Journal of Computing, Science &Technology
https://focjournal.unidel.edu.ng/
editorjcst@unidel.edu.ng
ISSN: 3043-6818 Print
https://focjournal.unidel.edu.ng

Building Ontology for Medicinal Fruits.

Iyamah B. E.[1], Okpako E. A.[2]
Department of Software Engineering, University of Delta, Agbor, Delta State.[1], Department of I.C.T University of Delta, Agbor, Delta State.[2]
boyce.iyamah@unidel.edu.ng[1], ejaita.okpaka@unidel.edu.ng
Corresponding Author's Email: boyce.iyamah@unidel.edu.ng

Article Info

Date Received: 14-09-2024
Date Accepted: 25-11-2024

ABSTRACT

This work suggests a way to use ontology approaches to build a knowledge base about medicinal fruits and create its application as a semantic search system. Fruit prescription has often gained popularity due to an increased demand for natural medicinal supplements found in fruit. Since information about medicinal fruits which is used to treat human ailments is frequently provided in a flat manner, often scattered and without meaning and lacks taxonomy, it is difficult to have a comprehensive understanding of illnesses and medicinal fruit treatment. In order to solve this issue, the explicit state of medicinal fruit knowledge was created through gathering data about fruit domain and the deployment of ontology engineering techniques in the health domain. In domain knowledge, ontology techniques aid in identifying important ideas or concepts and the connections that binds them. Concepts derived from medicinal fruits and the diseases they can treat are covered by the resulting application, Ontology for Medicinal Fruits (ONMF). The use of Protégé as an ontology tool was employed to model the knowledge base and OntoQA to evaluate its effectiveness and alignment with the system. As a semantic search ONMF application was found to improve the efficiency of information query by excluding non-relevant information.

1.0 INTRODUCTION

Damage to the human body occurs when the cells become sick and unstable, affecting the tissues, organs, and systems as well. Atoms or molecules containing one or more electrons not coupled in an orbital are called radicals which are mostly responsible for most human illnesses. Positive, negative, or neutral charges are all possible for radicals. It is common to define a radical as a species with an odd number of electrons while a Free Radical is a radical which moves from its original vicinity where it has been created. [1]
Actually, the World Health Organization advises those over 35 to consume at least five different fruits each day and vegetables daily to provide the body with enough nutrients.[2]
Vitamins and minerals are abundant in fruits. For instance, "beta carotene," which is often the precursor of vitamin A, is a significant component in mango and pawpaw. According to the findings of current studies, a large number of phytochemicals included in fruits function as potent antioxidants, preventing the harm that diseases inflict to the body's cells and organs. They are the biologically active ingredients that give plants their color, flavor, and odor as well as their defense against diseases that afflict both humans and plants. As a result, the potential of hundreds of these plant compounds to prevent cancer and other degenerative diseases is being studied. Bioflavonoid (vitamin P), phenolics, lycopene, carotenoids, antioxidant vitamins C and E, and glucosinolates are a few of the intriguing phytochemicals that have antioxidant properties.
Generally speaking, everyone would favor an efficient decision support system given the capabilities of computers, such as their speed, persistence, vast storage capacity, and communication capabilities, as well as the advantages of information and communication technology (ICT). Nowadays, electronic information is accessed, retrieved, and syntactically arranged [5]. The issue with this plan is that a significant amount of man-hour is still needed to obtain the pertinent resources, and the majority of the search results are useless. Therefore, it seems that a system of manual storage, access, and retrieval that is well-organized might not be any less effective. Consequently, it is clear that this electronic method simply solves the space issue; the main issue of ineffective access and retrieval of pertinent information remains unaddressed [6] [7]. We therefore need a scheme though electronic that will guarantee or ensure high relevance of fruits (information) retrieved.

1.1.1 CONCEPT OF ONTOLOGY

Because ontologies contain information—that is, data and its meaning within a domain—they are both machine and human comprehensible, unlike databases, which store data without meaning. In this era of information overload and knowledge management for better information processing, interpretation, and application, ontology is particularly advantageous since it makes it easier for humans and machines (computers) to join forces in the processing and application of information as knowledge.

Philosophy defines "ontology" as "the science of what is, of the kinds and structures of objects, properties, events, processes, and relations in every area of reality" [8], a definition that was appropriated for use in computer science and artificial intelligence (AI) in particular. The most popular definition of ontology in AI was provided by [9], who characterized it as a "explicit specification of a conceptualization." [10]. According to the work of [11], ontology can take many various forms, but it always consists of a vocabulary of concepts and an explanation of what they mean.

Together, these definitions and an example of the connections between ideas provide the field with structure and restrict the number of meanings that terms can have. Ontology is defined as "a formal explicit description of concepts in a domain of discourse (classes), properties of each concept describing various features and attributes of the concept (slots), and restrictions on the slots (facets)" in [12], which this paper followed because it seems clearer and more useful [13]. Semantic retrieval systems are essentially composed of three fundamental components, as shown in Figure 1.1: search engine, ontology, and corpus [14]. Semantic retrieval systems, even within the same area, are clearly distinguished by their ontologies, which are also the foundation of the semantic web [15]. Stated differently, an American semantic retrieval system can only be applied in Nigeria or Ghana if the ontology is specially modified and constructed for those countries' legal systems.

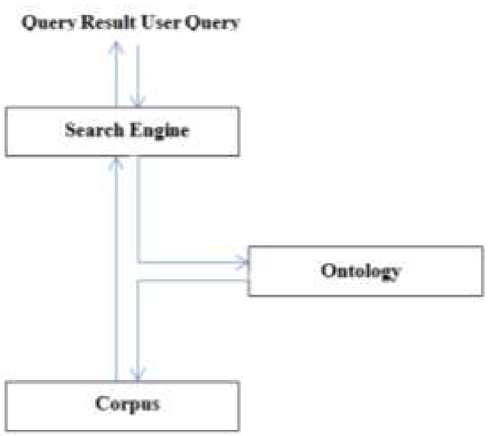


Fig 1.1 : Schematic View of a Semantic Retrieval System [16]

1.1.2 CLASSIFICATION OF ONTOLOGY

In the literature, there are several different ontologies [16]. The typical way of classifying ontologies is by commitment or goal, however there are other approaches as well [17]. The scope of the things the ontology defines, formality or complexity, and other factors are examples of additional classes. In [18], three types of ontological commitments are described: task commitment, method commitment, and domain commitment.

1.1.3 BUILDING ONTOLOGIES

There are several reasons for building ontology. Some of which are as follows [12] and [19]:

i. Share common understanding of the structure of information among people or software agents.
ii. Enable reuse of domain knowledge.
iii. Make domain assumptions explicit.
iv. Separate domain knowledge from the operational knowledge.
v. Analyze domain knowledge.

There are several established methods for building and evaluating ontologies at the moment. Since the authors [12] made it clear that there is no one correct ontology design technique, the ontology designer is free to select the one they want to utilize. These methods include the KATUS and dentistry methodology, the Lenat and Guha methodology, the Noy and McGuiness methodology, the Gruninger and Fox methodology, and others [12], [13], and [20–24]. All of these techniques can be used to the three basic ontology rendering methods: top-down, middle-out, and bottom-up. This paper employed the top-down approach and the methodology of Noy and McGuinness for the sake of simplicity and reliability. [25], [13].

1.1.4 ONTOLOGY LANGUAGES

In contemporary information systems, ontologies are state-of-the-art instruments for representing and organizing knowledge. There are several ways to represent ontologies, which fall into three categories: web-based, web standards, and traditional. For creating ontologies, Web Ontology Language (OWL), which is supported by the World Wide Web Consortium (W3C), is the most popular of them. With different expressiveness levels, OWL contains three sub-languages: OWL Lite, OWL DL, and OWL Full. The best option for this study is OWL DL since it guarantees ontological consistency [14] [17] [18]

1.1.5 BACKGROUND OF MEDICINAL FRUITS

Although most of us have a good idea what fruits and vegetables are when we eat them, it would be difficult to provide a definition for someone of just what makes one food a vegetable and another a fruit. For a botanist, the definitions are easier. [19]

"A fruit is a reproductive structure of an angiosperm which develops from the ovary and accessory tissue, which surrounds and protects the seed".

Fruit categorization and structure are discussed in this section. In botany, the definition of a fruit is simple. Fruit and seed development are both started by the fertilization process. Fruit develops as a result of a complicated series of changes in the ovary tissue, while seeds are formed from the ovules. A lot of fruits have flesh and sweets, which draw animals that spread the seeds to other places. Alternative seed dispersal strategies are employed by various non-fleshy fruits. Certain plants have the ability to produce fruit without fertilization. [22] [20]. These fruits, known as parthenocarpy, have no seeds. In order to avoid seeds, some of the grapes and watermelons that you purchase from farmers are grown in this manner. Its wall frequently thickens and differentiates into three roughly separate layers as the ovary grows into a fruit. Three layers come together to create the pericarp, which envelops the developing seed or seeds, according to [24]. The three fruit layers are:

i. Exocarp—the outermost layer often consisting of only the epidermis.
ii. Mesocarp—or middle layer, which varies in thickness
iii. Endocarp—which shows considerable variation from one species to another, is the inner-most layer of the fruit

1.1.6 FRUIT CLASSIFICATION

All fruits may be classified into three major groups on the basis of the number of ovaries and the number of flowers involved in their formation. The following outline includes most of the common types of fruits. [26]

1. SIMPLE FRUITS—In a single flower, a single matured ovary gives rise to simple fruits. The ovary is joined to a different floral component in accessory fruits. Fleshy fruits and dry fruits are the two main categories of simple fruits..
2. AGGREGATE FRUITS— comprise several fully developed ovaries scattered throughout the outer layer of a single container that developed in a single flower. Fruitlets are each of the aggregate fruit's separate ovaries. Every fruitlet will have a stony hole inside of it. In actuality, an aggregation fruit is made up of numerous little drupes.
3. MULTIPLE FRUITS— comprise a mass that is essentially made up of the mature ovaries of several lots of flowers. Almost always, many fruits are accessory fruits.

2.0 METHODOLOGY

Protégé, Ontoedit, Chimaera, Ontolingua, WebODE, OilED, pOWL, SWOOP, and OntoGen are a few of the editors available for rendering ontology [27–32]. The reason Protégé was selected from among these editors is [13]:

i. A knowledge base framework and ontology editor that are open source.
ii. Bolstered by a robust developer community that includes everyone from government officials and academics to business users seeking knowledge solutions in fields like corporate modeling, biomedicine, and intelligence gathering.
iii. Built on the Java platform, offering a plug-and-play environment that makes it an adaptive foundation for quick application development and prototyping.
iv. Suitable with Web Ontology Language (OWL), Extensible Mark-up Language (XML), and Resource Description Framework (RDF) and is platform independent. In particular, the ontology was modeled using Protégé OWL for expressivity and consistency purposes. Since OWL (RDF/XML) is widely used for rendering ontologies, our codes are in this format [24].

A detailed table comprising of the unique features could be seen as follows

Feature/Criteria	Ontology Language Support	User Interface	Semantic Web Support	Ease of Use	Open Source	Unique Advantages
Protégé	OWL, RDF(S), XML, Frames	User-friendly graphical interface	Fully supports semantic web technologies	High; well-suited for beginners and experts	Yes	Plugin ecosystem, active community, and support for instance data
OntoEdit	RDF(S), DAML+OIL	Text-based; complex	Limited	Moderate; suited for advanced users	No	Focuses on RDF(S) integration
Chimera	OWL, RDF(S)	Minimal, basic interface	Basic	Simple; beginner-friendly	Yes	Lightweight and simple
Ontolingua	KIF, Ontolingua	Text-based interface	Limited	Steep learning curve	Yes	Early framework for ontology sharing
WebODE	OWL, RDF(S)	Intuitive but less modern	Good support	Moderate	Yes	Supports ontology evaluation
OILed	DAML+OIL, OIL	Basic	Limited	Moderate	Yes	Lightweight editor
pOWL	OWL	User-friendly	Strong integration	High	Yes	Web-based with PHP integration
Swoop	OWL	Lightweight but developer-focused	Good support	High	Yes	OWL-centric, debugging tools
OntoGen	OWL	Semi-intuitive	Moderate	Moderate	Yes	Semi-automated ontology generation

A series of competency questions was used to assess the ontology, which was constructed using the technique of Noy and McGuinness [12]. Competency questions define the ontology task and are typically considered at the beginning of ontology building, much like functional software requirements.

The steps involved in the Noy and McGuiness' methodology are as follows:

Step I: Establish the ontology's domain and extent
Step II: Take consideration for recycling pre-existing ontologies
Step III: List the key terms in the ontology.
Step IV: Establish your framework and classes
Step V: Describe the classes' attributes
Step VI: Specify the slots' values
Step VII: Develop examples

We were able to maintain concentrate on the ontology's goal thanks to the competency questions that were developed, and as a result, they were utilized to guarantee that the ontology was properly designed and evaluated. The competency questions that have been developed are:

i. What are the classes of fruits we have?
ii. What are the types of fruits we have under each class of fruit?
iii. What fruits cure what disease?
iv. What is the nutritional content of each fruit?

Furthermore, the high-level design of the ontology was modeled using a Directed Acyclic Graph (DAG) as shown in Figure 2.

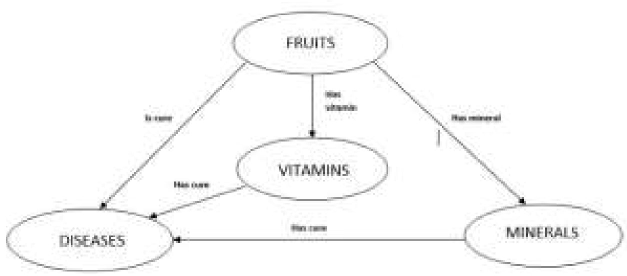
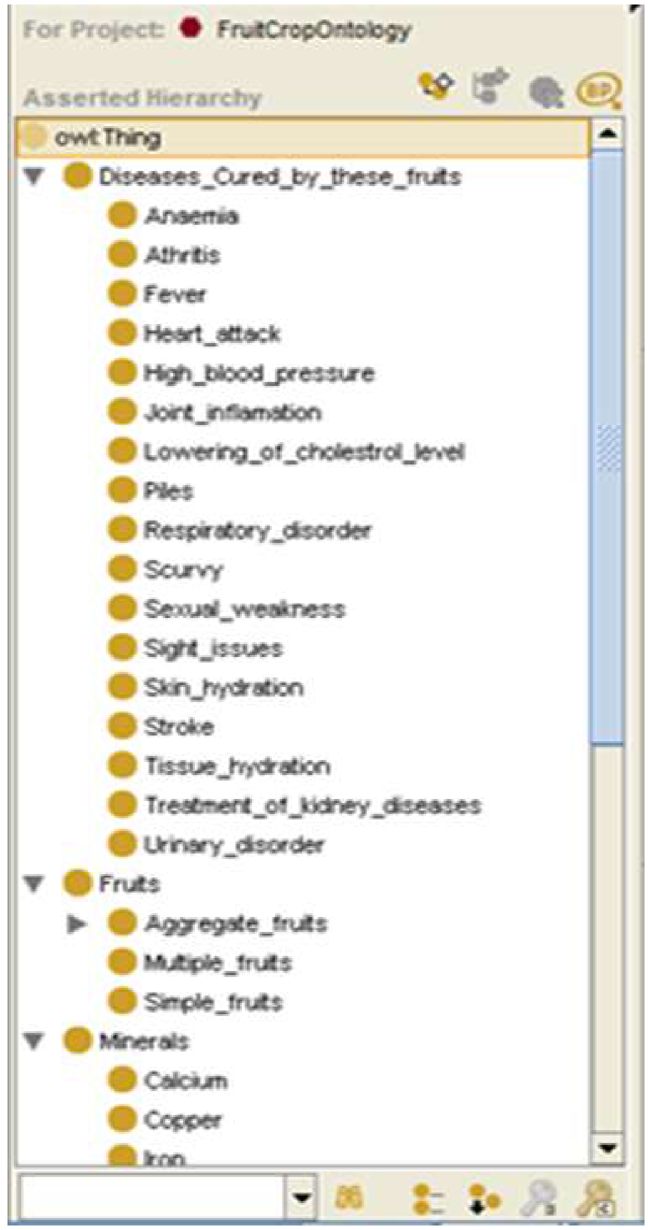
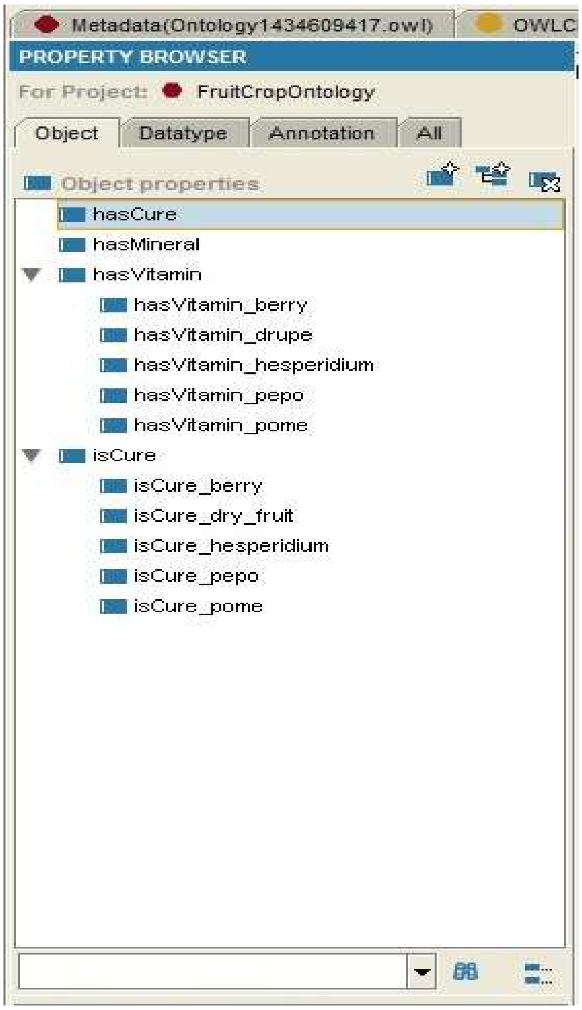


Figure 2. depicts the ONFC high level concepts and their relationships. The concepts are Fruits, Diseases cured, Minerals and Vitamins and the relations are IsCure, HasVitamin, HasMineral, and HasCure.

The concepts and the properties of ONMF as implemented in Protégé 5.5 are as shown in Figure 3a and 3b respectively. Figure 4 shows the Domain/Range of the ONMF and the class/individual tree of the ONMF is as shown in Figure 5.





In Protégé, an owl ontology begins with a root that is typically represented by "owl: Thing." As shown in Figure 3a, the superclass "owl:Thing" incorporates the concepts "Fruits, Minerals, and DiseasesCureByTheseFruits," each of which has its own subclasses. It is evident from Figure 3b, which displays the properties linking the ON MF ontology's concepts, that the attribute "is State" is the opposite of has State.

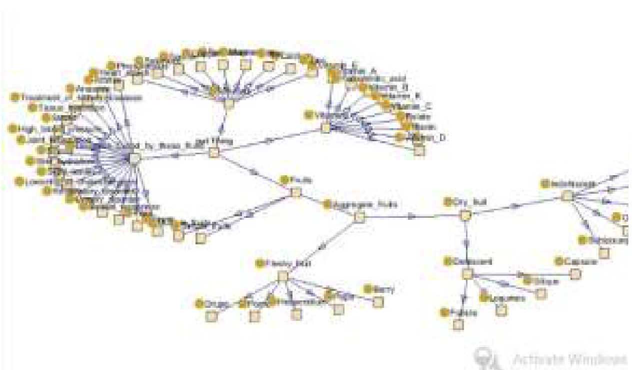


Fig. 4. ONMF Domain/Range

Since ontologies are usually expressed as triples (i.e. subject, predicate, object), the Domain/Range is typically utilized to demonstrate which concept is the subject and which is the object at a given instance. Range is the object, properties are the predicate, and the domain is the subject. Based on Figure 4,

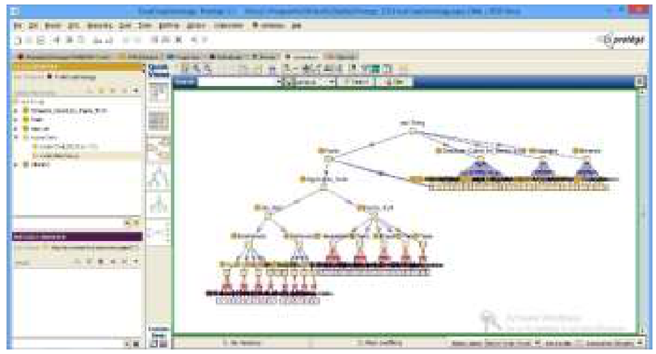


Figure 5 is a visualization of class and individual tree defined in the ONMF. This tree simply shows the hierarchical relationships among the concepts of the ontology and the instances of the concepts.

3.0 RESULT

The complexities of the ontology constructed (such as concepts, properties, domain, range, class, and instance) are revealed in section 3, which constitutes one of the standard approaches to presenting ontology results [13, 14, 34]. In order to capture the ontological commitment, an ensemble of competency questions is now frequently used to query the ontology in addition to this method [12, 17, 33]. In addition to being comprehensible to humans, particularly non-experts, this novel method of presenting ontology results continues to gain popularity since it illustrates the legitimacy of the ontology in relation to its goal.

As with similar researches [17, 37], we queried ONMF in line with the competency questions as shown in section 3 using Protégé tool query and export tab plug-in. A few of these queries and their results were captured in Figure 6a, Figure 6b and Figure 6c.

QUERY 1 - What are the types of fruits we have under each class of fruit?

also have trouble comprehending and exchanging knowledge from other users who utilize different terminology and semantics. Building a system that can satisfy the needs of every user is tricky, particularly when ambiguity is involved. In order to express knowledge, RDF and Ontology are suggested tools that enable better solutions to speed up and streamline search techniques for both computers and people. Thus, ONMF was constructed. To do this, we used the methods established by Noy and Mcguiness. Using Directed Acyclic Graphs (DAG) as a model, ONMF was constructed with the help of the Web Ontology Language (OWL) editor, Protégé 5.5. Using OntoQA, we further demonstrated that ONMF is highly compatible with its goal of effective and efficient storage, access and retrieval of Nigerian medicinal fruit crops. Among the limitations of the ONMF, some key challenges include the ambiguity in data representation, scalability issues, limited generalizability and dependency on domain expertise. Some key areas for future enhancement of the system include multi ontology support, feedback loops and scalability to accommodate more data. ONMF offers unique contributions that distinguish it from other methods. Unlike traditional ontology engineering techniques which groups items purely based on mathematical patterns, ONMF incorporates domain specific ontologies to guide factorization, ensuring outputs are semantically meaningful. Furthermore, compared to machine learning models, ONMF provides enhanced semantic context as it aligns results with domain knowledge instead of relying solely on statistical patterns. Unlike supervised models, ONMF does not require extensive labeled datasets making it more practical for domains with limited data. The system would be beneficial to researchers, healthcare professionals, educators and agricultural planners.

6.0 REFERENCES

[1] Setchi, R. M. and Tang, Q. (2019). Concept Indexing using Ontology and Supervised Machine Learning. International Journal of Computer, Electrical, Automation, Control and Information Engineering. Vol. 1, No. 1, pp. 89-94.

[2] Maedche, A. and Staab, S. (2018). Measuring similarity between ontologies. In Proceedings of the 13th International Conference on Knowledge Engineering and Knowledge Management: Ontologies and the Semantic Web, EKAW02, Springer Verlag, LNAI. Berlin, 2473:pp. 251-263.

[3] Smith, B. (2023). Ontology. Blackwell Guide to the Philosophy of computing and information, Oxford, Blackwell, pp 1-19.

[4] Gruber, T. R. (2003). A Translation Approach to Portable Ontology Specifications. Stanford University, Stanford, California, pp. 1-23.

[5] Studer, R., Benjamins, V. R. and Fensel, D. (2018). Knowledge Engineering: Principles and Methods. Data Knowledge Engineering, Vol. 25, No. 1, pp. 161-197.

[6] Uschold, M. and Jasper, R. (2019). A Framework for Understanding and Classifying Ontology Applications. The Proceedings of the IJCAI-99 Workshop on Ontologies and Problem-Solving Methods (KRR5), Stockholm, Sweden, 12pp.

[7] Noy, N. F. and McGuinness, D. L. (2021). Ontology Development 101: A guide to Creating Your First Ontology. Stanford University, Stanford, CA, USA, pp. 1-25.

[8] Agapito, G., Simeoni, M., Calabrese, B., Caré, I., Lamprinoudi, T., Guzzi, P. H., Pujia, A., Fuiano, G., & Cannataro, M. (2018). DIETOS: A dietary recommender system for chronic diseases monitoring and management. *Computer Methods and Programs in Biomedicine*, *153*, 93–104. https://doi.org/10.1016/j.cmpb.2017.10.014

[9] Collins, P. F., Yang, I. A., Chang, Y.-C., & Vaughan, A. (2019). Nutritional support in chronic obstructive pulmonary disease (COPD): an evidence update. Journal of Thoracic Disease, 11(S17), S2230–S2237. https://doi.org/10.21037/jtd.2019.10.41

[10] Costa, D. L., Collins, M. L., Perl, S. J., Albrethsen, M. J., Silowash, G. J., & Spooner, D. L. (2014). An ontology for insider threat indicators development and applications. CEUR Workshop Proceedings, 1304, 48–53.

[11] Datta, S. K., Shaikh, M. A., Srihari, S. N., & Gao, M. (2021). Soft-Attention Improves Skin Cancer Classification Performance. http://arxiv.org/abs/2105.03358

[12] Filippi, P., Jones, E. J., Wimalathunge, N. S., Somarathna, P. D. S. N., Pozza, L. E., Ugbaje, S. U., Jephcott, T. G., Paterson, S. E., Whelan, B. M., & Bishop, T. F. A. (2019). An approach to forecast grain crop yield using multi-layered, multi-farm data sets and machine learning. Precision Agriculture, 20(5), 1015–1029. https://doi.org/10.1007/s11119-018-09628-4

[13] Guntur, S. R., Gorrepati, R. R., & Dirisala, V. R. (2018). Internet of Medical Things. In Medical Big Data and Internet of Medical Things (Issue October 2018, pp. 271–297). CRC Press. https://doi.org/10.1201/9781351030380-11

[14] Hamad, A. A., Thivagar, M. L., Alshudukhi, J., Alharbi, T. S., Aljaloud, S., Alhamazani, K. T., & Meraf, Z. (2021). Secure Complex Systems: A Dynamic Model in the Synchronization. Computational Intelligence and Neuroscience, 2021, 1–6. https://doi.org/10.1155/2021/9719413

[15] Hurt, A. (2019). Internet of Medical Things emerges. Dermatology Times, 40(10), 52–58. http://ezproxy.uct.ac.za/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=cin20&AN=138944526&site=ehost-live

[16] Kakhi, K., Alizadehsani, R., Kabir, H. M. D., Khosravi, A., Nahavandi, S., & Acharya, U. R. (2022). The internet of medical things and artificial intelligence: trends, challenges, and opportunities. Biocybernetics and Biomedical Engineering, 42(3), 749–771. https://doi.org/10.1016/j.bbe.2022.05.008

[17] Khaki, S., Wang, L., & Archontoulis, S. V. (2020). A CNN-RNN Framework for Crop Yield Prediction. Frontiers in Plant Science, 10. https://doi.org/10.3389/fpls.2019.01750

[18] Ma, C., Wang, H., & Hoi, S. C. H. (2020). Multi-label Thoracic Disease Image Classification with Cross-Attention Networks. Singaporean Journal of Radiology, 21, 1–9.

[19] Manickam, P., Mariappan, S. A., Murugesan, S. M., Hansda, S., Kaushik, A., Shinde, R., & Thipperudraswamy, S. P. (2022). Artificial Intelligence (AI) and Internet of Medical Things (IoMT) Assisted Biomedical Systems for Intelligent Healthcare. Biosensors, 12(8). https://doi.org/10.3390/bios12080562

[20] Roehrs, A., da Costa, C. A., da Rosa Righi, R., Rigo, S. J., & Wichman, M. H. (2019). Toward a Model for Personal Health Record Interoperability. IEEE Journal of Biomedical and Health Informatics, 23(2), 867–873. https://doi.org/10.1109/JBHI.2018.2836138

[21] Ekuobase, G. O. and Ebietomere, E. P.(2013). Ontology for Nigeria Case Laws. African Journal of Computing and ICTs. Vol.6, No.2, pp.177-194.

[22] [14] Breuker, J., Elhag, A., Petkov, E. and Winkels, R.(2022). Ontologies for Legal Information Serving and Knowledge Management. Legal Knowledge and Information Systems, Jurix '02: The fifteenth Annual Conference, Amsterdam, IOS Press, pp. 73-82.

[23] D. Riano, F Real, J. A. Lopez-Vallverdu, F. Campana, S. Ercolani, P. Mecocci, R. Annicchiarico, and C. Caltagirone. An ontology-based personalization of health-care knowledge to support clinical decisions for chronically ill patients, J. Biomed. Inform. 45(3): 429-446 (2022).

[24] V. Varma. Use of ontologies for organizational knowledge management and knowledge management systems. In: R. Sharman, R. Kishore, and R. Ramesh (eds.), Ontologies a Handbook of Principles, Concepts and Applications in Information Systems. Springer Science+Business Media, New York, 2009, pp. 21-47.

[25] M. K. Smith, C. Welty, and D. L. McGuinness. OWL Web Ontology Language Guide, 2009 [cited 2015 May 1]; Available from: http://www.w3.org/TR/owl-guide/.

[26] RDF Working group. RDF - Semantic Web Standards, Semantic web, [cited 2015 May 1]; Available from: http://www.w3.org/RDF/

[27] Caralt, C. N. (2018). Modelling Legal Knowledge through Ontologies. OPKJ: the Ontology of Professional Judicial Knowledge. Ph.D. Thesis, Department de Ciencia Political I Dret Public, Universitat Autonoma De Bercelona, 527pp.

[28] Su, X. and Ilebrekke, L. (2022). A Comparative Study of Ontology Languages and Tools. Springer, Verlag, pp. 3-12.

[29] Horridge, M., Jupp, S., Moulton, G., Rector, A., Stevens, R. and Wroe, C. (2017). A Practical Guide To Building OWL Ontologies Using Protégé 4 and CO-ODE Tools Edition 1.1. University of Manchester, UK, 103pp.

[30] Weiten, W. (2017). Psychology: Themes and Variations. Wadsworth Publishing, 7th ed., 880pp.

[31] G. O. Ekuobase, and I. Anyaorah, "Tail Tolerance of Web Services Solution Built on Replication Oriented Architecture (ROA)," *Canadian Journal of Pure and Applied Science*. SENRA: British Columbia, vol. 8, no. 2, pp. 2943–2954, 2014.

[32] C. D. Manning, M. Surdeanu, , J. Bauer, J. Finkel, J. B. Steven and D. McClosky, "The Stanford CoreNLP Natural Language Processing Toolkit," *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations,* pp. 55–60, 2014. https://doi.org/10.3115/v1/P14-5010

[33] M. Khabsa, S. Carman, S. R Choudhury, and C. L. Giles, "A Framework for Bridging the Gap between Open Source Search Tools," In SIGIR 2012 Workshop on Open Source Information Retrieval, pp. 32–39, 2012