

## THE DESIGN AND APPLICATION OF DOMAIN ONTOLOGY IN THE FIELD OF DERMATOSIS

XINGLIANG QI<sup>1</sup>, YANG ZHOU<sup>2,\*</sup> AND ZHENGUO WANG<sup>2</sup>

<sup>1</sup>The Second Affiliated Hospital of Shandong University of Traditional Chinese Medicine  
No. 1, Jingba Road, Jinan 250001, P. R. China

<sup>2</sup>Institute for Literature and Culture of Chinese Medicine  
Shandong University of Traditional Chinese Medicine  
No. 4655, Daxue Road, Jinan 250355, P. R. China

\*Corresponding author: zhouyangjn@163.com

Received March 2018; accepted June 2018

**ABSTRACT.** *Based on the actual process of clinical diagnosis and treatment, by analyzing the clinical characteristics of dermatosis, we determined the construction principle of domain ontology in the field of dermatosis, and applied it to the construction of clinical diagnosis system. We set up a priori database and extracted the key nodes in the process of diagnosis and treatment with the skin disease of erythematous symptoms as the research object. We created the characteristic rules and their application methods in the process of dermatology diagnosis and treatment, and constructed the tree structures, the classification dictionary and the domain ontology in the field of dermatosis. On this basis, we set up a clinical diagnosis and treatment system for dermatosis with skin erythema symptoms as an example.*

**Keywords:** Ontology, Dermatitis, Clinical diagnosis and treatment, Erythema

**1. Introduction.** With the increasing number of medical data, the application of AI is more and more popular. It is becoming more and more aware that using the computer to deal with some data related work is the inevitable trend of medical development. The automation of the diagnosis and treatment system is the inevitable result of the application of the computer.

Compared with the complicated diagnosis and treatment of clinical diseases, the diagnosis process of skin diseases is relatively clear and concise. The diagnosis of most skin diseases is based on the manifestations of skin lesions and the histopathological characteristics of skin. The automation of the clinical diagnosis and treatment of dermatosis based on domain ontology can be used as the starting point of the research. In the process of research, the design and construction of the domain ontology in the field of dermatosis, is the key and important point of the whole research. The establishment of the domain ontology depended on the decomposition and reconstruction of the corpus of dermatosis. The establishment of a dermatology professional dictionary is the basis for the construction of domain ontology. After the classification and attribute setting of the dictionary, we decomposed and reconstructed the corpus of dermatosis under controllable rules.

**2. Purpose and Background.** The process of clinical diagnosis and treatment is a necessary experience for both doctors and patients to complete the diagnosis and treatment. The complete process consists of the following aspects: the first contact between doctors and patients, the clinical data collection and summary, analysis, identification, disease induction, timely supplement, the diagnosis, the treatment, the feedback and correction

of diagnosis and treatment, the final assessment. The process is now still mainly relied on by the doctor himself.

With the increasing number of medical data, the application of AI is more and more popular. It is becoming more and more aware that using the computer to deal with some data related work is the inevitable trend of medical development. With the characteristics of medical treatment, doctors directly face human health and life safety. Everyone wants the best diagnosis and treatment. Because the medical data is huge and messy, an important question is how to choose the right method to process the data efficiently and accurately.

In this field, some scholars have put forward some theoretical and practical methods to try to solve these problems.

Ma et al. put forward the knowledge-based medical auxiliary diagnosis system. The system established the invented index table by using word segmentation of the disease information in medical knowledge base; secondly, the similarities between the input symptoms and the related diseases were calculated and sorted; then, the system used the weight ratio of the feedback information and the original was determined dynamically to optimize the diagnosis based on user feedback; finally, the Bayes classification algorithm was used to make recommendations based on illness case [1].

Zhang et al. put forward a model of hypertension diagnosis and treatment system based on ontology and case-based reasoning to use computers as an assistant way of diagnosis and treatment of hypertension. Firstly, hypertension domain ontology and reasoning rules are constructed. Next, Jess, which is an inference machine, is used to reason. The knowledge base is enriched by reasoning to acquire the new knowledge and some instances of the ontology are extracted as small case base. Then, Jena is used to complete the parallel query of ontology base and case base. Finally, the query and reasoning results are returned and a prescription of drug treatments and non-drug treatments is given, which is suitable for the patient [2].

Yang et al. transformed Clinical Practice Guideline (CPG) into a set of computerized process with extending standards and available ontology publicly. By extracting subject elements of clinical trials from the CPG semi-automatically, a general mapping between subject elements and ontology model was provided, and the instantiation operation was founded for completing the diagnosis and treatment processes [3].

However, the ever-changing clinical diseases are not simply right and wrong. Most patients do not come to the hospital after all symptoms appear. Even the diagnostic items of the disease may not be fully expressed in one patient. Patients may seek medical treatment at any stage of the disease, and clinical symptoms and signs may also be any form of disease. This complexity cannot be solved by one algorithm. Therefore, the above system may be somewhat different from the ideal in the accuracy of diagnosis.

The accurate, comprehensive and timely realization of computer-aided diagnosis and treatment must follow the development frontier. The following was the focus of this project: reasonable application of data model and professional data processing technology, how to combine the rigor of medicine, how to make the professional data as accurate and comprehensive as possible.

There were three important elements. First, the knowledge contained in the literature would be discovered in a comprehensive, rapid and accurate way from the mass and disordered medical literature resources. Second, the knowledge would be displayed from a different dimension, and the various associations between the knowledge carried by the literature would need to be discovered. Third, we would need to realize the leapfrog development of the literature service, and the result would be that the information retrieval service would be raised to the knowledge service of the literature.

We organized the text at the knowledge level based on the theory and method of natural language processing. By analyzing the semantic environment of domain ontology

knowledge and the structure of resource document, we have set up data classification structures and attributes using the rich semantic environment information expressed by the domain ontology. We built a data dictionary of dermatology, realized the natural language processing of document resources in the medical field, and provided the driver based on ontology for the diagnosis and treatment of skin diseases.

**3. Research Methods and Steps.** Through the research on the clinical characteristics of skin diseases and the analysis of the clinical process, we have identified the key points in the construction of the clinical diagnosis and treatment system. That would be to take the symptoms as the dominant, and to diagnose and treat the disease as the main purpose. The doctor's clinical diagnosis coincides with this in fact. In order to complete the datamation and automation of the diagnosis and treatment system, we must start with the practice and needs of clinical diagnosis and treatment so as to make a diagnosis and treatment system that really helps clinicians. It was one of the most important and time-consuming processes to develop the diagnosis and treatment system that made the symptoms fully detailed and considered the complex and varied clinical conditions of the clinic. Therefore, we determined the skin disease with erythematous symptoms as the breakthrough point. By establishing a priori database, we made a detailed analysis of the common skin lesions of erythema in department of dermatology, and analyzed the results into the construction of domain ontology and treatment diagnosis system [4].

**3.1. The choice of the corpus of dermatosis.** We used two authoritative books on dermatology as the basic corpus: *China Clinical Dermatology* and *Routine Medical Diagnosis and Treatment in Peking Union Medical College Hospital – Routine Diagnosis and Treatment of Department of Dermatology*. These two books are the most common authoritative books in clinical department of dermatology, and are widely used in clinical practice.

The field of dermatology involves all aspects of the field of clinical medicine. From the basic to the clinical, from modern to traditional, the field of medicine is very large and involves many subjects. We were committed to the clinical diagnosis and treatment of dermatology, focusing on the process of diagnosis and the choice of treatment, taking the etiology, clinical symptoms, diseases, diagnosis and treatment as the core contents. We chose *China Clinical Dermatology* (1st Edition, Zhao Bian editor) as the main source of data. A few of the contents referred to the "skin pathology diagnosis" (3rd Edition), "skin pathology" (2nd Edition), "international dermatology classification and name".

**3.2. The establishment of the dermatology professional dictionary.** We analyzed the corpus of dermatology in detail, extracted conceptual vocabularies, and then established a dermatology professional dictionary.

Considering the domain characteristics of dermatology knowledge and the complexity of knowledge structure, we constructed the dictionary based on ontology. Through this classification dictionary, unstructured text data can be formatted into readable, querying structured data. We can make a semantic analysis of the knowledge of dermatology combined data mining technology.

After designated professional fields and categories we determined the selection criteria and principles and methods. We categorized the following contents separately: professional medical vocabulary, medical loanwords, anatomic vocabulary, etiological vocabulary, skin erythema specific vocabulary, Chinese medicine vocabulary.

**3.3. The classification of the dermatology professional dictionary.** Based on MeSH, combined with the clinical specialty of dermatosis, we established the tree structures of the dermatology professional dictionary. There were 22 categories of first class catalogues, as follows: pathological anatomy, dermatology, basic pharmacology, histology,

skin diseases, symptoms, diagnosis, treatment, drugs, organisms, epidemiology, season, population, geography, metabolic and nutritional disorders, endocrine disorders, physical factors, chemical factors, social factors, infectious pathways, lesions, erythematous features.

According to the characteristics of dermatology and the need of diagnosis and treatment system in this study, this tree structure improved the classification structure of MeSH, and had been formed a more intuitive and practical classification tool. The first class catalog was sub classified according to the professional requirements. For example, the skin diseases were subdivided into 37 subcategories referring to the classification of *China Clinical Dermatology*.

**3.4. Attribute setting.** According to the tree structures of the dermatology professional dictionary, we have set up the attributes of the classification, including two categories of basic concepts and professional attributes. For example, the attributes of skin diseases were set as follows: definition, etiology, alias, pathogenesis, clinical manifestations, skin lesion characteristics, related diseases, complications, clinical examination, histopathology, diagnostic points, antidiastole, location, treatment and drugs.

The classification of dermatology was designed to meet the primary purpose of clinical diagnosis and treatment. It was divided into two types, which were expressed as 1 and 2, respectively. 1 was the concept attribute, which indicated that the attribute content was fixed from the corpus, and 2 was the professional attribute, which indicated that the attribute content had the corresponding range.

**3.5. The establishment of a priori database.** The priori database referred to the result of the erythematous dermatosis screened from all skin diseases by clinicians. The screening criteria were derived from the descriptions of the symptom details and the descriptions of the evolution of symptoms in *China Clinical Dermatology*, which were determined by the frontline clinicians. The priori database included 407 kinds of skin diseases in this study.

The priori database could be used as a feedback basis for data analysis. Compared with the running data, the data definition can be improved by correcting the data result error. Under the premise of improving the accuracy of data, this feedback can reduce the ambiguity of the semantic dictionary, and correct the deviation in the construction of ontology, can enhance the extraction of characteristic rules, and improve the key node settings and the accuracy of the results [5].

**3.6. Treatment of characteristic symptoms and signs.** Symptoms and signs that have characteristic significance for the diseases diagnosis were dealt with as special rules during the construction of the diagnosis and treatment process. There were 60 such rules such as: “slapped cheek” erythema suggestive of erythema infectiosum, erythema marginatum suggestive of skin manifestations of rheumatic fever, and facial butterfly erythema suggestive of systemic lupus erythematosus.

**3.7. The setting of key nodes.** Some symptoms or the characteristics of diseases that are of great significance in the diagnosis, were called key nodes. The addition of key nodes made the diagnosis process more concise, more accurate and more efficient, and they could also serve as a hint for clinical inquiry. All the key nodes were proved by the department of dermatology clinical staff and were perfected and revised in the confirmatory results. There were 6 key nodes as follows: color, category, accompanying symptoms, causes, prevalence, symptom distribution.

The clinical diagnosis and treatment of dermatosis is very complicated. This had made the setting of key nodes need to pay attention to a lot of rules. For example, pruritus is one of the accompanying symptoms, and includes severe pruritus and mild pruritus. The tolerance of different patients to pruritus is different. In the group with lower threshold,

mild pruritus may be described as severe one. Therefore, this kind of subjective accompanying symptom is meaningful only when the doctor can determine it. For another instance, the whole body consists of all parts of the body. Considering the location of the disease, that increases the uncertainty for the final result. For example, erysipelas, symptoms can occur in any part of the body; when the symptoms occur on the head and face, there will be a deviation according to the classification of the disease that only appears on the head and face. Scarlet fever often has a sequence of erythema symptoms, usually in the order of the neck, chest, trunk and limbs; when the symptoms of a patient have not been developed to the whole body, the selection of the area of the disease may also be misleading to the diagnosis. Therefore, in the logical setting, the complexity and randomness of the disease should be fully considered.

**3.8. The construction of domain ontology in the field of dermatosis.** After the establishment of prototype, knowledge reorganization, and the definition of characteristic rules, we created the domain ontology in the field of dermatosis.

Ontology is a philosophical concept, which is used to describe the nature of things, and to explain the system of objective existence. In the AI domain, Neches et al. interpret the ontology as: the ontology of a system consists of its vocabulary and a set of constraints on the way terms can be combined to model a domain [6].

In this study, the construction of the ontology of dermatology has the following characteristics.

- (1) Making full use of the professional thesaurus and the terminology dictionary
- (2) Using the classification results to realize the maximum unidirectional extensibility of the ontology
- (3) Application of authoritative and professional books
- (4) Effective participation of dermatology professionals

**3.9. The construction of the clinical diagnosis and treatment system.** The whole system was divided into six parts according to its functions: information collection, symptom analysis, disease diagnosis, disease information display, selection of treatment plan and data storage after diagnosis.

Our research process is shown in Figure 1.

**4. Summary.** Taking authoritative dermatology as corpus, based on the clinical diagnosis and treatment process, referring to MeSH we analyzed the clinical features of the erythematous symptoms of dermatosis, organized and summed up the relevant rules and key nodes, and have built the dermatology professional dictionary and the tree structures. At the same time, we have set up the property. On this basis, we successfully constructed the domain ontology in the field of dermatosis. Based on the ontology, we designed and realized the automation of the clinical diagnosis and treatment system. The construction of the system was closely combined with the actual clinical diagnosis and treatment process, and solved the problem of the combination of clinical needs and knowledge services, and successfully explored an ideal and practical theory and method, which provided a theoretical basis and method for the construction of clinical discipline diagnosis and treatment system.

**Acknowledgment.** This work is partially supported by 3 projects: 1) The Key Research Project of Shandong Province (Nos. 2016GSF202028 and 2016CYJS08A01-1); 2) Special Post-doctoral Innovation Program of Shandong Province (2016). The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

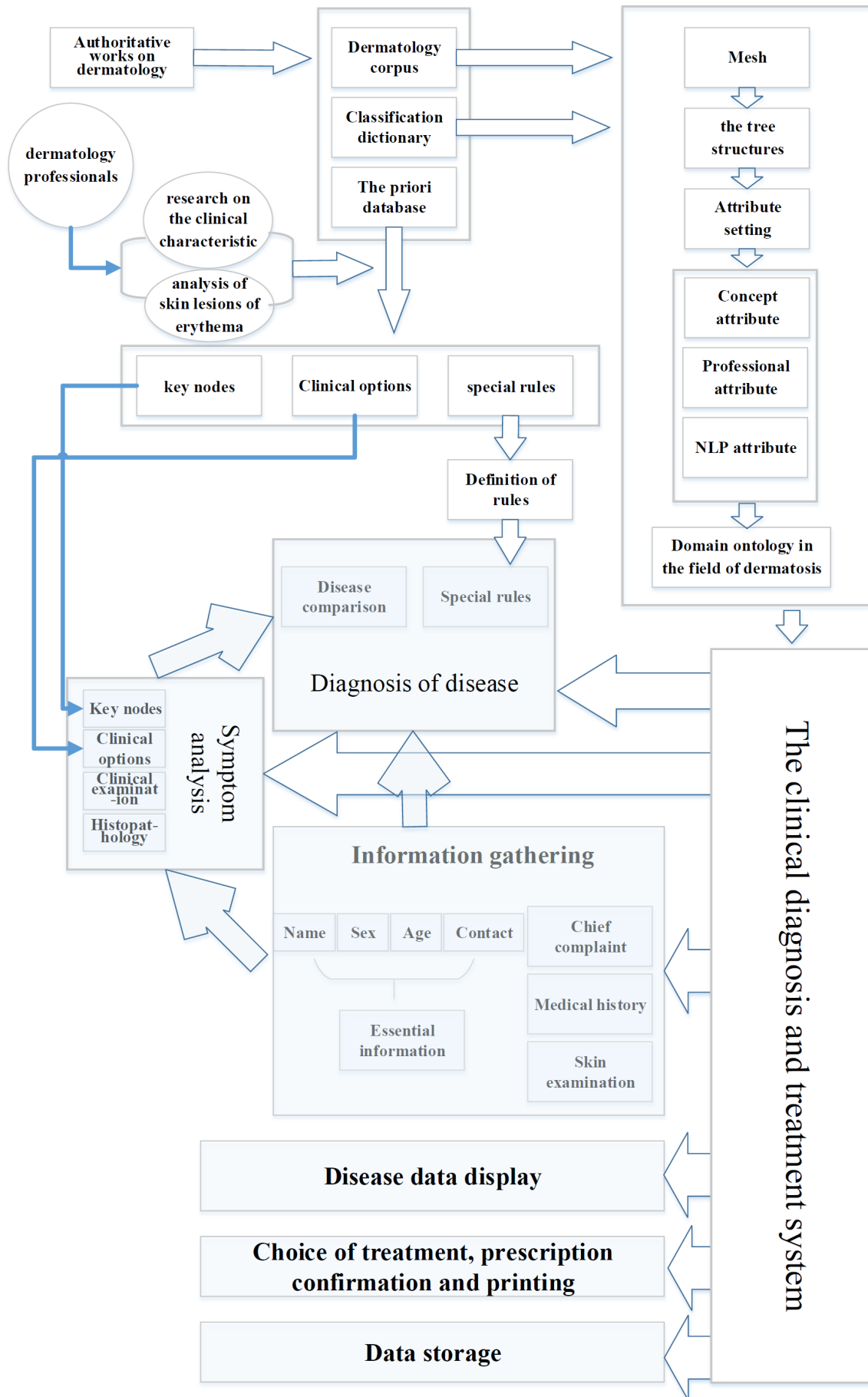


FIGURE 1. The chart of research process

**REFERENCES**

- [1] Y. Ma, H. Wang and H. Gao, Knowledge-based medical auxiliary diagnosis system, *Journal of Computer Applications*, vol.36, no.S1, pp.217-219, 2016.
- [2] W. Zhang, X. Zhang and J. Chen, Research on reasoning model of hypertension diagnosis and treatment system based on ontology, *Computer Engineering and Design*, vol.34, no.11, pp.4016-4020, 2013.
- [3] L. Yang, G. Kang and L. Zhang, Transforming clinical practice guideline into process, *Computer Integrated Manufacturing Systems*, vol.23, no.5, pp.1040-1049, 2017.
- [4] X. Qi, Y. Zhou, X. Fu and Z. Wang, Construction of the diagnosis and treatment process of dermatosis based on data-driven approach, *2017 IEEE/WIC/ACM International Conference on Web Intelligence*, pp.1143-1146, 2017.
- [5] X. Qi, C. Wei, T. Yuan, Z. Wang and Y. Zhou, Research on the diagnosis and treatment process of dermatosis based on data-driven approach, *ICIC Express Letters, Part B: Applications*, vol.7, no.11, pp.2367-2372, 2016.
- [6] R. Neches, R. Fikes, T. Finin, T. Gruber, R. Patil, T. Senator and W. R. Swartout, Enabling technology for knowledge sharing, *AI Magazine*, vol.12, no.3, pp.36-56, 1991.