Applying Ontology Knowledge Representation Technology and Semantic Searching Methods to Support the Production of High Quality Longan Fruit

Chapter · January 2020

DOI: 10.1007/978-981-15-1465-4_59

Winai Bangkhomned
Naresuan University

Applying Ontology Knowledge Representation Technology and Semantic Searching Methods to Support the Production of High Quality Longan Fruit

Winai Bangkhomned[1] and Janjira Payakpate [2]

[1,2]Department of Computer Science and Information Technology, Faculty of Science, Naresuan University, Phitsanulok, 65000, Thailand

*n3467@hotmail.com ; janjirap@nu.ac.th

Abstract. The objectives of this research were to develop and ontology and apply semantic search methods to the management of knowledge on the production of the tropical fruit, longan, in Northern Thailand. Factors affecting the quality of longan includes thinning of flowering crops, timing and methods of pruning the trees, the use of potassium chlorate to stimulate flowering, appropriate fertilization, and eradication of diseases and pest insects. The information gained from a substantial literature review on these matters is presented in three sections in the paper: (1) knowledge collection, and (2) the development of an ontology consisting of 7 major information classes, and (3) illustration of the development of ontology-based information accessing systems with semantic searching capabilities. Our testing of information retrieval from the ontology using semantic searching showed a Precision of 100% for information retrieved, with Recall of 93.50% and F-measure 96.64%. These results supported the further implementation of the Thai Longan Production Knowledge Service System.

Keywords: Semantic web, Production, Ontology, Knowledge-base, Longan

1 Introduction

The longan is a tropical fruit that is a significant economic crop in the Northern Part of Thailand (and elsewhere in Southeast Asia). Supplying a highly competitive current market, in 2017, 743 kilograms per rai were produced, and 800 kilograms per rai were produced in 2018: a rai = 0.16 hectares or about 0.45 acres. The total production in 2017 was 625,750 tons and in 2018 was 682,249 tons. Market prices vary considerably, with the 2017 crop averaging a selling price of 22.35 baht in 2017 and 26.72 baht in 2018[1].

While seasonal variations in price can fluctuate, the quality of the longan fruit is also an important factor affecting to the selling price. Our purpose in the research was to investigate and systemize the essential information regarding quality improvement for longan crops. We found that many farmers who regularly produce a crop of longan have very little scientific knowledge to assist them in producing high quality fruit. For example, fruit thinning is necessary as a quality control measure for the final fruit crop, yet most farmers have avoided this process out of ignorance and the fear of losing longan quantity, often resulting in large quantities of poor quality fruit being produced.

To assist in overcoming this knowledge deficit, we undertook the development of our Longan Information Service System (LISS) as an on-going project developing a knowledge retrieval model based on an ontology [2]. The information in the ontology consists of knowledge on pruning, fruit thinning, sodium chlorate treatment for stimulating fruit flowering, pest control and general crop fertilization, all of which are important in producing quality fruit. The project included the identification of the relationships between these quality longan production factors, and the presentation of a prototype of an ontology of this knowledge and the development of a quality knowledge search system based on the principles of semantic searching.

2 Longan Production

A quality longan can be defined as a longan fruit with a 3 cm. circumference and a golden yellow shell [3] (Figure 1).

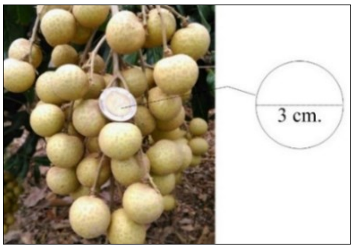

Fig. 1. A Quality Longan

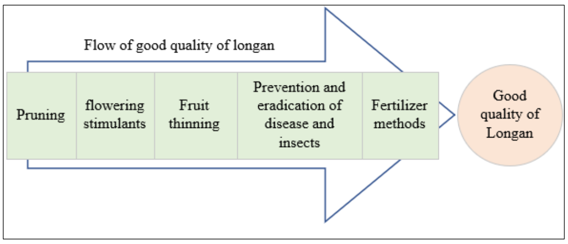

Fig. 2. Flow of good quality longan.

To achieve the production of high quality longan fruit, suitable for the international market and able to attract high prices, it is essential to follow a cropping and crop treatment process [2] that is illustrated in Figure 2. This process is comprised of five essential steps. First, pruning, cutting away some branches of the longan tree to allow fresh air and sunlight to reach all of the developing fruit. Pruning is also necessary to reduce the outbreak of diseases and to eradicate insect pests [4],[5],[6]. The second step in the process is to apply Secondly, the correct treatment is to apply appropriate amounts of potassium chlorate or sodium chlorate at appropriate times and intervals to stimulate flowering [3],[7],[8],[9],[10]. Fruit thinning is the next step when removing a certain amount of the developing fruit will assist in the development of quality of the remaining fruit with a larger circumference [11]. Pest control is a continuing step in the process to prevent an outbreak of pest insects, as well as limiting the spread of diseases. The amount and type of chemicals appropriate at different times in the growing cycle must be correct and properly controlled [12],[13],[13],[15]. Also, care must be taken not to apply chemicals of a type and at times that are detrimental to the beneficial insects such as bees that are essential in the pollination of the trees and during flowering[16] . Finally, an understanding of the types, amounts, and correct application methods and timings of the application of fertilizers is essential as nutrients of different types and formulations are necessary at different times in the growing period [7],[17],[18].

If the farmers should follow these processes and apply these techniques to produce get good quality fruit, it was essential to correctly and comprehensively gather the necessary information, store it in a manner that was easily and "intelligently" accessible by the farmers: thus an ontology based system with semantic searching.

3 Ontology & Semantic Web

Semantic search systems using RDF (Resource Description Framework) are systems for data collection that use SPARQL, short for "SPARQL Protocol and RDF Query Language", for searching the knowledge base. The SPARQL-based system that we developed for quality longan production information was developed using the following tools.

3.1 Semantic Web.

A semantic web is an information structure form which a computer program gains understanding of the meaning of the information concepts defined and related in the web, and of the processes of evaluating and associating the information concepts. The semantic web also includes the technology used to store and present structured content.

The ability to understand the data or information semantics, and apply this understanding to data searches helps non-experts in that field to search for information contained in the web in a manner that they would otherwise be unable to carry out, given their restricted understanding of the data. The computer can automatically understand of the meaning of web content and processing [19].

As stated by the W3C, the international organization that sets the development guidelines and standards for information on Semantic Web, referenced discussed [20], the architecture of a Semantic Web must be based on multiple languages and logical constructs, which we suggest includes URI / IRI and Unicode, XML Language, Namespaces, XML Schema, XML Query, RDF Model and Syntax, Ontology, Rules, Query Language, Logic. This multi-tiered structure is illustrated in Figure 3, taken from [21].

3.2 Web Ontology Language (OWL)

The World Wide Web Consortium (W3C) Web Ontology Language (OWL) is a Semantic Web language designed to represent rich and complex knowledge about things, groups of things, and relations between things. for describing ontologies and defining relationships between data by processing the information content as well as presenting information to users in a meaningful way. The data is categorized in the form of classes, class properties and relationships between classes, allowing the computer to read values, comprehend the associations and relationship between and among the classes, and understand the semantics of the data [22].

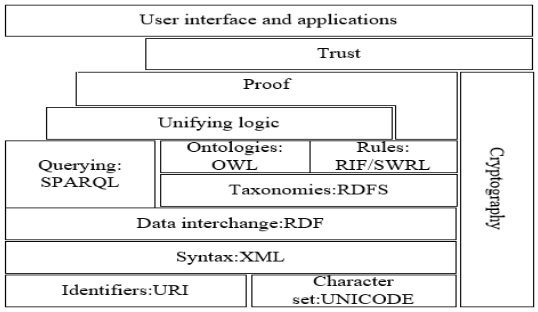

Fig. 3. Semantic web architecture in layers [21].

3.3 Ontology

An ontology is essentially database technology that is used to store knowledge on a particular subject domain and describe the semantics inherent in or defined for the data, extending and increasing the understanding of the scope of the stored knowledge and the relationships between the knowledge constructs [22] by using the descriptions and specifications of the knowledge concepts to illustrate the conceptualizations inherent in or described for the objects that exist in that domain [23],[24]. As suggested in [25], the ontology contains descriptions of the knowledge base structures for the purpose of information sharing by interested participants in that knowledge domain, and the between those involved with the domain and can be used to process applications in various applications to have the expertise and increase the automation of the process more[25][26]. The significant and essential factor in the development of an ontology is the involvement of knowledge engineers, together with specialized experts in the knowledge domain, to build the ontology and the associated computer and information systems [27].

3.4 Ontology Application Management Framework (OAM)

An Ontology Application Management Framework (OAM) is an application framework that simplifies the development of applications that use semantic web technologies and access the ontology. The user can import existing data from various information systems currently stored in various data formats (XML, .CSVs, RDBMS tables etc.), or from other domain ontologies, and define recommendation rules, together with configuration and access information. This can be achieved using an application template developed by the National Electronics and Computer Technology Center (NECTEC) in Thailand. The OAM Framework is integrated with software for processing the data according to the semantic web standards currently available under the defined application framework [27]. These software tools include:

1. Apache Jena: a Java Framework for semantic web application development.
2. D2RQ: a software tool that contains language for defining data conversion standards between relational database schema and data in the form of an ontology according

to OWL and RDF Schema standards. D2RQ was developed by the University of Berlin, Germany.

3. Apache Jena TDB: RDF database management software that supports storing and retrieving of RDF data using the SPARQL language via the Jena API and Jena Fuseki (part of Apache Jena).
4. Apache Jena Inference Engine: a software tool for processing based on a Rule-based Inference Engine used for RDF data (part of Apache Jena).

The interaction and structural relationship between these tools in illustrated in Fig 4.

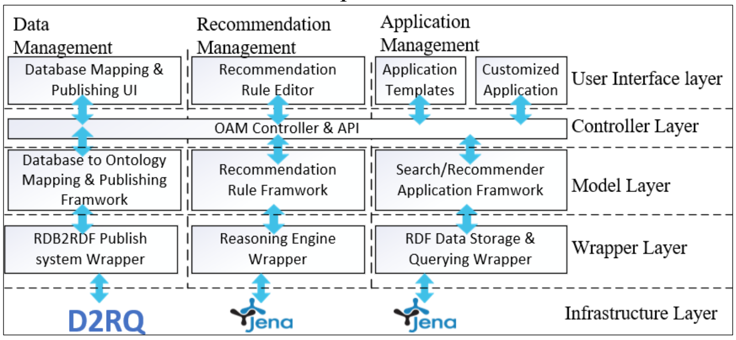

Fig. 4. An overview of platform components Supporting the development of knowledge-based applications[27].

4 Semantic Search on Longan Information Service System (LISS).

This paper proposes a mechanism of the quality longan production information system, especially the important factors in the production of longan for quality by dividing into 3 parts: Knowledge Acquisition, Knowledge-based Ontology, and Ontology-based Application Development as shown in the Figure 5.

4.1 Knowledge Acquisition

The adoption of semantic web technologies is dependent upon the availability of existing semantic resources. The collection of knowledge is an important issue. For our research and ontology-development purposes, we conducted an extensive literature search and selected knowledge from the research articles that we found as well as from textbooks on the subject. We also conducted in-depth interviews with experts regarding matters related to quality longan production. We identified the appropriate techniques that included fruit thinning, pruning, prevention and eradication of diseases and insects, use of flowering stimulants, and fertilization methods, timing and products [2]. Using this information we created a simple interface that helps convert the knowledge gained from experts into easily accessible ontological structures [28]. This information acquisition and recording process is illustrated in Figure 5, Section A: Knowledge Acquisition.

After initial analysis of the information and the inter-relationships between and amongst the information, we found that the production period was a variable that linked all 5 techniques, allowing us to divide the production period of longan into 7 phases: Preparation, Before flowering, Flowering, Fruiting, Early fruiting, Medium-term fruiting and Harvest.

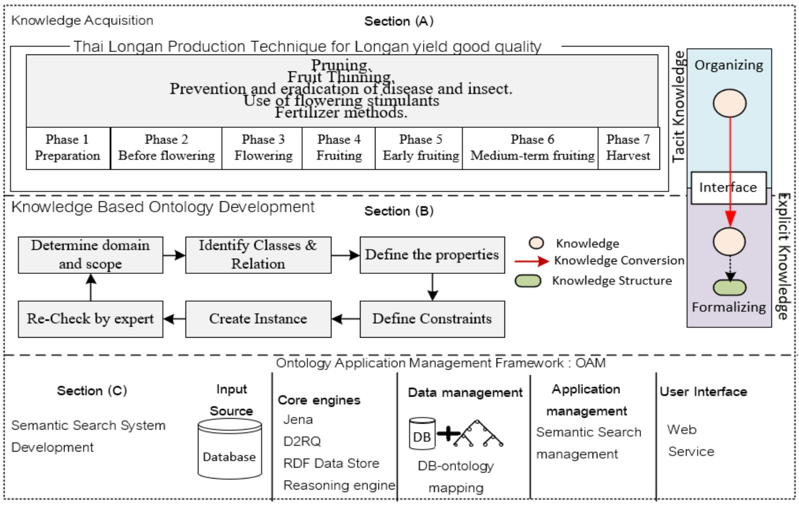

Fig. 5. An overview of the research framework.

4.2 Knowledge-based Ontology development

In pursuit of our purpose of developing a knowledge base and an ontology to inform growers on producing quality logan fruit, existing knowledge needed to be comprehensively identified and gathered [2]. The properties of various information classes needed to be identified and defined, and transformed into appropriate structures in the ontology.

The steps involved in developing this knowledge base were as follows: 1) Determine the domain and the scope of the domain, or what is called the Universe of Discourse elsewhere, 2) Identify the Classes and Relationships between and within the Classes 3) Define the properties of the Classes 4) Define Constraints 5) Create Instances, and 6) Recheck: subject these information structures and the information content to be reexamined by an expert as an information quality control measure as illustrated in Figure 5, Section B: Knowledge Based Ontology Development.

There are many tools available for the development of an ontology knowledge base, such as the Protégé program developed by Stanford University United States [29] and the Hozo-Ontology Editor Program [30] developed by Osaka University, Japan. Both of them are opened-source software packages that conform to the OWL Semantic Web Standards of the World Wide Web Consortium (W3C) for ontology development. Using the knowledge base and ontology of Thai longan production presented in [2], the Hozo-Ontology editor was the implementation tool used for developing our ontology because it particularly applies to the domain of agriculture and works with the Ontology Application Management Framework (OAM) [27],[31],[32].

4.3 Development of knowledge searching systems

The first step in developing the system with OAM is Converting source data from data from the original user's database to RDF format: the Data Management Component. This works at the model level and consists of Schema mapping according to the OWL standard, dividing the specification of Properties of Classes into two types: the Datatype Property is a property that has a relationship type of Attribute-of (a/o) other Classes, with datatype such as string or integer, and Object Property that has a relationship type of Part-of (p/o) with other Classes. This Schema Mapping also includes Vocabulary mapping based on external ontologies by matching data from a source database with synonymous Classes. The outcomes of this Schema Mapping process are generated in the RDF (Resource Description Framework), as shown in Figure 6.

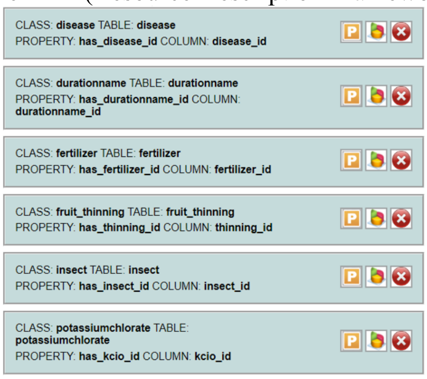

Fig. 6. Program configuration

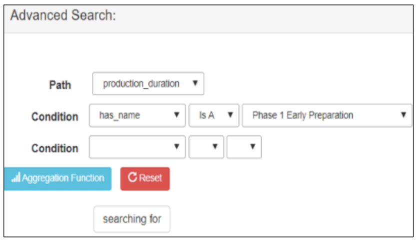

Fig. 7. The user interfaces for searching

The second step in developing the system with OAM is Application Management in which search patterns and search result formats are defined. The design approach is divided into a layer of work to allow the system to be flexible and easy to modify. Three layers are defined that have different functions.

First, there is the Model layer that is the part of the data structure that is independent of the software that is in the infrastructure. The layer that will be supported makes it more flexible. Which must develop the data conversion between the model and each software in the Wrapper layer.

The Controller layer links the data and functions in the model of each part of the system to be compatible, including the design of the OAM API so that the parts of the program, tools and applications can be connected so that users can manage Information within the system.

UI layer is a section developed to allow users to manage configuration information and use data within the system via the Application Template or customized application which will use the management functions via the OAM API as shown in Figure 4.

In the process of developing our Longan Information Service System (LISS) we accessed the information using the production period as the main parameter. The main factors affecting the quality of longan products are all related to the production period. These main are shown in Figure 7. This process will be part of the LISS system as illustrated in Figure 5, Section C: Semantic Search System Development.

5 Experiments & Results

Using the production period as the main condition for accessing the system and displaying the information based on the production factors, we were able to evaluate the efficiency and accuracy of queries by measuring Precision (1), Recall (2) and F-measure (3) [33]using the following equations:

$$\text{Precision P}=\frac{|A|}{|B|} \tag{1}$$

by |A| is the number of correct answers. |B| is the total number of system responses.

$$\text{Recall R}=\frac{|A|}{|C|} \tag{2}$$

by |C| is the total number of answers that the expert identified.

$$\text{F-measure F}=\frac{2x(RxP)}{(R+P)} \tag{3}$$

5.1 Search Performance Tests

Search Performance Tests were carried out to evaluate responses to queries regarding the variety of information of interest, including recommendations regarding (a) Fertilizer formulae, (b) Pruning, (c) Use of flowering stimulants, (d) Fruit Thinning, and (e) disease and insect infestation control.

Based on the production phase indicated, queries were presented and the correctness of the answers was evaluated by 3 experts rating the answers as correct or incorrect. Where a majority, meaning at least 2 of the 3 experts, agreed, then the assessment of the majority as to correct or incorrect prevailed.

The results achieved were Precision, 100% and Recall, 93.50%. Analysis of the responses showed that each Rule defined in the ontology had a Precision of 100%, with the information presented to satisfy queries included 3 methods of Fruit Thinning, 3 methods of Pruning, 2 methods of flowering stimulant use, 13 diseases, 23 insect pests and 14 fertilizer formulae. However, the experts had suggested, in the early stages of knowledge acquisition, that there were more fertilizer formulations than the system suggested, and in the fruiting period there were more diseases and pests suggested by the experts than were recalled by the system. Because the database does not cover all disease and fertilizer information which can be updated later not affect the knowledge base structure. consequently, the developed system can be used for real use.

6 Conclusion

We developed an ontology of knowledge regarding the cropping and production of longan fruit for the purpose of enhancing fruit quality and growing quality longan fruit suitable for the international market. Seven main classes of production information were defined: disease control, insect eradication, fertilizer formulations, pruning requirements, chemical assistance to flowering, fruit thinning and production duration. This enabled the creation of production calendars as guidelines for farmers to improve the quality of their crops. Following these guidelines resulted in lower costs of production and higher quality fruit, resulting in higher profits.

References

1. Office of Agricultural Economic. Agricultural statistics of thailand. 2018.
2. Bangkhomned W, Payakpate J. A conceptual design of thai longan production via an ontology. International Journal of Latest Trends in Engineering and Technology Special Issue ICRMR-2018. 2018;12(6):018-23.
3. Manochai P, Khaosumain Y, Sritontip C, Changjaraja S. Technology on longan production, Bangkok: physicscenter; 2004.
4. Manochai P. The control of longan canopy size by pruning, Bangkok: The Thailand Research Fund; 2007.
5. Thorp TG, Stowell B. Pruning height and selective limb removal affect yield of large 'hass' avocado trees, HortScience. 2001;36(4):699-702.
6. Hudson RL. Sunset pruning handbook. Menlo Park, California, Lane Books; 1971.
7. Yuttana Khaosumain, Chiti Sritontip, Santi Changjaraja. Effect of kcio3 (potassium chlorate) on off-season flowering in longan (dimocarpus longan lour.) cv. Daw. bangkok: Kasetsart University Research and Development Institute; 1999.
8. Manochai P, Sruamsiri P, Wiriya-alongkorn W, Naphrom D, Hegele M, Bangerth F. Year around offseason flower induction in longan (dimocarpus longan, lour.) trees by kclo3 applications: Potentials and problems. Scientia Horticulturae. 2005;104(4):379-90.
9. Kanaree V, Pankasemsuk T. Effect of potassium chlorate on quality of longan fruit. Journal of Agriculture. 2003;1.
10. Potchanasin P, Sringarm K, Naphrom D, Bangerth KF. Floral induction in longan (dimocarpus longan, lour.) trees: Iv. The essentiality of mature leaves for potassium chlorate induced floral induction and associated hormonal changes. Scientia Horticulturae. 2009;122(2):312
11. Jarassamrit N, Manochai P, Jarenkit T, Suton V, Wiriyaalonggon W. Effects of fruit thinning on longan quality. The 2nd National Horticultural Congress; Charoen Thani Hotel, Khon Kaen2002.
12. Visitpanich J. Determinations of the causal agents of decline, witches' broom symptom and sudden death of longan and their control. Bangkok: The Thailand Research Fund(TRF). 2003.
13. Visitpanich J. Pest management of economic important diseases and insect pests of longan and developing diagnostic tool for producing pathogen free longan plant. Bongkok: The Thailand Research Fund:TRF.(in Thai); 2000.
14. Babpraserth C, Klinkong S, Chiemsombat P, Paradornuwat A. Development of disease detection techniques for longan witches' broom and disease epidemiology. Bangkok: Kasetsart University Research and Development Institute.(in Thai); 1994.

15. Anupunt P, Sukhvibul N, editors. Lychee and longan production in thailand2005: International Society for Horticultural Science (ISHS), Leuven, Belgium.
16. Narjes ME, Lippert C. Longan fruit farmers' demand for policies aimed at conserving native pollinating bees in northern thailand. Ecosystem Services. 2016;18:58-67.
17. Khaosumain Y, Sritontip C, Changjeraja S, editors. Effects of nitrogen concentration on growth and leaf nutrient content of longan in lysimeter2008: International Society for Horticultural Science (ISHS), Leuven, Belgium.
18. Yuttana Khaosumain, Chiti Sritontip, Santi Changjaraja, Pawin Manochai, Theeranuch Jarenkit. On-farm trial for fertilizers management to enhance off-season longan production. Songklanakarin Journal of Plant Science. 2016;3(1(January-March)):63-9.
19. Berners-Lee T, Hendler J, Lassila a. The semantic web: Scientific American; 2001.
20. Kifer M, Bruijn Jd, Boley H, Fensel D. A realistic architecture for the semantic web. Proceedings of the First international conference on Rules and Rule Markup Languages for the Semantic Web; Galway, Ireland. 2179200: Springer-Verlag; 2005. p. 17-29.
21. Harth A, Janik M, Staab S. Semantic web architecture. In: Domingue J, Fensel D, Hendler JA, editors. Handbook of semantic web technologies. Berlin, Heidelberg: Springer Berlin Heidelberg; 2011. p. 43-75.
22. Uschold M, Gruninger M. Ontologies: Principles, methods and applications. The Knowledge Engineering Review. 1996;11(2):93-136.
23. Gruber TR. A translation approach to portable ontology specifications. Knowledge Acquisition. 1993;5(2):199-220.
24. Noy NF, McGuinness DL. Ontology development 101: A guide to creating your first ontology. Stanford Knowledge Systems. 2001.
25. Lee C-S, Kao Y-F, Kuo Y-H, Wang M-H. Automated ontology construction for unstructured text documents. Data & Knowledge Engineering. 2007;60(3):547-66.
26. Srisawat, C., & Payakpate, J. (2016). Comparison of MCDM methods for intercrop selection in rubber plantations. Journal of Information and Communication Technology, 15(1), 165-182.
27. Buranarach M, Supnithi T, Thein YM, Ruangrajitpakorn T, Rattanasawad T, Wongpatikaseree K, et al. Oam: An ontology application management framework for simplifying ontology-based semantic web application development. International Journal of Software Engineering and Knowledge Engineering. 2016;26(01):115-45.
28. Di Iorio A, Rossi D. Capturing and managing knowledge using social software and semantic web technologies. Information Sciences. 2018;432:1-21.
29. Musen MA, Protégé T. The protégé project: A look back and a look forward. AI matters. 2015;1(4):4-12.
30. Mizoguchi R, Sunagawa E, Kozaki K, Kitamura Y. The model of roles within an ontology development tool: Hozo. Appl Ontol. 2007;2(2):159-79.
31. Bangkhomned W, Payakpate J. The development of ontologies of semantic search technology for disease management in longan. Phranakhon Rajabhat Research Journal(Science and Technology). 2019;14(2).
32. Chariyamakarn W, Boonbrahm P, Boonbrahm S, Ruangrajitpakorn T, editors. An expert system using ontology as knowledge base for personalized rice cultivation suggestion2018; Cham: Springer International Publishing.
33. Sirisom, P., Payakpate, J., & Wongthai, W. (2017) A system design for the measurement and evaluation of the communications security domain in ISO 27001:2013 using an ontology. Vol. 424 (pp. 257-265).