

Ontology Data Modeling of Indonesian Medicinal Plants and Efficacy

Received:
14 Juni 2022
Accepted:
23 June 2022
Published:
13 August 2022

¹Rizky Maulidya Afifa, ²Wisnu Ananta Kusuma, ³Annisa
^{1,2,3}*Computer Science, Institut Pertanian Bogor*
E-mail: ¹rizkyafifa@apps.ipb.ac.id, ²ananta@apps.ipb.ac.id,
³annisa@apps.ipb.ac.id

*Corresponding Author

Abstract—People use medicinal plants for as early prevention and treating disease. Medicinal plants must be careful not to cause side effects, so knowledge is needed. Medicinal plant knowledge is stored using an ontology data model. In some ontology studies, there are still shortcomings in managing information, namely the absence of a relationship between scientific terms related to medicinal plants and phrases already known to the public. Hence, it is necessary to have this relationship. In other studies, there is no information related to disease protein, so this research also develops ontologies to enrich knowledge about medicinal plants and their efficacy. Based on the results, the developed ontology test can build a relationship between scientific terms of therapeutic pants and phrases that are known to the public. The public also knows which proteins affect a disease, so public knowledge about medicinal plants is getting wider.

Keywords—Data Model; Knowledge; Medicinal Plant; Ontology

This is an open access article under the CC BY-SA License.



Corresponding Author:

Rizky Maulidya Afifa,
Computer Science,
Institut Pertanian Bogor,
Email: rizkyafifa@apps.ipb.ac.id



I. INTRODUCTION

National Food and Drug Agency (BPOM) 2020 [1] states that Indonesia has 30.000 plant species, and around 9600 are known to have medicinal properties. Plants are known to have efficacy, as herbal medicines have not been used optimally because of the lack of medicinal plant knowledge. Medicinal plant knowledge tends to be degraded so that, if it is not well documented, it can be lost and eroded as natural medicine [2]. Knowledge of medicinal is not only limited to their benefits. It must be viewed from a pharmacology perspective, such as side effects, contraindications, and the content of therapeutic plant compounds. Information about compound content is needed to determine the biological activity of medicinal plants because one medicinal plant can have more than one pharmacological effect [3]. Information on side effects and contraindications of medicinal plants is needed to be more careful in the use of medicinal plants.

Medicinal plant data is the form of knowledge descriptions, so it requires data modeling that can store hierarchical knowledge descriptions and perform knowledge inference. The inference is making conclusions to generate new knowledge [4]. The data modeling method used is ontology. "An Ontology is an explicit specification of a conceptualization," The ontology explains the concept of knowledge of domain such as classes, relations, and objects that humans can understand well to improve semantic interoperability [5]. Ontology can perform knowledge inference, which is suitable for managing medicinal plant data [6]. Conceptual ontology explains the meaning of a domain, and ontology is more concentrated on how semantic concepts are interrelated [7]. Ontology is a link between users and computers so that users get knowledge of a domain such as medicinal plants. Designing an ontology model such as a medicinal plant ontology allows users to obtain information about medicinal plants, such as side effects and the correlation between medicinal plants and diseases [8].

Previous research has designed an ontology based on medicinal plant ethno medicine and produced a prototype ontology of medicinal plants [9]. This research expects the addition of new domains in medicinal plant knowledge. Building an ontology from scratch is unnecessary in developing an understanding of medicinal plants. However, it can use existing ontologies to produce complete and accurate information about the domain of knowledge [10]. Fazriani's research used the ontology data modeling method to build an understanding of mobile-based medicinal plants [11]. The drawback of this study is there is no relationship between the activity of the compound and the term efficacy known to the public. This relationship is needed so that people can more easily understand scientific terms. Several studies used ontology data models for the development of medicinal plants other than Indonesia medicinal plants, including Thai herbal

medicine [12], Persian herbal medication [13], Indian medicinal plants with standardized medical terms [14], and individual ontology of Thai ginger based on taxonomy [15].

This study aims to design an ontology model of medicinal plants by building a link between the activity of compounds and the efficacy of medicinal plants, as well as being able to perform knowledge inference, which has not been done in previous studies, namely the research of Wardani [9] and Fazriani [11]. This research also adds a new domain, namely the disease protein domain, to expand medicinal plant knowledge. This research will produce a medicinal plant ontology data model that is more flexible in adding relationships and entities. The ontology data model uses the Web Ontology Language (OWL) format. Testing ontology model of medicinal plants uses the Simple Protocol and RDF Query Language (SPARQL). SPARQL is a standard query language for extracting information. SPARQL is structurally similar to SQL queries, but the performance of SPARQL is more semantic compared to SQL, because SPARQL can extracting information from data that has an OWL format [16]. OWL has a triple data structure, namely subject, predicate, and object, so that SPARQL can produce information in the form of class hierarchies from the ontology of medicinal plants [17]. In future research, the ontology data model with OWL format is expected to assist the development of the web-based semantic system or Knowledge Management System (KMS).

II. RESEARCH METHOD

This research has several stages: the first step is data collection, the second step is data analysis, the third step is ontology design, the fourth step is ontology implementation, and the final step is ontology testing seen in Figure 1. This research uses qualitative techniques for data analysis because collected medicinal plant data describe knowledge [18] and uses ontology model data to describe the medicinal plant domain and hierarchical structure clearly [19].

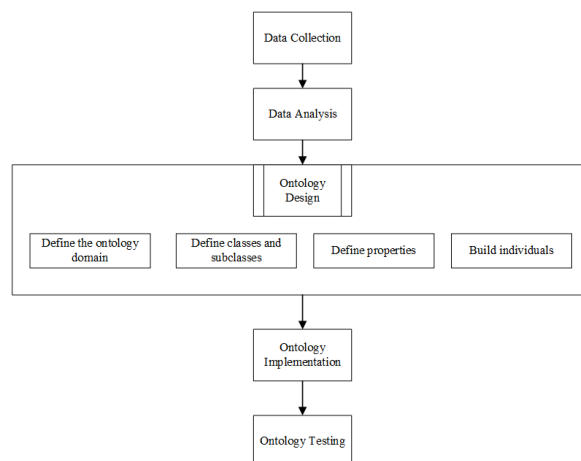


Figure 1. RESEARCH FLOW

The following is an explanation of the research flow in Figure 1:

1. Data Collection

Data collection is the first stage in designing the medicinal plant ontology data model. Data collection to collect all information about medicinal plants and their efficacy. Data collection is sourced from literature and website about medicinal plants. Data collection is done manually by downloading data on medicinal plants from several websites. The data needed is related to medicinal plants and diseases, such as plant taxonomy, side effects, use of plants, the content of plant compounds, and proteins that affect specific diseases. The data source can be seen in Table 1.

Table 1. RESEARCH DATA

Data	Source
Medicinal Plant Data	1. Literature of 100 Top Tanaman Obat Indonesia (National Institute of Health Research and Development 2011) 2. Literature of Tanaman Obat (Indonesian Agency for Agricultural Research and Development 2019)
Plant Efficacy Data	Literature of Ramuan Obat Tradisional Indonesia (Republic of Indonesia Ministry of Health 2016 and 2017)
Plant Taxonomy Data	http://plantamor.com/
Medicinal Plant Name Data	http://sioba.pom.go.id/
Plant Usage Data	Literature of Daftar Tanaman Obat Herbal Indonesia (Spiritual Youth Sapta Darma Sragen 2014)
Plan Compound Data	https://phytochem.nal.usda.gov/phytochem/search http://www.knapsackfamily.com/KNAPSAcK_Family/
Disease Protein Data	https://www.genecards.org/

2. Data Analysis

Data analysis used qualitative techniques because medicinal plant data described knowledge [18]. There are several processes at the data analysis stage: first, reviewing all data on medicinal plants that have been collected. Second, perform data reduction to retrieve core data related to medicinal plants. The Third is the presentation of data in tabular form to produce information related to medicinal plants. Fourth, draw conclusions and information from the tables presented to support knowledge in ontology design.

3. Ontology Design

In this research, the ontology design adopted the research that Mitsis [20] had done in designing ontology for nutrition and food. A simple ontology design on the ontology of medicinal plants and their efficacy can be seen in Figure 2.

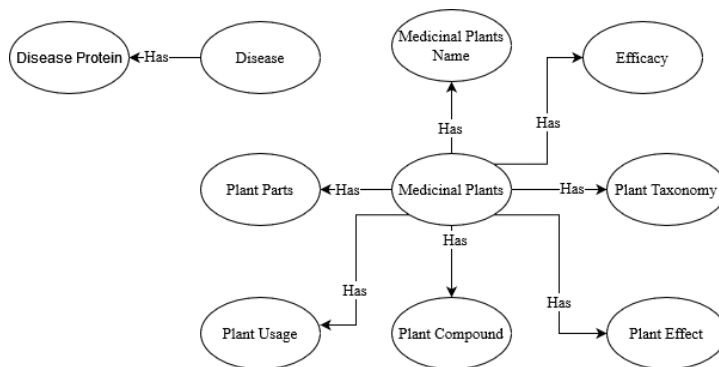


Figure 2. ONTOLOGY DESIGN

The ontology design has several stages: determining the ontology domain, defining classes and sub-classes, defining properties, and building individuals. Determination of the ontology domain provides boundaries or scope in developing medicinal plant ontologies. The formation of the ontology scope is based on data grouped based on any information related to medicinal plants. The public domain of the ontology forms a general class, and then a subclass of the public course is created—the top-down method is used in developing lessons and sub-classes. The top-down approach builds an available class and makes a more specific type to complete information about medicinal plants. The next step is to define the properties. A property describes the relationship between classes or classes with the value of an object, thus forming a conceptual model of medicinal plants. Ontology is more concentrated on how semantic concepts or data concepts are interrelated, so the stage of defining properties is the most critical stage of ontology design [7]. The last step of ontology design is to build the individual. Individuals are made according to their respective classes.

4. Ontology Implementation

The ontology implementation stage, using the protégé 5.5 tools. Protégé tools to implement the ontology design into the software so that the results of the ontology model can be visualized [21].

The ontology implementation has several processes: building classes and subclasses, then building properties between classes. The property in protégé consists of object properties that connect or map between one type and another and data properties that connect each class with a value [12]. The last process of implementing the ontology is to build individuals according to their respective categories.

5. Ontology Testing

Ontology testing is the final stage of designing medicinal plant knowledge. Ontology testing utilizes the Simple Protocol and RDF Query Language (SPARQL) to query and retrieve knowledge information from the ontology model. Furthermore, in the testing stage, ontology

model inference is also carried out to validate whether the ontology model follows the research objective, namely to generate new knowledge of the medicinal plant domain.

III. RESULT AND DISCUSSION

A. Data Analysis

Data analysis used qualitative techniques. Data analysis is based on data collected about medicinal plants. The results of the data analysis showed that therapeutic plant data consisted of 110 medicinal plants, medicinal plant name data consisted of 384 plants, disease protein data consisted of 477 proteins, compound content data consisted of 456 compounds, compound activity data consisted of 280 activities, plant efficacy data consisted of 223 properties. Data on plant parts consists of 12 pieces, data on plant families consists of 53 families, data on plant genus consists of 46 genera, and data on kingdoms, divisions, classes, and orders of plants consists of 33 data. Data on how to use plants consisted of 79 benefits, and data on side effects and contraindications of plants consisted of 200 data. Data on processing and utilization of plants consists of 238 data. The data from the analysis were grouped to obtain the general domain of medicinal plants.

B. Ontology Design

The ontology design consists of several stages :

a. Define the ontology domain

The ontology domain is needed to determine the limits or space of medicinal plant knowledge. The ontology domain scope is determined based on the results of data analysis. The results of the data analysis obtained that the public part of drug data became the public domain, namely medicinal plant classes, data on medicinal plant names became the public domain, namely therapeutic plant name classes, compound content and activity data became the public domain, namely plant content classes. Data on plant parts become a public domain, namely plant parts classes, kingdom, divisions, classes, orders, families, and genera data into the public domain, namely plant taxonomy classes, data on how to use and process into a general environment, namely grow use classes. Benefit and disease data became the public domain, namely the plant effect class, and protein data became the public domain, namely the disease protein class. Side and contraindications data became the public domain, namely the plant effect class, so the plant knowledge domain consisted of 9 fields from the public domain.

b. Define classes and sub-classes

The public domain that makes up the general class is built on subclasses to form a more specific type. The style of medicinal plant names comprises several sub-classes, namely Latin, regional, and foreign. Plant taxonomic class contains several sub-classes: kingdom, division, class,

order, family, and genus. The plant efficacy class consists of two sub-classes: disease and plant benefits. The plant content class consists of two sub-classes: the method of processing and the process of utilization. The course of plant effects consists of two sub-classes: side effects and contraindications.

c. Define properties

Mapping and integrating classes and subclasses will produce a relationship between other courses. Class mapping in the medicinal plant domain can be seen in Figure 3.

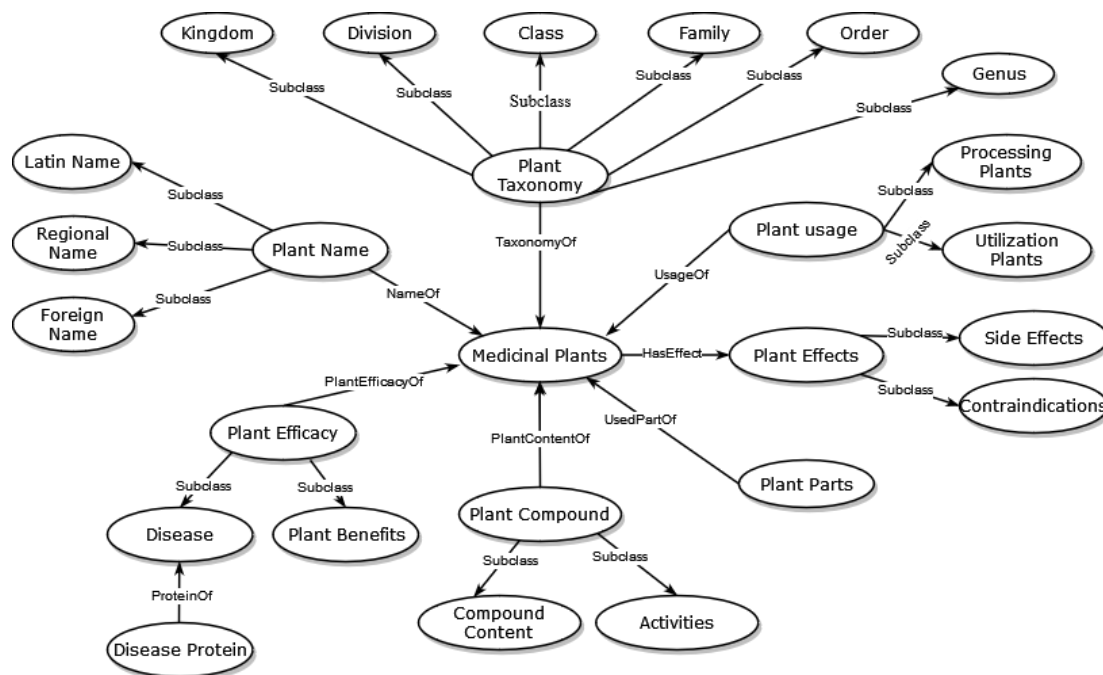


Figure 3. MAPPING DOMAIN OF MEDICINAL PLANT

Each class has properties with the medicinal plant class; then, the sub-class relates to the therapeutic plant class domain. In Figure 3, the plant efficacy class has sub-classes, namely disease and plant benefits. The sub-class of illness is medicine for illness, and the sub-class of plant benefits is an additional non-medicinal plant efficacy—examples of the effectiveness of plants for medicinal disease such as medicinal plants treating fever. The efficacy of non-medicinal plants for conditions such as medicinal plants can darken eyebrows and increase milk production. The class of plant use also has a sub-class of processing methods and ways of utilization, so the sub-class will also have properties with the medicinal plant class. Subclass mapping can be seen in Figure 4.

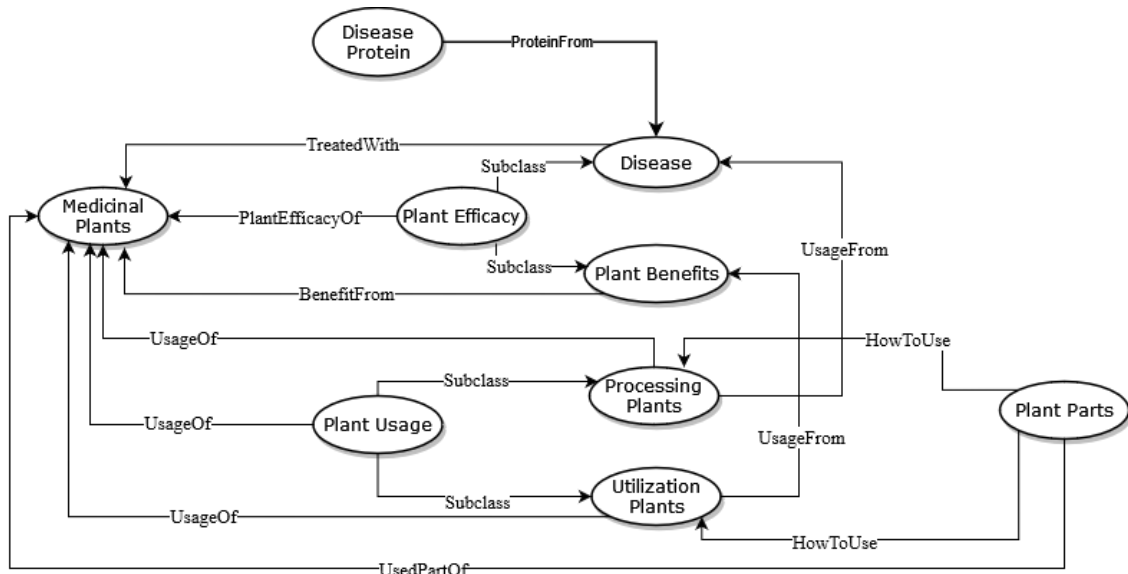


Figure 4. MAPPING SUBCLASSES WITH MEDICINAL PLANTS

Mapping between classes will form a rule that produces a conceptual model of medicinal plants. The conceptual model of medicinal plants creates a hierarchy between types to have new information in the form of medicinal plant knowledge. The conceptual model of medicinal plants can be seen in Figure 5.

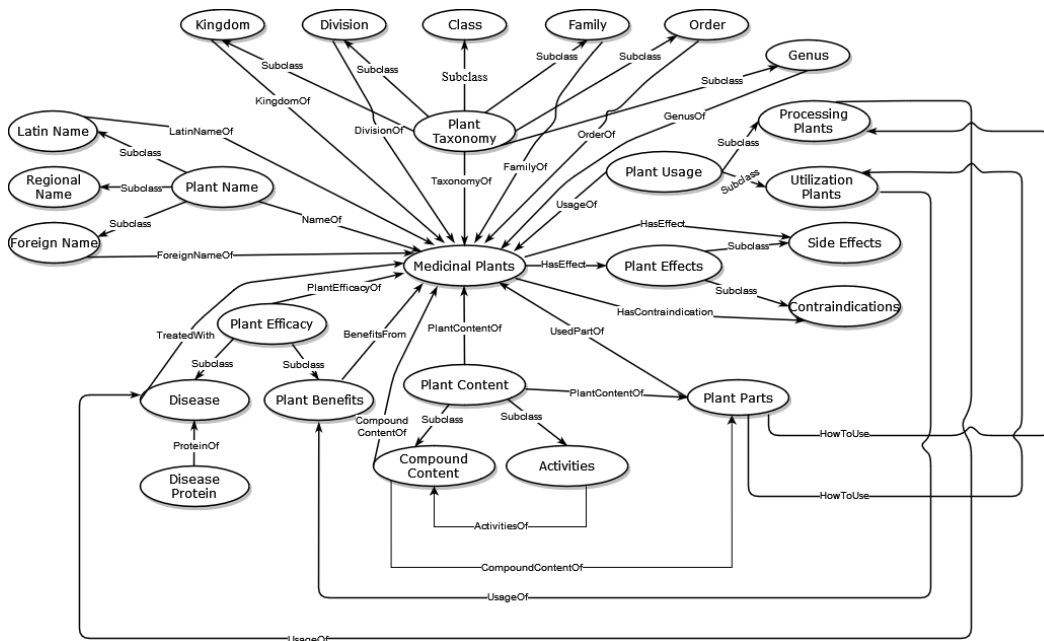


Figure 5. MEDICINAL PLANT CONCEPTUAL MODEL

d. Build individuals

The final stage in ontology design is building individuals. Individuals constructed according to their respective classes have individual ginger, fennel, cumin leaves, etc.

C. Ontology Implementation

The ontology implementation uses protégé 5.5 tools to visualize the ontology model. The first stage of ontology implementation is to define the ontology domain of medicinal plants into classes whose representation using graphs can be seen in Figure 6.



Figure 6. SCREENSHOT OF CLASS AND HIERARCHY IN PROTEGE

The next stage of ontology implementation in protégé is defining properties. Properties in the protégé tool are divided into two, namely, object properties and data properties. Besides that, the protégé will build ontology property inverse. Property implementation on protégé can be seen in Figure 7.

Object Property	Domain	Range	Inverse	Data Property	Func	Domain	Range
owl:topDataProperty				owl:topDataProperty	<input type="checkbox"/>		
memiliki_kontraindikasi	tanaman_obat	tanaman_obat	tanaman_obat	memiliki_kontraindikasi	<input checked="" type="checkbox"/>	tanaman_obat	xsd:string
memiliki_efek_tanaman	tanaman_obat	tanaman_obat	tanaman_obat	memiliki_efek_tanaman	<input checked="" type="checkbox"/>	tanaman_obat	xsd:string
memiliki_nama_daerah	tanaman_obat	tanaman_obat	tanaman_obat	memiliki_nama_daerah	<input checked="" type="checkbox"/>	tanaman_obat	xsd:string
deskripsi_protein	protein_penakit	protein_penakit	protein_penakit	deskripsi_protein	<input checked="" type="checkbox"/>	protein_penakit	xsd:string

Figure 7. SCREENSHOT OF PROPERTY IMPLEMENTATION IN PROTEGE

The final stage of ontology implementation is individual development. According to each class, personal development can be seen in Figure 8, an example of the particular action in the medicinal plant class.



Figure 8. SCREENSHOT OF BUILD INDIVIDUALS IN PROTEGE

D. Ontology Testing

Ontology testing uses the Simple Protocol and RDF Query Language (SPARQL). SPARQL functions to generate queries from the conceptual model ontology that has been designed. SPARQL query generation based on the built property. SPARQL query consists of triples (subject, predicate, object); for example, to make a SPARQL query asking “what is the effect of medicinal plants,” then SPARQL is formed: ?MedicinalPlant medicinal:HasEffect ?PlantEffects.

?MedicinalPlant as subject, medicinal: HasEffect as property and ?PlantEffect as object. This test will refer to the “what” question. For example, SPARQL asks about “what parts of medicinal plants are used to treat diseases, what are the additional benefits of medicinal plants, what is the content of medicinal plants, what is the Latin name of medicinal plants, what effects are produced from medicinal plants, what disease proteins that affect disease can cause treated with medicinal plants, as well as what are the names of the terms of medicinal plants known to the general public.” The SPARQL query and results can be seen in Figure 9.

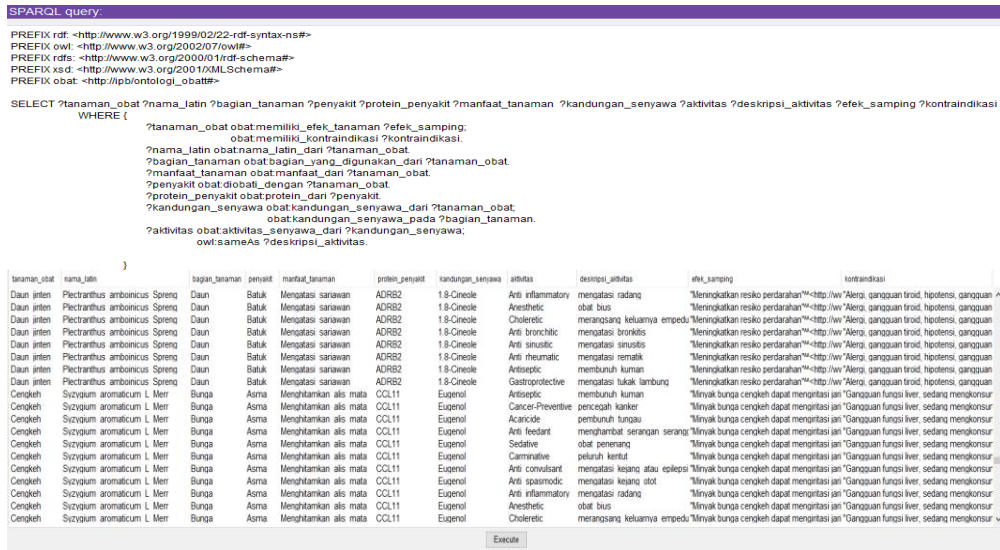


Figure 9. SCREENSHOT SPARQL TESTING MEDICINAL PLANT ONTOLOGY

Query testing is based on the conceptual model of medicinal plants, namely the Latin name of medicinal plants, parts used, treatment of diseases and additional benefits, disease protein, compound content, activity, and description of the action of compounds needed by the community as well as side effects and contraindications of medicinal plants. SPARQL query testing generates knowledge about medicinal plants. Figure 10 shows a knowledge graph from the SPARQL query on medicinal plant knowledge.

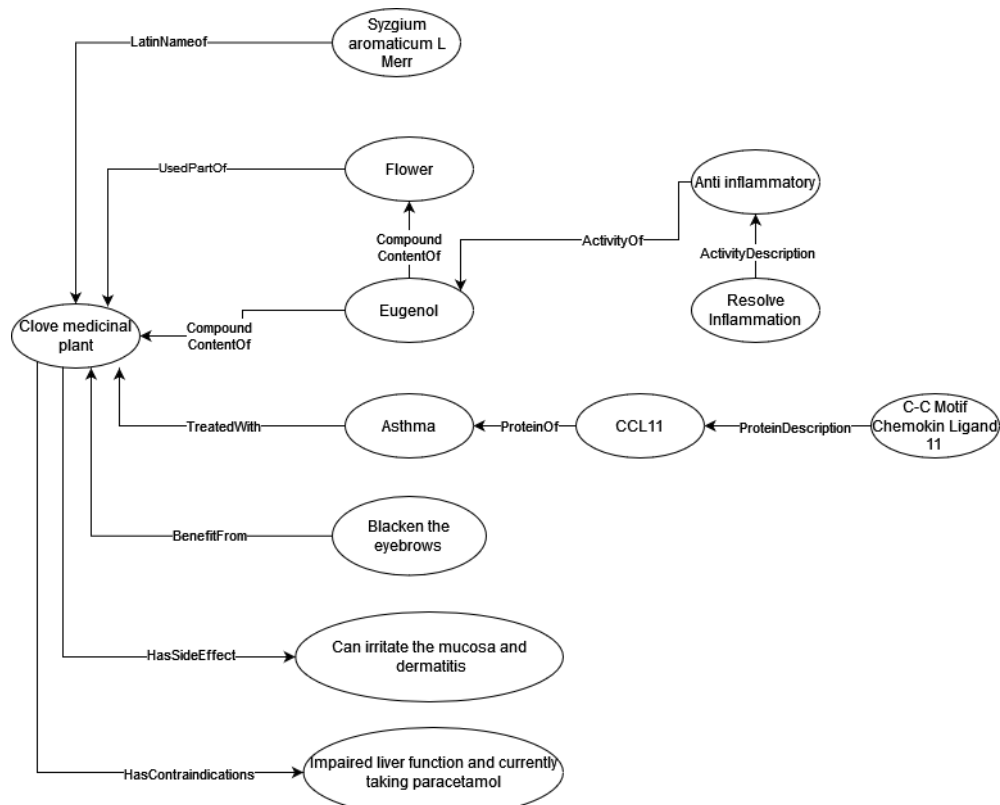


Figure 10. CLOVE MEDICINAL PLANT KNOWLEDGE

Figure 10 shows the results of the SPARQL query, for example, taking knowledge of the clove medicinal plant. The graph shows that the clove plant has the Latin name *Syzygium Aromaticum*; the part used by the clove is the flower. The clove can treat asthma and has other benefits. Namely, it can blacken the eyebrows. Cloves have side effects that can irritate the mucosa and have contraindications that are not allowed in patients with a history of impaired liver function and patients who take paracetamol. Clove plants contain Eugenol compounds, found in flowers with Anti-Inflammatory compound activity, which people know as overcoming inflammation. From the results of the SPARQL query, the ontology model built is by the research objective: to build a relationship between compound activity and the term efficacy known to the public so that people understand the term medicinal plant efficacy.

In protégé, there are reasoner tools. The reasoner function is to validate the medicinal plant ontology model and check whether or not there are errors in making the ontology model. In this research, the reasoner tool used is Pellet. The Pellet tools will validate the ontology model and perform the inference process. The inference is used to generate new knowledge of the medicinal plant domain. An example of the validation of this ontology model is checking on individual cumin leaves. The knowledge inference process and results can be seen in Figure 11 and Figure 12.

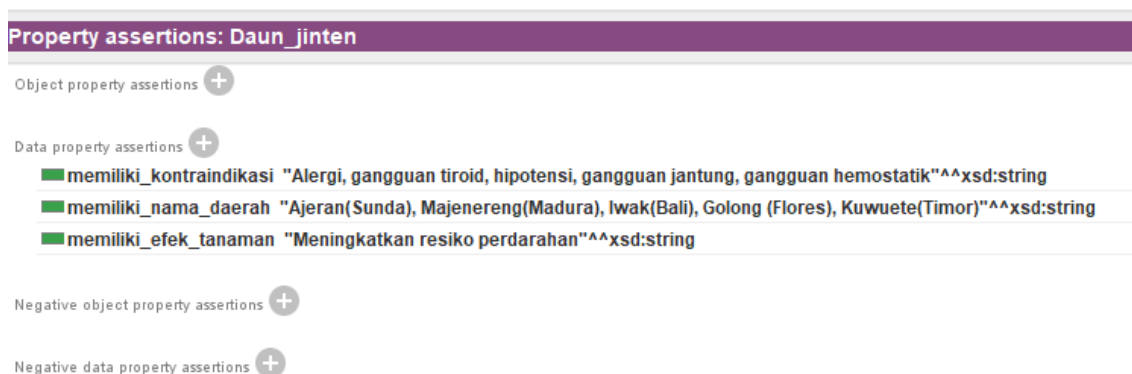


Figure 11. SCREENSHOT OF INDIVIDUAL BEFORE INFERENCE PROCESS IN PROTEGE

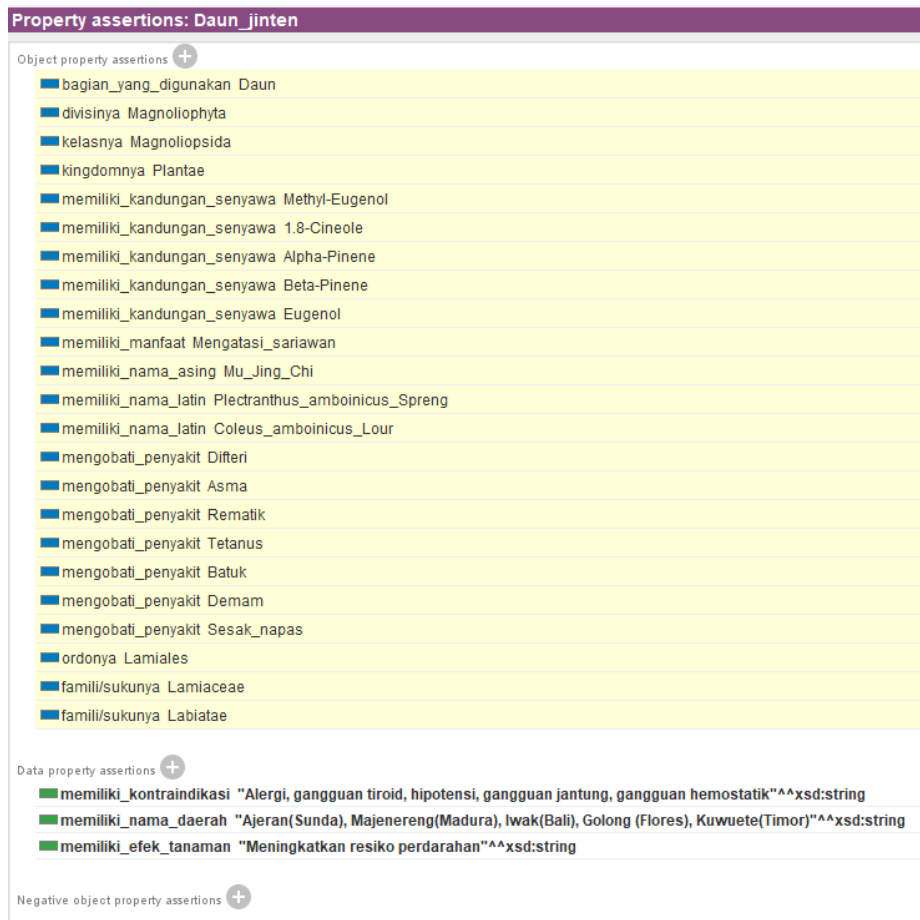


Figure 12. SCREENSHOT OF THE INDIVIDUAL AFTER THE INFERENCE PROCESS IN PROTÉGÉ

The inference process generates new knowledge from individual cumin leaves. The new knowledge generated is the efficacy of plants, plant parts, compound content, and plant taxonomy. Based on the testing and validation, the ontology model can perform knowledge inference. It can reuse the medicinal plant ontology data model for further system development, such as a Knowledge Management System (KMS) because the ontology representation can be in the form of an OWL/RDF format that can be understood or easily understood by the system.

IV. CONCLUSION

The ontology model that has been designed can perform knowledge inference and build new domains, namely disease proteins, to enrich the knowledge of medicinal plants. The resulting ontology model can also build relationships between compound activities based on terms that are known to the public. The query test on the ontology model can provide results from the stored knowledge about medicinal plants. Future research hopes that additional domains of medicinal plant knowledge, such as nutritional content and more detailed knowledge such as clinical testing,

require experts in data validation so that the medicinal plant ontology model is expected to support information in the manufacture of Indonesian standard herbal medicines.

REFERENCES

- [1] Badan Pengawas Obat dan Makanan (BPOM), “Potensi Obat Herbal Indonesia,” *Bpom Ri*, vol. 11, no. 88. 2020.
- [2] Haeruddin, H. Johan, U. Hairah, and E. Budiman, “Ethnobotany database: Exploring diversity medicinal plants of dayak tribe borneo,” in *International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, 2017, pp. 1–6, doi: 10.1109/EECSI.2017.8239094.
- [3] N. Hilal A. Syahrir, F. Mochamad Afendi, and B. Susetyo, “Efek Sinergis Bahan Aktif Tanaman Obat Berbasis Jejaring dengan Protein Target,” *J. Jamu Indones.*, vol. 1, no. 1, pp. 35–46, Mar. 2016, doi: 10.29244/jjdn.v1i1.30594.
- [4] E. E. Ogheneovo and P. A. Nlerum, “Knowledge Representation in Artificial Intelligence and Expert Systems Using Inference Rule,” *Int. J. Sci. Eng. Res.*, vol. 11, no. 4, pp. 1886–1900, 2020, [Online]. Available: <https://www.researchgate.net/publication/347504395>.
- [5] T. R. Gruber, “A Translation Approach to Portable Ontology Specifications,” *Knowl. Acquis.*, vol. 5, pp. 199–220, 1993, [Online]. Available: http://ac.els-cdn.com/S1042814383710083/1-s2.0-S1042814383710083-main.pdf?_tid=28c5c43c-df11-11e2-b19c-00000aacb35f&acdnat=1372327738_c3518b0aab22af436aef0c5e3e2dd48e.
- [6] F. Z. Laallam, M. L. Kherfi, and S. M. Benslimane, “Using Ontologies to Overcoming Drawbacks of Databases and Vice Versa: A Survey,” *Comput. Sci. Eng. An Int. J.*, vol. 3, no. 2, pp. 1–21, 2013, doi: 10.5121/cseij.2013.3201.
- [7] M. Uschold, “Ontology and database schema: What’s the difference?,” *Appl. Ontol.*, vol. 10, no. 3–4, pp. 243–258, 2015, doi: 10.3233/AO-150158.
- [8] R. Gunawan and K. Mustofa, “Finding knowledge from Indonesian traditional medicine using semantic web rule language,” *Int. J. Electr. Comput. Eng.*, vol. 7, no. 6, pp. 3674–3682, 2017, doi: 10.11591/ijece.v7i6.pp3674-3682.
- [9] D. W. Wardani, S. H. Yustianti, U. Salamah, and O. P. Astirin, “An Ontology of Indonesian Ethnomedicine,” *Int. Conf. Information, Commun. Technol. Syst.*, pp. 47–52, 2014.
- [10] H. Satria, R. S. Priya, L. H. Ismail, and E. Supriyanto, “Building and Reusing Medical Ontology for Tropical Diseases Management,” *Int. J. Educ. Inf. Technol.*, vol. 6, no. 1, pp. 52–61, 2012.
- [11] A. Fazriani, W. A. Kusuma, and I. Batubara, “Sistem Berbasis Pengetahuan Tumbuhan Obat Pusat Studi Biofarmaka,” *J. Jamu Indones.*, vol. 4, no. 1, pp. 17–27, 2019, doi: 10.29244/jji.v4i1.88.
- [12] W. Tungkwampian, A. Theerarungchaisri, and M. Buranarach, “Development of Thai herbal medicine knowledge base using ontology technique,” *Thai J. Pharm. Sci.*, vol. 39, no. 3, pp. 102–109, 2015.
- [13] H. Shojae-Mend, H. Ayatollahi, and A. Abdolahadi, “Developing a mobile-based disease ontology for traditional Persian medicine,” *Informatics Med. Unlocked*, vol. 20, 2020, doi: 10.1016/j.imu.2020.100353.
- [14] G. Vadivu and S. Waheeta Hopper, “Ontology mapping of Indian medicinal plants with standardized medical terms,” *J. Comput. Sci.*, vol. 8, no. 9, pp. 1576–1584, 2012, doi: 10.3844/jcssp.2012.1576.1584.

- [15] N. Kaewboonma, T. Supnithi, and J. Panawong, "Developing ontology for Thai Zingiberaceae: From taxonomies to ontologies," 14th Int. Conf. Electr. Eng. Comput. Telecommun. Inf. Technol., pp. 596–599, 2017, doi: 10.1109/ECTICon.2017.8096308.
- [16] A. Khamparia, B. Pandey, and V. Pardesi, "Performance analysis on agriculture ontology using SPARQL query system," 2014 Int. Conf. Data Min. Intell. Comput. ICDMIC 2014, 2014, doi: 10.1109/ICDMIC.2014.6954258.
- [17] N. Kumar and S. Kumar, "Querying RDF and OWL data source using SPARQL," 2013 Fourth Int. Conf. Comput. Commun. Netw. Technol. ICCCNT, 2013, doi: 10.1109/ICCCNT.2013.6726698.
- [18] S. Gqibani, N. Clarke, and A. L. Nel, "Motivation for developing a qualitative methodological basis for the analysis of historical curriculum changes," IEEE Glob. Eng. Educ. Conf. EDUCON, pp. 637–644, 2016, doi: 10.1109/EDUCON.2016.7474617.
- [19] X. Li, Y. Zhang, J. Wang, and Q. Pu, "A Preliminary Study of Plant Domain Ontology," Proc. - 2016 IEEE 14th Int. Conf. Dependable, Auton. Secur. Comput. DASC 2016, 2016 IEEE 14th Int. Conf. Pervasive Intell. Comput. PICom 2016, 2016 IEEE 2nd Int. Conf. Big Data, pp. 109–112, 2016, doi: 10.1109/DASC-PICom-DataCom-CyberSciTec.2016.36.
- [20] K. Mitsis, K. Zarkogianni, N. Bountouni, M. Athanasiou, and K. S. Nikita, "An Ontology-Based Serious Game Design for the Development of Nutrition and Food Literacy Skills," Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS, pp. 1405–1408, 2019, doi: 10.1109/EMBC.2019.8856604.
- [21] A. Nugroho, "Membangun Ontologi Jurnal Menggunakan Protege," J. Transform., vol. 10, no. 1, pp. 20–25, Jul. 2012, doi: 10.26623/transformatika.v10i1.66.