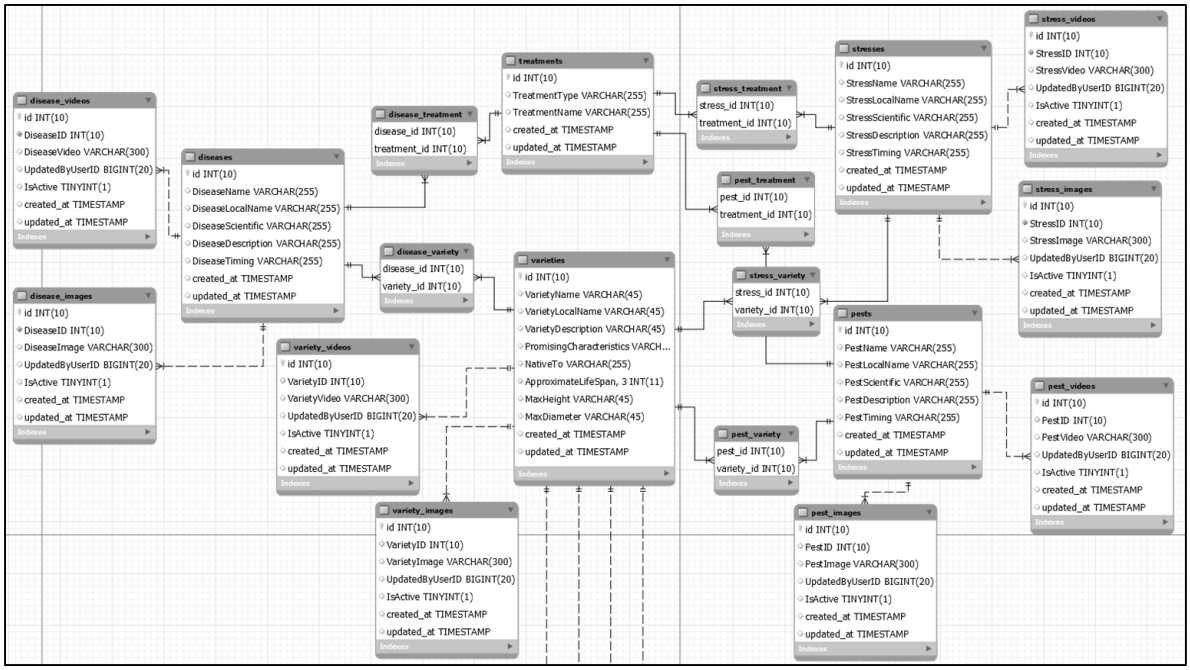

Figure 7. The Database Schema for Pili Variety Information

The schema depicted in Figure 7 will hold the information about the variety and other pertinent information of individual Pili NSIC registered variety.

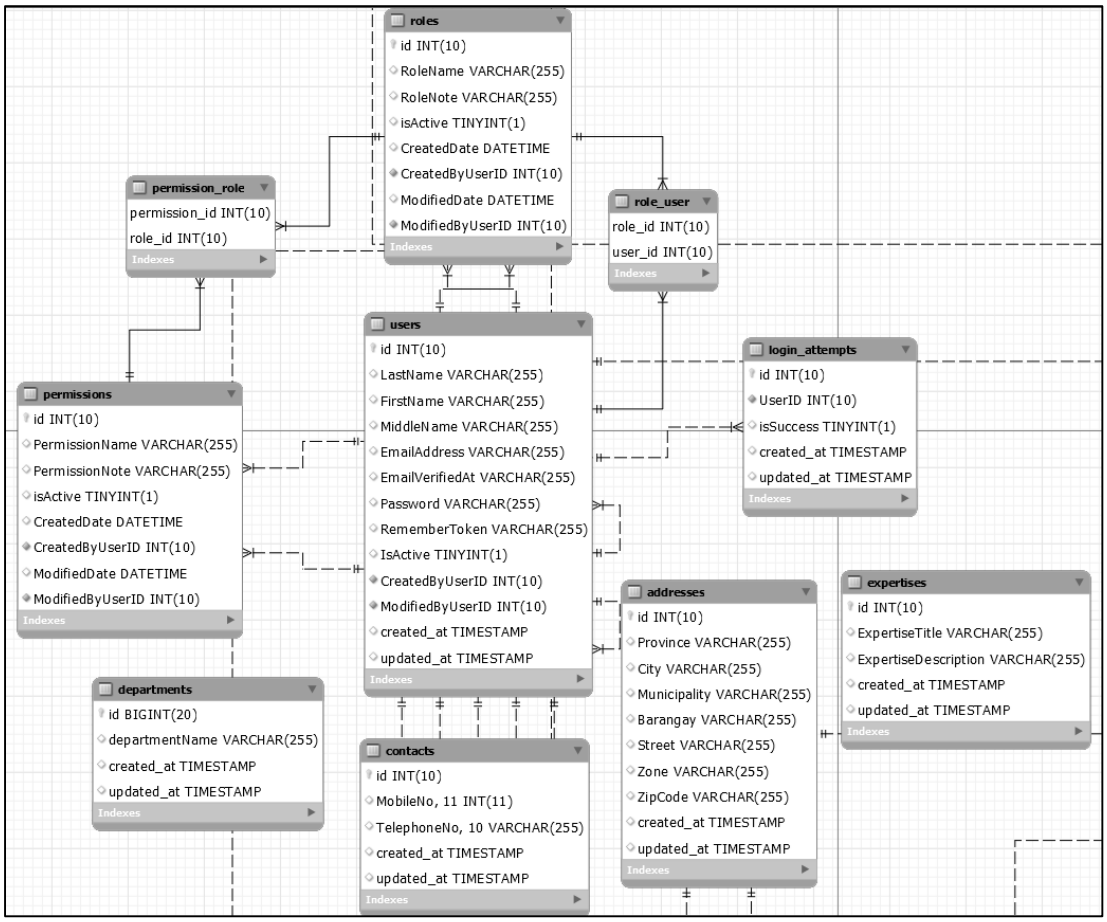

Figure 8. The Database Schema for Users, their Roles and Permission

The schema depicted in Figure 8 holds the information about the users, its roles and permissions. This includes LoginAttempt which stores all user's login attempts; User, user information and other relevant information; Contact, user's contact information; Role, list of user's roles for authorization; Permission, list of user's permission; UserRole, Links the user's to the UserRole table.

3. Results and Discussion

The Pili database serves as the central repository of data from the collaborative build-up of information. The data model is constructed in a structured method containing a variety of data types including texts, numbers (integers and floating-point number), dates, images, and hyperlinks. In Section 2, showed the schema for Morphological Characteristics, Taxonomic Hierarchy, Variety, User (their role and permissions). The database was expounded the Basic Information into three (3) subcategories of database tables, and come up with Botanical Information, Geographic Location, and Cultivation and Uses as shown in Figure 9. The tables are linked (represented by straight lines) showing the relationship of respective category to one another.

Nonetheless this is only a preliminary structural physical design of the database, the researchers have constructed, using the Third Normal form of Normalization, 180 tables, all interconnected by cardinalities. We are expecting that the database will expand and be optimized as inputs from our development team and the stakeholders arrive depending on the need.

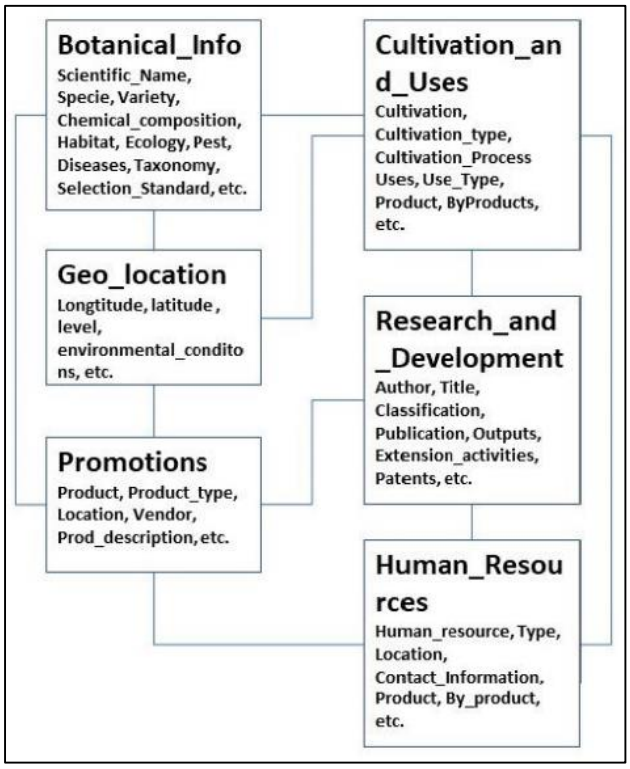

Figure 9. Representation of the Relational Database System of Pili Database

4. Conclusion

Presented in this paper is the design and development of the Pili database system towards KM-buildup through the use of the Structured Relational Database Model. It aims to provide a comprehensive centralized digital repository of Pili data that is not easily accessible to the general public using the

Internet as a medium for disseminating pili information to the public. Moreover, this study also facilitates users in inferring and gaining new data insights.

It is our hope that researchers, students, policy-makers, and other stakeholders around the world will find this repository useful for collaborative studies. The Pili processors, in particular, will dive into the largest existing compilation of nearby pili farms to become the supplier of raw products; marginalized farmers will find numerous information illustrating the facts, prospects, and potentials of the Pili industry.

References

[1] G. McLaren, A. Portugal, A. Cosico, W. Eusebio, T. Ulat, M. A. Sallan, V. J. Ulat, R. Bruskiewich, "*Integrated information systems for crop research and improvement*", in Proceedings of the World Rice Research Conference, Tokyo and Tsukuba, Japan, November 4-7, 2004, pp.579-583.

[2] S. Rajan, T. K. Sahu, L. P. Yadava, "*Mango resources information system: An open access portrayal of phenotypical, genetic and chemical information on mango*", Acta Horticulturae, no. 992, pp.99-104, doi: 10.17660/ActaHortic.2013.992.11.

[3] T. Řezník, V. Lukas, K. Charvát, K. Charvát Jr., Š. Horáková, M. Kepka, "*Foodie Data Models for Precision Agriculture*", 13th International Conference on Precision Agriculture, St. Louis, Missouri, USA, July 31–August 4, 2016.

[4] V. N. Villegas, R. E. Coronel, "*Note: cytology of pili and barobo [Diplodiscus paniculatus Turcz., study conducted in the Philippines, nut tree, angiosperm tree]*" Philippine Agriculturist, vol. 63, no. 2, 1981, pp.174-178.

[5] R. E. Coronel, "*Pili nut Cavarium ovatum Engl.*", in Promoting the Conservation and Use of Underutilized and Neglected Crops, J. Heller, J. Engels, K. Hammer, Eds., Institute of Plant Genetics and Crop Plant Research (IPK)/International Plant Genetic Research Institute (IPGRI), 1996, pp.57.

[6] M. Mendioro, M. G. Diaz, V. M. Velasco, M. Alcaraz, R. Lalamunan, K. Amoloza, L. Villameal, "*Genetic Characterization of Pili (Canarium ovatum Engl.) from Albay, Camarines Norte, and Camarines Sur Through Isozyme Analysis*", Philippine Journal of Science, vol. 137, no. 2, December 2008, pp.115-125.

[7] L. Xia, D. Zou, J. Sang, X. Xu, H. Yin, M. Li, S. Wu, S. Hu, L. Hao, Z. Zhang, "*Rice Expression Database (RED): An integrated RNA-Seq-derived gene expression database for rice*", Journal of Genetics and Genomics, vol. 44, no. 5, May 2017, pp.235-241, doi: 10.1016/j.jgg.2017.05.003.

[8] E. Tray, A. Leadbetter, W. Meaney, A. Conway, C. Kelly, N. Maoiléidigh, E. de Eyto, S. Moran, D. Brophy, "*An open-source database model and collections management system for fish scale and otolith archives*", Ecological Informatics, vol. 59, September 2020, pp.101-115, doi: 10.1016/j.ecoinf.2020.101115.

[9] Y. Rosen, R. Rimor, "*Using a Collaborative Database to Enhance Students' Knowledge Construction*", Interdisciplinary Journal of E-Learning and Learning Objects, vol. 5, 2009, pp.187-195.