

# Ontology-based Big Data Analysis for Orchid Smart Farming

**Nattapong Kaewboonma**

Department of Information Systems, Faculty of Management Technology,  
Rajamangala University of Technology Srivijaya, Thailand  
*nattapong.k@rmutsv.ac.th*

**Wirapong Chansanam**

Department of Information Science, Faculty of Humanities and Social Science,  
Khon Kaen University, Thailand  
*wirach@kku.ac.th*

**Marut Buranarach**

Data Science and Analytics Research Group,  
National Electronics and Computer Technology Center,  
Pathum Thani, Thailand  
*marut.bur@nectec.or.th*

## ABSTRACT

*Background.* Precision agriculture or smart farming is becoming more and more important in modern orchid farming in Thailand. Sensing and communication technologies have witnessed explosive growth in the recent past. These technologies are empowering information systems from many domains such as health care, environmental monitoring and farming, to collect and store large volume of data.

*Objectives.* The research aims to develop an ontology for big data analysis for the smart farming in Rajamangala University of Technology Srivijaya (RUTS), Nakhon Si Thammarat campus.

*Methods.* The ontology design and development process comprises: (1) Ontology design: the domain ontology provide vocabularies for concepts and relations within the orchid domain, and information ontology which specifies the record structure of databases; (2) Ontology development, which consists of five processes: (i) defining the scope, (ii) investigating the existing ontologies and plan to reuse, (iii) defining terms and its relations, (iv) create instances, and (v) implementation and evaluation.

*Results.* The research outcome is the domain ontology and information ontology wherein 11 concepts of smart farming were identified and classified into classes and sub-classes.

*Contributions.* The system is designed for assisting orchid farmers by giving recommended measures and expected results based on the knowledge extracted from best practices.

## INTRODUCTION

Thai orchids are popular in the global market. It has the largest export volume among Thailand's cut flowers. In 2019, the value of exported orchid cut flowers, orchid plants and orchid products from Thailand was US\$6.81 million (Department of International Trade Promotion, 2019). Thailand, the world's largest orchid producer and exporter, has an abundance of exotic orchids. There are more than 1,000 species of orchids in Thailand, though some species are rarely seen now.

91% of the orchid flowers exported from Thailand are orchid inflorescences or stems. The *Dendrobium* alone, accounted for 96% of orchid genera that are exported as stems. The rest are *Mokara*, *Aranthera*, *Oncidium*, *Aranda*, *Vanda* and *Arachnic*. Other orchid products are loose blooms, garlands, dried flowers, bouquets, and corsages (Supnithi et al., 2011). Japan, USA, Vietnam, China and Italy are Thailand's major trade partners for orchid flowers, while the Netherlands, Brazil, India, Germany and Japan are the major trade partners for orchid plants.

Nowadays, more and more modern orchid farms in Thailand are adopting precision agriculture or smart farming. This development emphasizes the use of information and communication technology in the cyber-physical farm management cycle. Sensing and communication technologies have witnessed explosive growth in the recent past (Wolfert, Ge, Verdouw, & Bogaardt, 2017). These technologies make it possible to collect and store large volumes of data from many domains including healthcare, environmental monitoring and farming.

New technologies, such as, the Internet of Things (IoT) and cloud computing, are expected to leverage this development and introduce more robots and artificial intelligence in farming. This is encompassed by the phenomenon of big data, where massive volume of data with a wide variety can be captured, analyzed and used for decision-making (Wolfert et al., 2017). In addition, the recently developed sensor Web technologies enable sensor measurements from all kinds of sources to be available for sharing through Web services.

The main challenge of orchid production is the higher costs of production while the selling price remains the same or goes lower. Moreover, there is a scarcity of labor, so farmers have to rely on non-Thai workers. However, both government agencies and private sector organizations support the orchid industry by providing knowledge of orchid production technology to farmers to increase the output quality.

This paper describes an ontology-based big-data analysis system that provides recommendations based on environmental factors, to improve orchid production.

## LITERATURE REVIEW

Ontologies can usefully support information systems as explicit conceptual knowledge models that represent domain knowledge. Moreover, for semantic Web applications, ontologies play a key role of providing semantic vocabulary which is meaningful to annotate websites for machine interpretation. In the context of information systems, ontologies borrow from the fields of symbolic knowledge representation in artificial intelligence, from formal logic and automated reasoning, and from conceptual modeling in software engineering, while also building on Web-enabling features and standards (Grimm, Abecker, Völker, & Studer, 2011). "Although hierarchy is not a defining characteristic of ontologies, it is an important component in the representational model prescribed by the Resource Description Framework (RDF) model and syntax specification" (Jacob, 2003, p. 20).

“RDF defines a model and a set of elements for describing resources in terms of named properties and values. More importantly, however, it provides a syntax that allows any resource description community to create a domain-specific representational schema with its associated semantics. It also supports incorporation of elements from multiple metadata schemas. This model and syntax can be used for encoding information in a machine-understandable format, for exchanging data between applications and for processing semantic information. RDFS complements and extends RDF by defining a declarative, machine-processable language—a meta-ontology or core vocabulary of elements—that can be used to formally describe an ontology or metadata schema as a set of classes (resource types) and their properties; to establish relationships between classes, between properties and between classes and properties; and to specify constraints on properties.” (Jacob, 2003, p. 20-21).

This research has constructed an ontology-based big data analysis system using Web Ontology Language (OWL) to describe the scope of knowledge and relationships of data on orchid farming. Meanwhile, a rule language (Rattanasawad, Runapongsa, Buranarach, & Supnithi, 2013) was employed to construct rules to enable the computer to infer and provide recommendations that are consistent with each growing orchid factors.

## **RELATED WORK**

Three previous case studies are very close to our current research. They suggest how the ontology-based search and semantic technologies can be exploited and used to enhance the functionality of smart farming.

Ngo, Le-Khac and Kechadi (2018) proposed a knowledge framework to be used in an agriculture ontology which can be employed in smart agriculture systems. Starting from the needs of a knowledge model for wheat, the AgriOnt is then built. There are three steps in the ontology development progress: (1) building a domain-specific knowledge hierarchy, (2) defining slots of the categories and representing axioms, and (3) knowledge acquisition filling in the value for slots of instances. Four thematic subdomains are presented in the new agricultural ontology: agriculture-based subdomain, geographical ontology, IoT-based subdomain, and business subdomain. Classes in each subdomain derive from a general class, entity, and two of its sub-classes. For instance, the agricultural subdomain covers basic classes in the agriculture domain, such as Farm, Crop, Product, Fertilizer, or Condition. The IoT-based subdomain has the main classes: ObserveSystem, Sensor, FeatureOfInterest, and ObservationValue. This ontology also incorporates weather data into agricultural datasets as the weather condition is one important factor that impacts crop yields.

A study by Khummanee, Wiangsamut, Sorntepa and Jaiboon (2018) focused on developing automated smart farming for *Dendrobium Sonia Bomjo* cultivation by applying fuzzy logic and IoT to control all the essential environment variables inside a greenhouse. The research on procedures to build the automatic smart greenhouse, proposed seven steps, labeled A to G: (A) Collecting information from orchid farmers, (B) Translating crisp inputs to fuzzy sets, (C) Creating a knowledge-base, (D) Simulating human reasoning, (E) Transforming the fuzzy sets into a crisp output, (F) Controlling actuators, and (G) Loop to D step. Input sensors (A) include the sensors for capturing environment variable values record of temperature, humidity, light, and soil moisture. The actuators include fogs, light bulbs, fans, sprinkler pumps, LEDs, and motors for controlling plastic curtains. This proposed system can automatically control the growth factors of orchids' inflorescences. The final results show that orchids can thrive constantly, with an average growth rate of 27.38 cm per week.

Muangprathub et al. (2019) developed wireless sensor networks for crop watering. They enhanced the control system between the node sensors in the field and data management via smartphone and Web application. Three important components are: (1) the control box to connect and obtain data of crops, (2) Web application to handle crop data and field information, and (3) mobile application to control crop watering. Data mining was applied to the data to predict suitable temperature, humidity, and soil moisture for future crops. The result of this work was useful for determining suitable moisture content in the soil for vegetables and thus helped to reduce costs and increase agricultural productivity.

## METHODS

The *ontology development* process consists of: (1) defining the scope, (2) investigating the existing ontologies and plan to reuse, (3) defining terms and its relations, (4) create instances, and (5) implementation and evaluation.

Based on the primary resources and the domain experts, the ontology was divided into two sub-domains of *ObjectType* (domain ontology) and *Measurement* (information ontology). *ObjectType* consists of the concepts, attributes and instances associated with Thai orchids. The objective of the *ObjectType* is to support semantic matches when searching for knowledge objects. *Measurement* is a meta-model that describes knowledge objects such as *Measure\_Day*, *Measure\_Sensor*, *Temperature\_Value*, *Humidity\_Value*, *Lighting\_Value*, *SoilMoisture\_Value* and other related information.

For defining terms, we grouped similar meaning words into concepts, with the assistance of domain experts. We defined relations among concepts based on the primitive relations *is-a*, *part-of* and *attribute-of*.

The *is-a* relation is a class-subclass relation that specifies a hierarchical relation between broader and narrower concepts, connecting them into a taxonomy. A *part-of* relation that represents a whole-part relation and attribute-of relation which represents the properties of a class are designed to construct a concept definition. *Is-a* and *part-of* connect classes. The *attribute-of* relation is defined by primitive data types such as, string, integer, and Boolean. For example, root, leaves, flowers are *part-of* orchid.

## RESULTS

The ontology was developed using the HOZO ontology editor, a Java-based graphical editor created to produce heavy-weight and well thought out ontologies from Mizoguchi Laboratory, The Institute of Scientific and Industrial Research, Osaka University (Kozaki, Kitanura, Ikeda, & Mizoguchi, 2002). The scope of the ontology development focused on sensor measurements data and knowledge of Thai orchids in RUTS, Nakhon Si Thammarat campus.

The research outcome is the domain ontology and information ontology for big data analysis for smart farming. Eleven concepts of smart farming were identified, defined and organized into classes and sub-classes. The scope of the ontology development focused on Thai orchid data and orchid production, for example, the concepts of *Orchids*, *Sensor*, *Measurement*, *ObjectType*, *Object*, *Day*, *DayofWeek*, *Result* and *Record Result* (Figure 1).

A class of *Measurement* is defined as the main class in RUTS Orchids Smart Farming with defined properties: *Measure\_ID*, *Measure\_Day*, *Measure\_Sensor*, *Temperature\_Value*, *Humidity\_Value*, *Lighting\_Value*, *SoilMoisture\_Value* and *Suggest\_Results* (Figure 2). Modeling the class hierarchy relied on domain experts. A class hierarchy of *Object*, *ObjectType*, and *Orchid* class were defined. We applied the *is-a* relation and the *part-of* relation to define class hierarchy as shown in Figure 3. For example, the *Object* class, defined

by two *attribute-of* relation: *Object\_Type\_ID*, and *Object\_Type\_Name*, contains a *part-of* relation with *ObjectType* class.

### Rule Engine

The Jena rule and inference engine (Rattanasawad et al., 2013) is used to compose the measurement rules so that the query outcome can be delivered to the users. The Ontology Application Management (OAM) framework (Buranarach, Supnithi, Thein, & Ruangrajitpakorn, 2016) is used for rules development. By the arrangement of orchid and farming knowledge for particular flora cultivate conditions, clear rules can be developed. Those rules then can be used to clarify the appropriateness factors for an orchid smart farm with defined flora cultivation conditions. Figure 4 shows example rules, indicating that low temperature and medium humidity, lighting and soil moisture are inferred as good environment factors for *Cattleya* orchids. The rule-based knowledge was derived from recommendations from an orchid farming manual. The recommendations were subsequently transformed into rule-based knowledge using the OAM framework. Rule conditions can be defined based on the terminology defined in the Thai orchid ontology. Recommendation results must also be well-defined in creating inference outcomes. Figure 4 also shows the measurement conditions and recommendations rules that can be applied. According to the rules, the value constraints of orchid properties, such as temperature, humidity, lighting and soil moisture, will be used to assess the environment factors. The assessment is provided as the recommendation outcome.

### Application of Expert Recommender Engine

The Orchid Smart Farming Recommendation System (OSFRS) was developed to link information between the ontology and knowledge-based repository. The measured values from sensor devices (IoT) are stored in the relational database management system which will match concepts and that database in order to transform the recorded data into the ontology (classes and instances) in the RDF format. Consequently, the rule-based knowledge is used to generate recommended outcomes. The users can query the data to obtain the environment factor assessment results that are suggested (Figure 5). In querying for orchids data, query condition from users is transformed into a SPARQL query to retrieve the measurement and assessment results from the OSFRS. The query conditions depend on the defined properties and classes in the Thai orchid ontology.

The steps of system processing of a user request can be summarized as follows:

- 1) A user may select the *has\_object\_type* property and select a type of plant, such as *Orchids*, *Cattleya*, or *Vanda*, which are some defined classes in the ontology. The query in SPARQL format is then submitted to the recommender system.
- 2) The system subsequently retrieves the results from the knowledge base, which utilizes an RDF database. The knowledge base has applied the recommendation rules, which generated the recommendation results in the knowledge base for sensor measurements recorded for each day.
- 3) The user can subsequently see “Result”, which is recommended assessment result of each measurement record. For example, Figure 6 shows that two sensor measurements recorded for *Cattleya* on May 1, 2019 are considered Good and Moderate respectively. Thus, the user can see that only 50 percent of the measurements are considered good environment factors for the orchids. The results could lead the user to examine and improve on the moderate results subsequently.

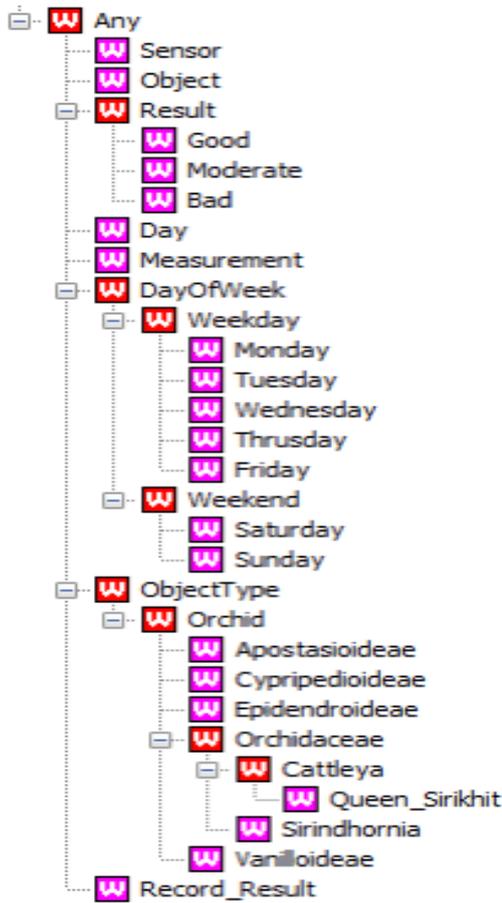


Figure 1. Concepts of *OrchidsSmartFarming* class

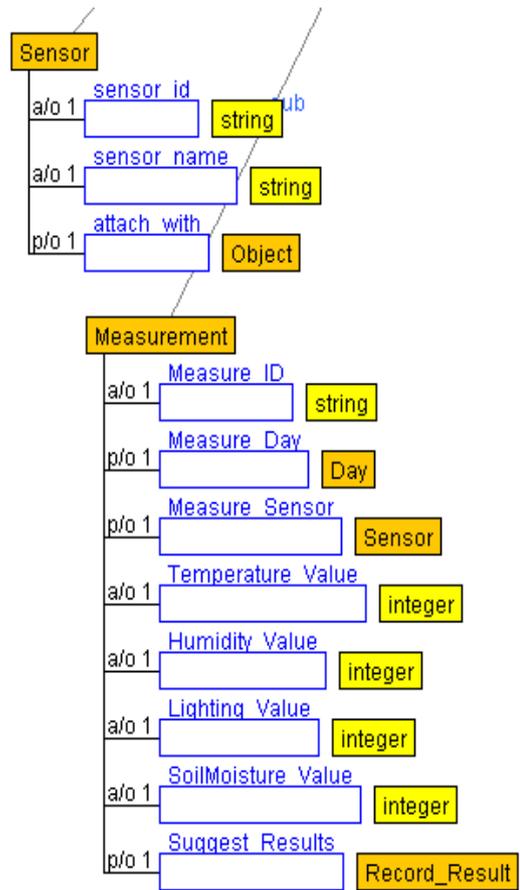


Figure 2. Properties of *Measurement* class

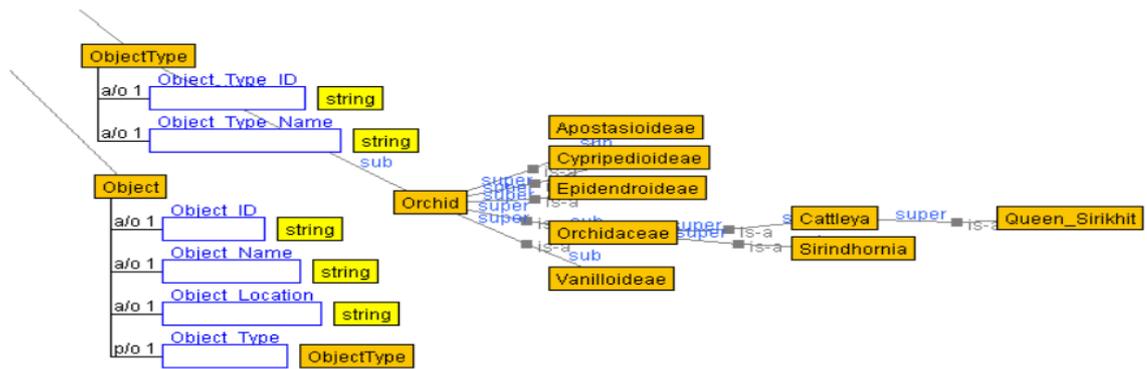


Figure 3. Class hierarchy of orchid

Orchids		Temperature			Humidity			Lighting			Soil Moisture			Result		
Cattleya	Vanda	Low <25	Medium 25-30	High >30	Low 45%	Medium 50-60%	High >60%	Low <50%	Medium 50-60%	High >60%	Low <50%	Medium 50-60%	High >60%	Good	Moderate	Bad
X		X				X			X			X		X		
X			X			X			X			X		X		
	X		X				X		X			X		X		
	X	X				X			X			X		X		

Recommended to "Measurement"

Recommendation of "Result"

Figure 4. Recommended measures and expected results

## CONCLUSION

This paper describes the development of an ontology-based big data analysis for orchid smart farming. It uses a recommender system with a knowledge-based framework to provide specific options for new farmers in Thailand. The ontology-based technique was proposed according to agriculture and emerging trends in big data technology. The knowledge-based creation emphasizes an ontology and rule-based knowledge development. An on-going effort is implementing real-time data collection with sensor devices and automatic updating of the data repository without farmer intervention. After the sensor input, recommendations can be updated for enhanced farmer requests. The knowledge-based orchid smart farming recommendation system can be extended to support some smart farm application improvement, for example, developing a training system for new young farmers or supporting the technologists and other professionals. Our future work will include the evaluation of the new Thai orchid ontology by orchid specialists, and the evaluation of the recommendation system based on an orchid farmer survey.

## REFERENCES

- Buranarach, M., Supnithi, T., Thein, Y.M., & Ruangrajitpakorn, T. (2016). OAM: An ontology application management framework for simplifying ontology-based semantic Web application development. *International Journal of Software Engineering and Knowledge Engineering*, 26(1), 115-145.
- Department of International Trade Promotion. (2019). *Thai orchids*. [https://www.ditp.go.th/contents\\_attach/539560/539560.pdf](https://www.ditp.go.th/contents_attach/539560/539560.pdf)
- Grimm, S., Abecker A., Völker J., & Studer, R. (2011). Ontologies and the semantic web. In J. Domingue, D. Fensel, J.A. Hendler (Eds.), *Handbook of semantic web technologies*. Springer.
- Jacob, E. (2003). Ontologies and the Semantic Web. *Bulletin of the American Society for Information Science and Technology*, April/May 2003, 19-22. <https://asistdl.onlinelibrary.wiley.com/doi/pdf/10.1002/bult.283>
- Khummanee, S., Wiangsamut, S., Sorntepa, P., & Jaiboon, C. (2018). Automated smart farming for orchids with the Internet of Things and fuzzy logic. In *Proceedings of International Conference on Information Technology (InCIT)*, 24-26 October 2018, Khon Kaen, Thailand (pp. 1-6). <https://doi.org/10.23919/INCIT.2018.8584881>
- Kozoki, K., Kitanura, Y., Ikeda, M., & Mizoguchi, R. (2002). Hozo: An Environment for building/using ontologies based on a fundamental consideration of "role" and "relationship". In *Proceedings of 13th International Conference Knowledge Engineering and Knowledge Management (EKAW2002)*, Sigüenza, Spain, 1-4 October (pp. 213-218). Springer. [https://doi.org/10.1007/3-540-45810-7\\_21](https://doi.org/10.1007/3-540-45810-7_21)
- Muangprathub, J., Boonnam, N., Kajornkasirat, S., Lekbangpong, L., Wanichsombat, A., & Nillaor, P. (2019). IoT and agriculture data analysis for smart farm. *Computers and Electronics in Agriculture*, 156, 467-474.
- Ngo, Q.H., Le-Khac, N.A., & Kechadi, T. (2018). Ontology based approach for precision agriculture. In *Proceedings of 12th International Conference on Multi-disciplinary Trends in Artificial Intelligence (MIWA2018)*, Hanoi, Vietnam, 18-20 November 2018 (pp. 175-186). Springer.

- Rattanasawad, T., Runapongsa, S.K., Buranarach, M., &Supnithi, T. (2013). A review and comparison of rule languages and rule-based inference engines for the semantic Web. In *Proceedings of International Computer Science and Engineering Conference (ICSEC)*, 4-6 September 2013, Nakorn Pathom, Thailand (pp. 1-6). IEEE Explorer. <https://doi.org/10.1109/ICSEC.2013.6694743>
- Supnithi, T., Paiboonrat, P.,Buranarach, M., Lekawatana, S., Intalak, A., Sangtien, S., ... Choptham, W. (2011). Ontology based orchid knowledge platform for knowledge services in orchid cluster. In *Proceedings of 2011 Annual SRII Global Conference, 29 March-2 April 2011, San Jose, CA, USA* (pp. 540-549). IEEE Explorer. <https://doi.org/10.1109/SRII.2011.6>
- Wolfert, S., Ge, L., Verdouw, C., & Bogaardt, M. (2017). Big data in smart farming: A review. *Agricultural Systems*, 153, 69-80.