The Pili knowledge bank: A comprehensive digital repository backend for Pili (Canarium Ovatum Engl.) information system

John Rey Lirag[1,2*], Ronnel Atole[1,2], Leo Constantine Bello[1,2], Rowell John Artiaga[2], Salvador Briones II[1,2]

[1] Information and Communication Technology Management Office, Partido State University, Camarines Sur, Philippines

[2] College of Arts and Sciences – Institute of Information Technology, Partido State University, Camarines Sur, Philippines

ABSTRACT

The organization and presentation of agricultural data are greatly facilitated by databases that are specially designed to allow ease of data entry and organized data display. The Pili knowledge bank together with its user-friendly interface was designed and implemented to showcase the knowledge collected from different Pili cultivars, enabling users around the world to see its prospects and potentials. This study, in general, supports and responds to the challenges entailed in the national government's having identified "Pili" (Canarium ovatum Engl.) as the "flagship commodity" of the bicol region, Philippines due to its competitive advantage for export and other potentials. This crop is also considered a "tree of hope" because of its usefulness from sap to roots. In particular, it seeks to address the dearth of accessible knowledge on the Pili industry through the creation of "Pili Knowledge Bank", a comprehensive digital repository, based on a relational database that stores information efficiently and in a standardized manner. The database also allows efficient queries, utilizing the social Web (Web 2.0) technologies, and is designed and structured towards knowledge management (KM) build-up. The authors observed that this intervention can be leveraged to design and implement the next generation of data, models, and decision support tools for the agricultural production system to benefit the agricultural community especially the marginalized farmers and vendors, and the country as a whole.

Keywords: Canarium ovatum Engl., Relational database, Web 2.0, Knowledge management.

1. INTRODUCTION

The fast-growing demand for food has transformed production systems around the world, and the digitalization of agriculture, as one of the biggest drivers of technological innovation, has played a major role in improving productivity. However, after reading studies related to the digitalization of agriculture, the majority of these digital solutions, focus on improving the productivity of major crops, as minor crops often are not included in the assessments and future recommendations of technological innovation in agriculture (Shepherd et al., 2018), particularly for vulnerable livelihoods (Capital, 2019). Minor or underutilized crops constitute a portion of agricultural biodiversity that is mainly produced and consumed by local communities around the world and can play an important role in the future of agriculture (Mabhaudhi et al., 2016). The rate of adaption by other farming communities and industry has seen a decline for many reasons, chief among them lack of support from the research community and industry (Tadele, 2019).

Digital repository of agricultural data for minor or underutilized crops like Pili is still

lacking these days. It is because, agricultural researches, food supplying agencies, and other government agricultural entities focused on staples like corn and rice and stretches less attention to minor crops (Coronel, 1996). Pili information on the current state, like germplasm and genetic potentials, are limited, scattered, and are usually stored in private collections that are not accessible by the general public affecting the Pili industry in general. Despite limited, scattered information and absence of digital repository of Pili information, there is grey literature expressing that Pili products have great potential in the world market and even hoped that soon will be ranked equal with cashew and macadamia (Coronel, 1996). The pili nut kernel (Fig. 1) is the most important part of the tree and has many uses. Pili nuts are mainly used to manufacture candies and confectioneries, while pili nut oil is high in demand locally and in foreign countries like Guam, Australia, Canada, and the United States (Mendioro et al., 2008). Aside from kernel other pili products like pili oil extract from either the pulp and elemi are also capturing the attention of the processors and multi-national companies such as Olay Fragrance, Chanel France, and Christian Dior. The oil is being used as the base for their perfume collections. Locally made pili pulp oil is now becoming a popular alternative to olive oil. An enterprise in Bicol produces 1,000 liters to 5,000 liters of pili pulp oil per month.

Pili trees naturally grow abundantly in many parts of Bicol and such is attributed to the volcanic soil and the favorable climatic condition of the region. 80% of fruit-bearing pili trees are concentrated within this region (Sorsogon (55.6%), Albay (25.3%), Northern Samar (10.4%)). In three years, an average of 7,456 metric tons production for years 2016-2018 and an average of 75 metric tons exported to different countries (PSA, 2019).

The internet provides an excellent mechanism for the dissemination of information. Since information about pili is lacking, limited, and scattered, the development of this crop has been greatly affected. Interventions were challenging to formulate due to inadequate resources. This calls attention to designing and implementing the Pili Knowledge Bank, a digital repository that allows unequivocally linking of all information associated with pili through the use of Web and Internet technologies.

This Pili Knowledge Bank is not only intended to compile and disseminate pili knowledge but also, to organize the pili information in a form that is understandable by both humans and computers to accelerate the development of bespoke solutions and trigger new standards and semantics in the identification of existing pili knowledge gaps and motivates more research works in breeding, farm management, agricultural best practices, nutritional aspects, ethnobotanical uses, and potential for derived novel products and marketing. In this case, it will become more beneficial to farmers, vendors, processors, and companies to be conscious of the prevailing market value of pili products and possibly sparks interest and further investment in the pili industry. Scientists, researchers, and students will be given access to a comprehensive repository of knowledge about pili which can be used to develop new insights and interventions to further uplift the pili industry.

The integration of Web 2.0 and semantic web technologies in the design of the Pili Knowledge Bank will open opportunities for participatory culture and interoperability for all the stakeholders. Semantic web technology was incorporated to make it certain that every bit of knowledge related to pili will be aligned and make it a bigger interwoven resource of knowledge. In this setup, the knowledge will encompass all the information needed by the diverse group of stakeholders (Drury et al., 2019). This collaborative environment will provide a culture of motivation and sparks interests among key players in the cultivation, production, and commercialization to adapt standards and move to the pace of what the world market needs.
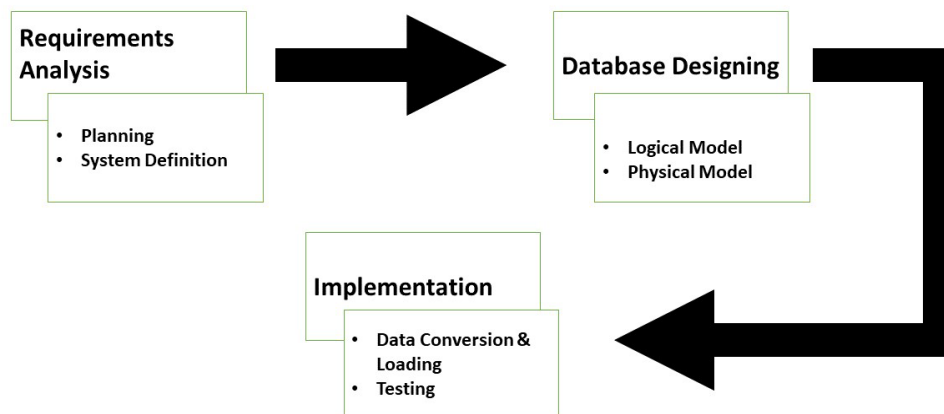

Fig. 1. Pili pulp, nut, kernel

Fig. 2. Database development life cycle

2. MATERIALS AND METHODS

2.1 Database Requirement Analysis and Scoping

The provision 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 Fig. 2 was considered as the most significant stage in the creation of the Pili Knowledge Bank. 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.

2.2 Data Gathering

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 in validating secondary data.

2.3 The Pili Knowledge Bank Design Concept

To facilitate ease of sharing and retrieval among various collaborating entities responsible for promoting, cultivating, and managing pili information this system architecture was observed as shown in Fig. 3. Entities, BCAARRD (Bicol Consortium for Agriculture, Aquatic and Natural Resources Research and Development), agriculture technician,

researches, dealers, farmers, and producers are part of the building process of pili knowledge and access through a collaborative interface (Web 2.0). Data and information gathered are stored in a structured centralized repository. Information not available in the repository is crawled/queried using an integrated Web 3.0 structure.

2.4 Data Vocabularies

A representation of different types of data can be generated from one pili sample as showed in Fig. 4.

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. This section of the database houses information about pests, diseases, and stress of the pili crop. The core structure accommodates basic information such as common names, local names, scientific names, and descriptions. Images, videos, and treatments are in separate tables, each contains a unique id to link the information as shown in Fig. 7. Information about its effects, timing, severity are also considered.

2. Product/use. This section will contain the different types of pili products and information associated with each product. Products are categorized into delicacies; beauty and wellness; fashion and accessories, and furniture as the product's major category, and each has sub-categories.

3. Agronomy. This part will hold the information about growth and development, lifecycle, reproduction, post-harvest methods, and other agronomic details related to the pili crop.

4. Research and development. This portion of the database will store the publications, utility models, inventions, articles, research outputs including news and updates on the pili crop.

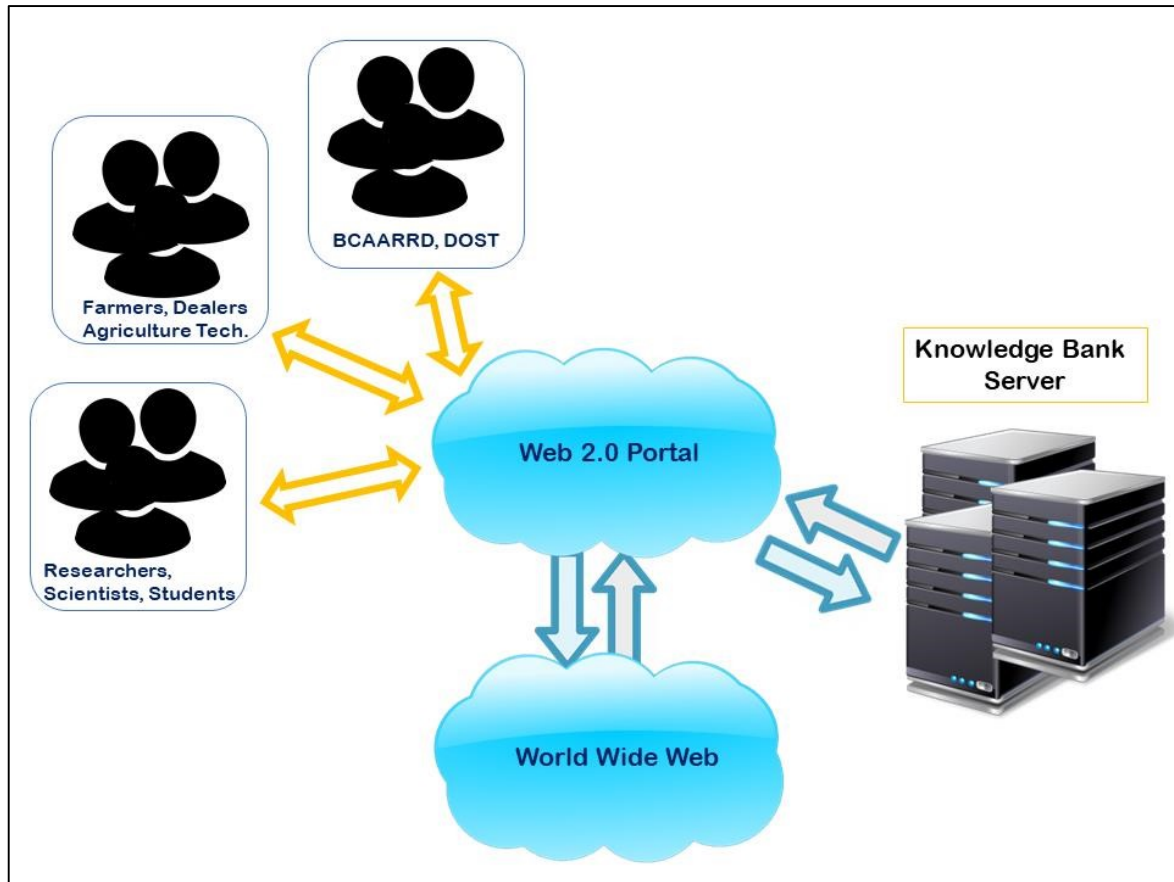5. Genetics. This portion will be accumulated by the pili genetic information.

Fig. 3. System architecture of the "Pili Knowledge Bank" information query and access
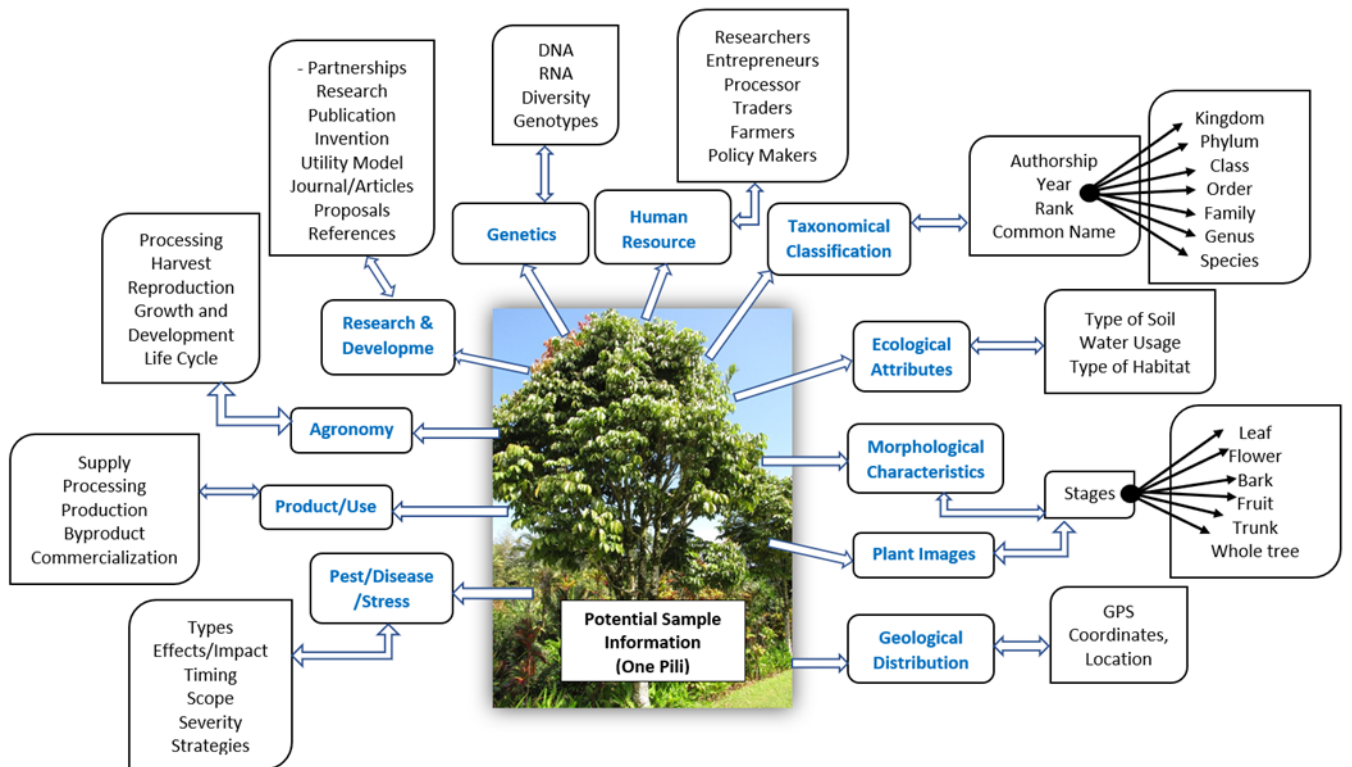

Fig. 4. Potential information associated with a pili sample

6. Human resources. This will store the different types of users e.g. students, researchers, policymakers, farmers, processors, vendor, etc., and their role and user's privileges.
7. Taxonomy. This section will be occupied by the taxonomic hierarchy information of the pili crop including its references.
8. Ecological attributes. This stores the ecological requirements of the pili crop.
9. Morphological characteristics. This part of the database will be stored the knowledge on pili morphological properties and also the references.
10. Geological distribution. This will be the basin of information about the location to GPS coordinates of individual pili crop or farm.

2.5 Database Construction and Content

Databases are considered as a knowledge medium, a representation of a combination of knowledge types, formats, and purpose of represented knowledge. Rosen and Rimor (2009) 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". In the study of Park et al. (2015), a database model has been developed for agricultural decision support systems for water-resource analysis, and for organizing experimental data by Driemeier et al. (2016). Although restructuring relational databases might pose complexity compared to newer technologies such as object-oriented or ontology-based databases, it is still considered a practical technology for storing data efficiently and permanently. Relational databases are well-suited with a wide array of software and applications for data distribution and or data propagation. It is also used to reduce redundancy where each record of data has its unique identifier. This section presents and discusses the database structure, design, and implementation.

The Pili Knowledge Bank database was implemented using MySQL, a relational database management system, fast, multi-threaded, multi-user, and robust. The data model for this database was developed through an iterative process of consultation with experts on the pili crop. This will store comprehensive collection and global information about Canarium ovatum Engl. The backend has been normalized to handled complex queries and avoid storage of data with logical inconsistencies. As data collection and curation through the database interface is a continuous process, the pili knowledge will become more comprehensive, and thus, this can provide rich information needed by the diverse stakeholders. For instance, a pili processor 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, and students will make use of journals, articles, utility models, inventions, research papers for their study references. As the data grows bigger, information queries can become very complex to retrieve many pieces of information stored from different tables. This ability makes this knowledge bank be able to answer common questions such as pili nutrition profile, potentials yields, lifespan, and many more. 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 as showed in Fig. 5 to Fig. 8.

At present, the database produced 180 normalized tables designed to handle complex queries needed to generate pili knowledge. The Pili Knowledge Bank is enriched by the collection of various pili related information ranging from its nomenclature, accession, National Seed Industry Council (NSIC) registration, promising characteristics, distribution, morphological characteristics, genetic resources, geo-mapping and characterization, propagation techniques, emerging pest, stress and diseases, growth and yield response, postharvest processing, to products, e-commerce, data analytics, technical experts and skilled technicians, partnerships, cooperatives, enterprises and other related data present in different cultivars.

The schema shown in Fig. 5 depicts the structure on how the information about the morphological characteristics of individual pili NSIC registered variety is constructed. 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).
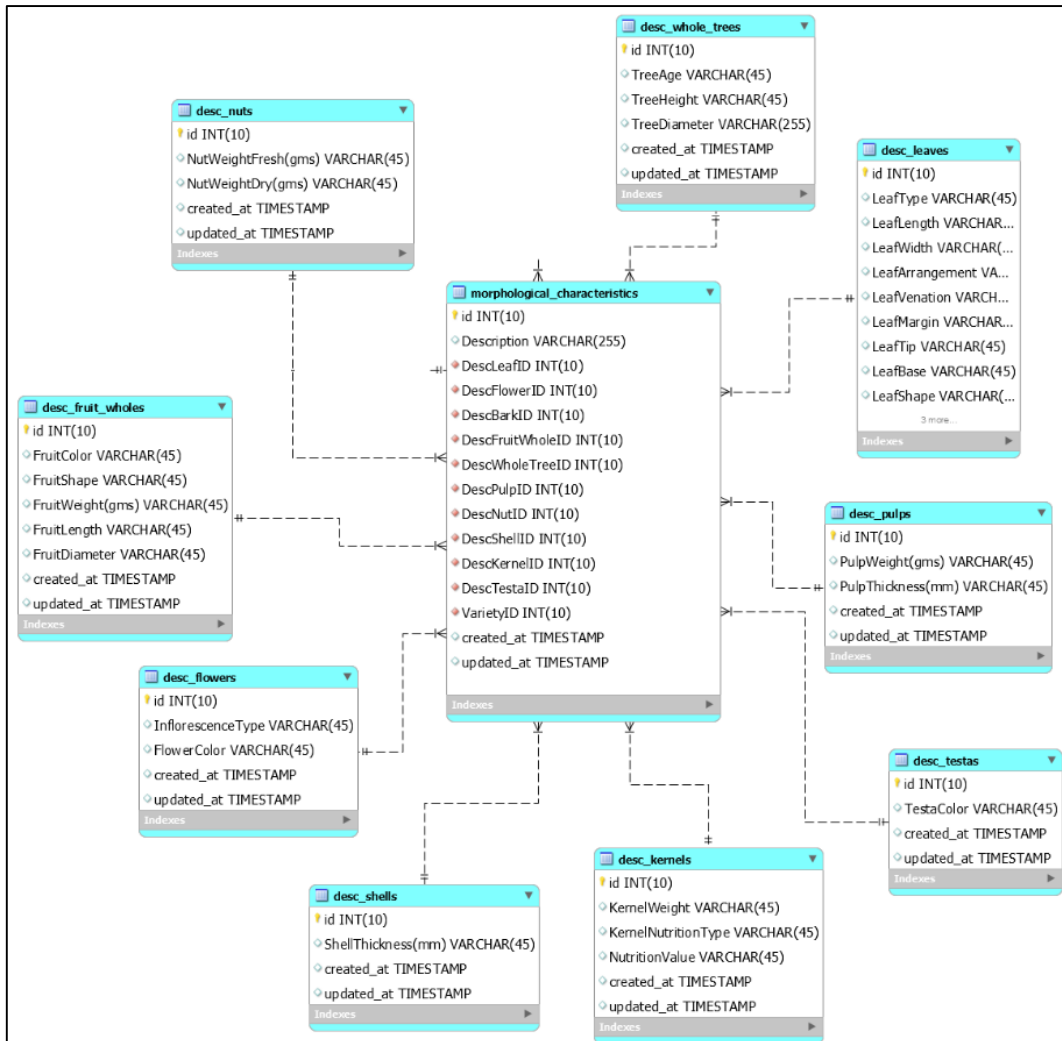
Fig. 5. Database schema for pili morphological characteristics

The schema shown in Fig. 6 presents the logical view of taxonomic information of pili in the database. Though this database will only hold one species (Canarium ovatum Engl.) in the current state, the taxonomic hierarchy schema was been set ready to cater for other crops soon to be associated in this Knowledge Bank.

The schema presented in Fig. 7 shows how pili variety information is organized systematically in the database. This also showcases how the data is organized and structured for pili pests and diseases including its treatment.

Fig. 8 showed how the users, their roles, and permissions are structured in the Pili database. 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 permissions; UserRole, Links the users to the UserRole table (Lirag et al., 2020).

The Pili Knowledge Bank database uses a windows server operating system and a laragon backend. The database is segmented into various tables to address every bit of information related to pili. The server is housed in a cooled, dehumidified server room at 20°C. Surge protection and 3 KVA uninterruptible power supply are provided. At any given time, at least three copies of the database are in existence to safeguard against catastrophic hardware failure or hacking.
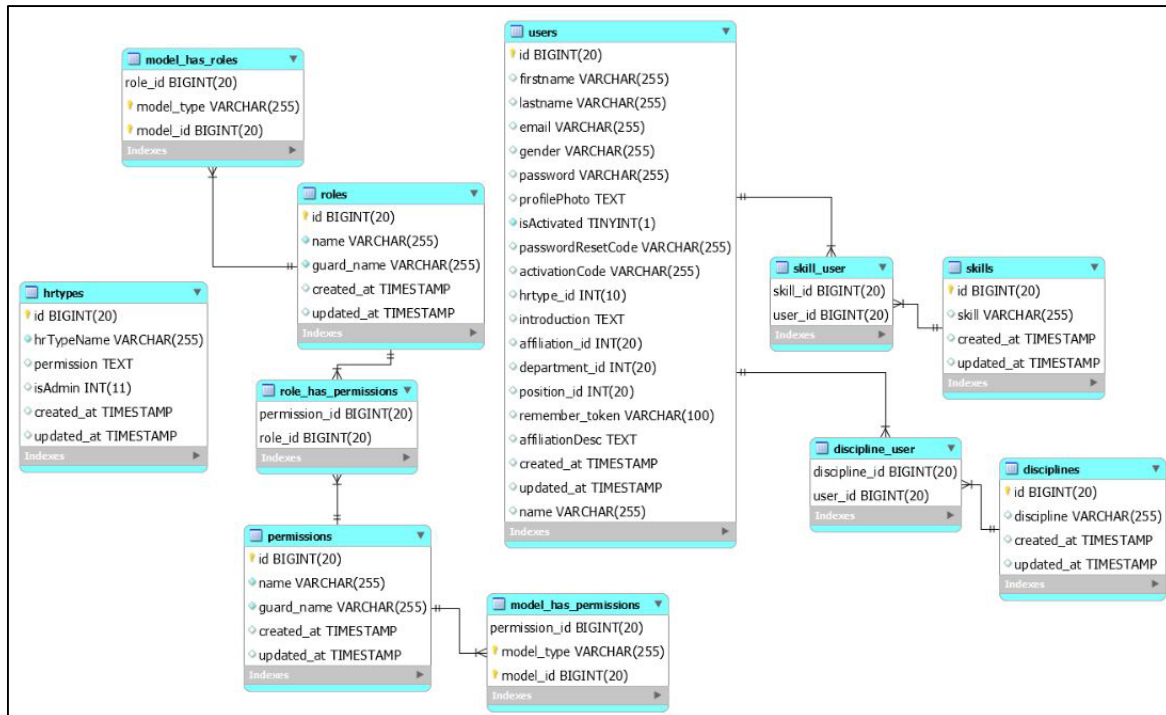
Fig. 6. The database schema for pili taxonomic hierarchy

Fig. 7. Schema for pili variety information

Fig. 8. The database schema for users, roles, and permission

3. RESULTS AND DISCUSSION

The presented Pili Knowledge Bank, a comprehensive digital repository, is based on a relational database that stores information efficiently and in a standardized manner. The repository references 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. The system is also equipped with a user-friendly interface for managing information as well as, ease of data entry and data display.

The interface as shown in Fig. 9 is programmed in PHP (Hypertext Preprocessor) using the Laravel framework. This conforms with the W3C XHTML 1.0 strict specification drafted by the World Wide Web Consortium. This results, 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 as showed in Fig. 10 and Fig. 11 and organized knowledge display as shown in Fig. 12. The interface includes a normal user interface for viewing only and an administrator functions overlay for logged-in users with adequate privileges.

Fig. 9 shows the landing page of the system. The users can search in News and Updates, Researches and Pili bioecology for easy retrieval of pili information stored in the knowledge bank. On the home page, can browse the uploaded pili news and updates. On the Bioecology page, can find the registered varieties; morphological characteristics; taxonomy; pest, diseases, and treatments; industry, and resources and tools. On the Pili R&D page, can be found the uploaded research studies for pili. The Pili Marketplace links to the e-commerce portal for pili products. This page also links to the signup page for user account creation.

Lists of Enterprises is one of the major information that can be found in this system including individual enterprise profiles. Information like enterprise name, type, product lines, enterprise address, owner's information, and other contact information is stored in the system.

Morphological characterization describes the structural features of pili. Fig. 11 provides an entry page for pili morphological characteristics. The entry page is categorized by pili parts and is labeled for ease of entry and for the content contributors to be guided on what information is needed in the specific entry field.

Fig. 12 shows how the system presents the morphological characteristics of a pili sample. The richness of information on this page will be of great help to the users needing this information and the likes.
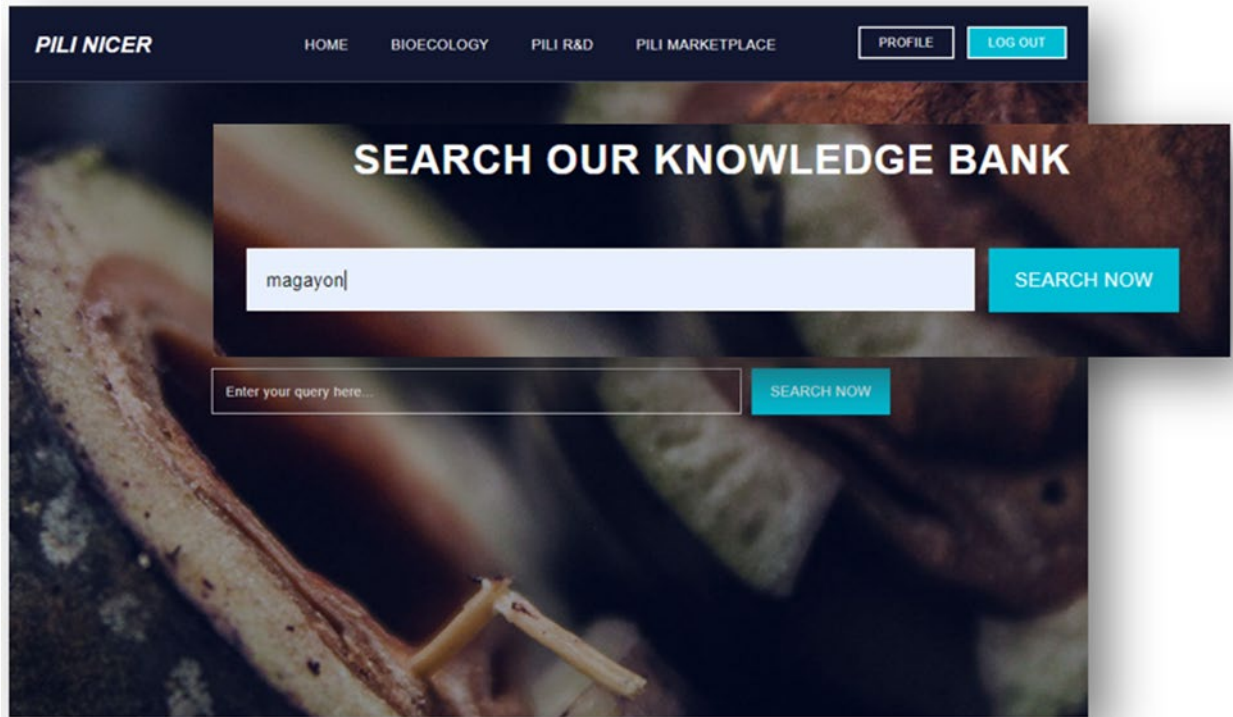
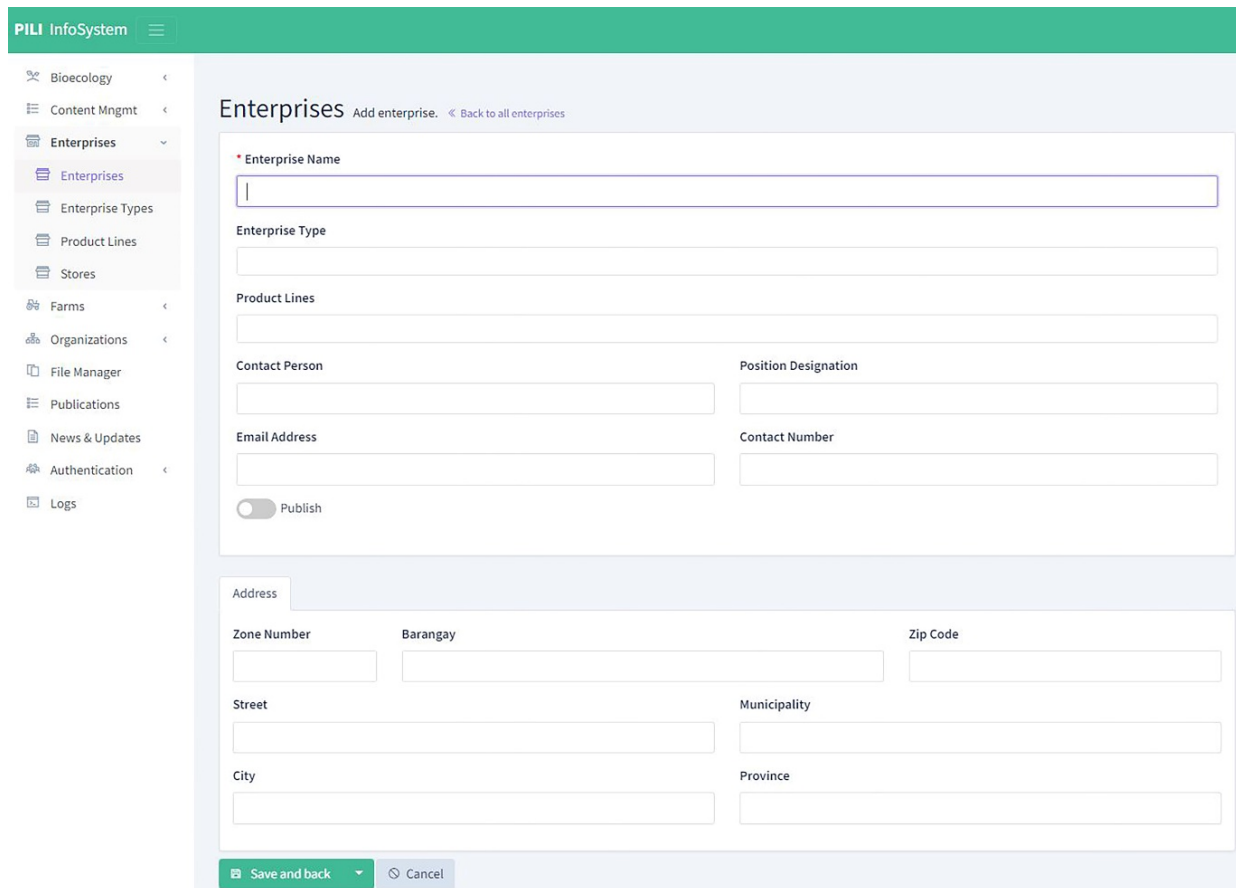**Fig. 9.** The Pili Knowledge Bank homepage interface

Fig. 10. Data entry page for pili enterprises

Fig. 11. Data entry page for pili morphological characteristics

Fig. 12. Morphological characteristics presentation of selected pili variety

4. CONCLUSION

We described the development of the Pili Knowledge Bank, a comprehensive digital repository for Pili (Canarium ovatum Engl.) that provides an easy way to store pili information to build its knowledge domain. After reviewing related pieces of literature, systems, and data standards, a novel data model was developed to allow seamless data entry and retrieval of pili information. An automated programming interface (API) and a graphical user interface (GUI) were designed and developed to allow for online access to the database for users and database managers. Also, this provides a convenient submission interface for the contribution of novel studies related to pili by independent researchers. Thus, the database will aid in the rapid and complete exploration of pili prospects and potentials as a commercial crop. Similarly, this tool can be an avenue to encourage decision-makers, government, and private sector partnerships to ensure the sustainability of the pili industry. The marginalized farmers, in particular, will find numerous pieces of information illustrating the facts, prospects, and potentials of the pili crop. As for the future direction of this study, further work will be done to refine the structure so that it can better accommodate scientific data and provide useful information in the future. Once the database has been populated with up-to-date information, this system can be a valuable resource to a diverse group of stakeholders. Also, this can open opportunities for comprehensive data analytics such as comparative morphology to the comparison of gene expression patterns.

ACKNOWLEDGMENT

REFERENCES

Capital, A. 2019. Digitalisation for agriculture: The case of MUIIS Uganda. Wageningen, The Netherlands: CTA. https://cgspace.cgiar.org/handle/10568/105931.

Coronel, R.E. 1996. 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), 57.

Driemeier, C., Ling, L.Y., Sanches, G.M., Pontes, A.O., Magalhaes, P.S.G., Ferreira, J.E. 2016. A computational environment to support research in sugarcane agriculture. Computers and Electronics in Agriculture 130, 13–19. https://doi.org/10.1016/j.compag.2016.10.002

Drury, B., Fernandez, R., Moura, M.F. et al. 2019. A survey of semantic web technologies for agriculture. Information Processing in Agriculture. http//doi.org/101016/j.inpa.2019.02.001

Lirag, J.R.S., Atole, R.R., Bello, L.C.S., Artiaga, R.J.B., Briones II, S.V. 2020. Design and development of a comprehensive, robust and KM-Scalable database (Backend) for a PILI (Canarium ovatum Engl.) Information System. Journal of Innovative Technology Convergence, 2, 63–74.

Mabhaudhi, T., O'Reilly, P., Walker, S., Mwale, S. 2016. Opportunities for underutilised crops in Southern Africa's post-2015 development agenda. Sustainability, 8, 302. https://doi.org//10.3390/su8040302.

Mendioro, M., Diaz, M.G., Velasco, V.M., Alcaraz, M., Lalamunan, R., Amoloza, K., Villameal, L. 2008. Genetic characterization of pili (Canarium ovatum Engl.) from albay, camarines norte, and camarines sur through isozyme analysis, Philippine, Journal of Science 137, 115–125, ISSN 0031–7683.

Park, J., Kang, M.S., Song, J.H., Jun, S.M., Kim, K., Ryu, J.H. 2015. Web-based data processing and model Linkage techniques for agricultural Water-Resource analysis. Journal of The Korean Society of Agriculture Engineers 57, 101–111. https://doi.org/10.5389/KSAE.2015.57.5.101

PSA, 2019. Supply utilization accounts (SUA) of selected agricultural commodities 2016 – 2018. Philippine Statistics Authority, 57–60. ISSN 2012–0508

Rosen, Y., Rimor, R. 2009. Using a collaborative database to enhance students' knowledge construction, Interdisciplinary Journal of E-Learning and Learning Objects, 5, 187–195.

Shepherd, M., Turner, J.A., Small, B., Wheeler, D. 2018. Priorities for science to overcome hurdles thwarting the full promise of the 'digital agriculture revolution. Journal of the Science of Food and Agriculture. 100, 5083–5092. https://doi.org/10.1002/jsfa.9346

Tadele, Z. 2019. Orphan crops: their importance and the urgency of improvement. Planta 250, 677–694. https://doi.org/10.1007/s00425-019-03210-6