Design and Implementation of Ontology Knowledge Base of Endemic Genera of Seed Plants in Yunnan Province

Lin Peng, Xu Li, lin-nan Yang

(College of Big Data, Yunnan Agricultural University, Kunming 650201)

Correspondence to: lin-nan Yang (lny5400@sina.com)

Abstract: There is complex terrain, various soil and climate in Yunnan Province. The unique natural condition produces the rich resources of the species. The spermatophyte endemic genus resource is the most quantity of China. It is essential to protect and use Yunnan spermatophyte endemic genus resource reasonably. According to Yunnan spermatophyte endemic genus resource various data formats, deficient integration, low-efficiency intelligent retrieval, lack of individualized information service mode and more problems, This paper builds Yunnan spermatophyte endemic genus ontology knowledge base. This ontology knowledge base is the solution of Yunnan spermatophyte endemic genus resource using technical barrier, it laid the foundation to build Yunnan spermatophyte endemic genus area terminology information retrieval system.

1. Introduction
The unique geographical location and climatic condition of Yunnan province, which enriched species resources. It is an important gene pool and is known as “plant kingdom”. According to the statistic, Yunnan has more than 17 thousand advanced plants, accounting for 62.9% in China. Among the 10 thousand kinds of seed plant, there are 151 plants rare and endangered which are listed as protected by China, accounting for 42.6% of the total. Among them, there are more than 190 species belonging to 130 genera of seed plants, accounting for 48.3% of the unique genus(269 genera) of seed plants in China, which is the region with the highest abundance of seed plants in China[1,2].

A lot of research has shown that the distribution of endemic seed plants have very close connection with geography, climate and surroundings, etc[3-8]. However, at present, the distribution patterns and physiological characteristics of endemic seed plants in Yunnan as well as related geographical, climatic and wild habitats have diverse access channels, scattered storage, and inconsistent data standards, unsystematic, poor sharing, and search. These problems greatly restricted the large-scale and systematic research on the unique genera of seed plans in Yunnan.

Ontology knowledge base is knowledge cluster that uses ontology to describe and organize domain knowledge. It can describe concepts of certain field and even a wider range; it can organize, connect, reason and reuse the related knowledge of the certain field. And, it can able to reveal the intrinsic relationship between this knowledge. Also, it is a new method to recognize the digital representation of domain knowledge. Owing to the characteristic of the special species and complex relationship of seed plants, combined with information technology, artificial intelligence and GIS technology, the author can build ontology knowledge base of special species of seed plants in Yunnan, which can realize the complex knowledge structure of digitalize-representation of special species in Yunnan. This paper provided a new way of study and a new theoretical basis for special species in Yunnan.
This paper introduces spatial data fusion technology and multi-level ontology integration technology based on geographic ontology to construct a unique genre ontology library of Yunnan seed plants. The ontology knowledge base to Yunnan seed plants is divided into two layers: the upper layer is a general ontology layer. Setting the geographic ontology as the core, the climate ontology, time ontology, geographical ontology, and ecological ontology such as water quality and soil are established respectively. The lower layer is the main body layer of the unique genus of Yunnan seed plants. The domain ontology layer on the base of domain terminology is composed of domain core concepts and relation sets and core information ontology resource base.

2. Endemic genera of seed plants in Yunnan
The basic units of plant classification are the boundaries, gates, classes, orders, families, genera and species. Plants are divided into two classes, one is “lower plants”, the other is “higher plants”. The “higher plants” are divided into four phyla: angiosperm, gymnosperm, fern and bryophyte. Seed plants refer to the genera of angiosperm and gymnosperm. Endemic genera of seed plants are the greatly important part of flora, accounting for 8.9% of endemic genera of seed plants in China. Referring to a large number of references, the endemic genera of seed plants in Yunnan Province has 59 families, including 125 genera and 246 species of seed plants. As can be seen from the items listed in Table 3-1, the families of endemic genera are listed. From the Table 3-1, Composiae is the most endemic family with 11 genera, the second is Gesnerianceae and umbelliferae contains 10 genera, the third is Gramineae and Labiatae contains 8 genera.

| families       | numbers | generic name | numbers |
|---------------|---------|--------------|---------|
| Compositae    | 11      | Fargesia     | 41      |
| Gesneriaceae  | 10      | Bambusa      | 26      |
| Umbelliferae  | 10      | Sinocarum    | 9       |
| Labiatae      | 8       | Ancylostemon | 6       |
| Gramineae     | 8       | Tremacron    | 4       |
| Cruciferiae   | 5       | Tibetia      | 4       |
| Ranunculaceae | 5       | Pterygiella  | 4       |
| Magnoliaceae  | 4       | Arcuatopterus| 3       |
| Orchidaceae   | 4       | Nannoglottis | 3       |
| Taxodiacea    | 3       | Ypsilandra   | 3       |

3. Domain ontology construction of the endemic genera of seed plants in Yunnan Province
The methods for constructing ontology in different domains are also different, and here is no standard ontology construction method. But, as far as large domain ontology concerned, because of the large amount of data, the simpler conceptual relationship structure is more conducive to the storage of data, such as the gene “GO” project ontology. Therefore, based on the characteristic of the unique genus field of Yunnan seed plants, this paper will use the seven-step method combined with the core concepts and relationship components in the domain ontology to construct the ontology. This not only ensures that the unique ontology structure is simple and practical, but also facilitates the future expansion of the domain ontology. The construction ideas and processes of the endemic genera domain of seed plants in
Yunnan are shown in Fig.1, and the specific construction process and actual examples are shown in Fig.2 to Fig.4 and Table 2.

Fig.1 The flow diagram of ontology construction

Fig.2 part of ontology and hierarchy of endemic genera of seed plants in Yunnan Province
Fig. 3 The example diagram of ontology relationship of endemic genera of seed plants in Yunnan Province

| Type         | Attribute relation | Relationship description | Domain of definition    | range   |
|--------------|--------------------|--------------------------|-------------------------|---------|
| Data attribute | shapes             | Display form             | Sheaths, etc            | string  |
|              | numbels            | Description quantity     | Stamens, etc            | string  |
|              | colors             | Display color            | Anthers, etc            | string  |
|              | types              | Description type         | Culm sheath, etc        | string  |
|              | longs              | Length description       | Culms, etc              | string  |

Fig. 4 Part of relationship schematic diagram of the domain ontology of endemic genera of seed plants in Yunnan
4 Conclusion

In view of the problems of many fields, different data formats, poor data integration, complex concepts and relationships among concepts involved in the information resources of endemic genera of seed plants in Yunnan, this paper constructs a multi-layer ontology knowledge base of endemic genera of seed plants in Yunnan based on the similarities and differences of different concepts in different fields of the information resources of endemic genera of seed plants in Yunnan. Through conceptualization and modeling of knowledge in various fields, ontology mapping technology is used to realize mapping and transformation from heterogeneous data sources to global patterns, thus realizing the integration and sharing of all kinds of digital resources unique to seed plants in Yunnan.

Acknowledgments

The findings and the opinions are partially supported by projects of the major science and technology projects of Yunnan province, No.2018Z1001. This work is supported by Engineering and Technology Research Center for Yunnan Agricultural Big Data and Yunnan Agricultural University, China.

References

[1] Feng Jianmeng, Zhu Youyong. On the genera of seed plants endemic to China in Yunnan[J]. Ecology and Environment, 2010,19(3): 621-62.

[2] Wang Hesheng, Zhang Yili. THE BIO-DIVERSITY AND CHARACTERS OF SPERMATOPHYTIC GENERA ENDEMIC TO CHINA[J]. Acta Botanica Yunnanica, 1994,16(3):209-220.

[3] ZHU H,MA Y, YAN L,et al. The relationship between geography and climate in the generic-level patterns of Chinese seed plants[J].Acta Phytotaxonomica Sinica, 2007, 45(2):134-166.

[4] Ying-Xiong Qiu, Cheng-Xing Fu, Hans Peter Comes. Plant molecular phylogeography in China and adjacent regions: Tracing the genetic imprints of Quaternary climate and environmental change in the world’s most diverse temperate flora[J]. Molecular Phylogenetics and Evolution, 2011,7: No.of Pages 21.

[5] Ma Wenhong, Yang Yuanhe, He Jinsheng. Biomass of temperate grassland in Inner Mongolia and its relationship with environmental factors[J].Science in China(Series C) 2008,38(1):84-92.

[6] Feng Jianmeng, Xv Dongcheng. The large-scale distribution pattern of seed plant species richness in China and its relationship with geographical factors [J]. Ecology and Environmental Science 2009,18(1):249-254.

[7] Song Qing. Application research of intelligent retrieval technology based on domain ontology[D]. Chinese Academy of Agricultural Sciences, 2011.

[8] Liu Wei. Spatial data service discovery and integration based on geographical ontology[D]. Doctoral dissertation of China university of mining and technology, 2010.

[9] Yang Jie. Ontology based knowledge modeling and reasoning of citrus pests and diseases[D]. Central China Normal University.2014.