Constructing Motif-index of China Mythologies Database design, implementation and potential applications

4 authors, including:

Cuixiao Guo
Chinese Academy of Social Sciences

3 PUBLICATIONS   **1** CITATION

SEE PROFILE

Qubumo Bamo
Chinese Academy of Social Sciences

10 PUBLICATIONS   **26** CITATIONS

SEE PROFILE

Constructing Motif-index of China Mythologies Database

Design, Implementation and Potential Applications

Guo Cuixiao, Wang Xianzhao, Bamo Qubumo
Institute of Ethnic Literature
Chinese Academy of Social Sciences
Beijing, China
guocx@cass.org.cn, wangxzh@cass.org.cn,
silver@cass.org.cn

Li Gang
Beijing Zhongyan Technology Co., Ltd.
Beijing, China
ligang@zhongyan.org

Abstract—**Mythology is a form of oral and intangible cultural heritage, carrying rich historical and cultural information. The 56 different ethnic groups in China have numerous myths, both diverse in content and handed down in various forms. The present paper will focus on the design, implementation and application of a mythological motif-index database, known as Wang's Motif-index of China Mythologies Database. There are 33,469 mythological motifs stored in the database, extracted from 12,600 myths from different ethnic groups in China, and divided into ten groups. The paper will give a detailed introduction to the process of motif extraction and database construction. It will also show the interface and search mechanisms provided by the database. With the open-access online, the thematic database is not only applicable to mythology researchers, but also mythology enthusiasts.**

Index Terms—**mythology, motif, index, database, ethnic groups in China**

I. INTRODUCTION

Mythology is an important early human art form, as well as a form of oral and intangible cultural heritage, carrying rich historical and cultural information.

A "motif" is the basic unit of which a mythology is composed, and motifs appear repeatedly in mythologies from different time periods, regions and cultures. The Deluge myth motif, for example, appears extensively within many cultures throughout the world. Motifs can exist independently in the process of handing down, and can also be represented or combined within other popular narrative genres such as epics, legends, folktales, and ballads. The "motif" is favored by scholars due to its representativeness and universality and has become a convenient tool, even viewed as "the best analysis unit" [1]for mythology and other related disciplines.

The 56 different ethnic groups in China have numerous mythologies, both rich in content and handed down in various forms. Historically, literatures of the Han people have enjoyed a high degree of attention, while the collection and research of mythologies from ethnic minorities has been poorer. Most of these ethnic groups have not developed their own written languages and their mythologies can only be passed on orally.

Even ethnic groups with written language forms often pass on their mythologies orally. Such oral mythologies are known as "living myths." Handing them down is an unstable process and this makes them extremely precious.

Important work was needed to provide the comprehensive sorting, archiving and preserving of mythologies from ethnic groups in China, to present a panorama of the handing down of these myths and to promote the establishment of the discipline of mythology and the reconsideration of ethnic culture. To this end, Dr. Wang Xianzhao, a Senior Research Fellow at the Institute of Ethnic Literature of the Chinese Academy of Social Sciences (CASS), has spent close to 30 years reading extensively into the various mythologies of different ethnic groups in China, going into ethnic minority regions and collecting living myths, as well as segmenting and statistically analyzing the data. He has applied motif methodology used in international mainstream folk literature research to 12,600 myths, for which he devised ten mythological motif groups, each with three sub-groups, and categorized into 33,469 mythology motifs and codes. He used graphs to show the comprehensive and logical relationship between motif groups and sub-groups, all of which he lays out in his book *Wang's Catalogue: Motif in China's Mythology (Wmotif)* [2]. This is a reference book and the first mythological book, containing mythological motifs for various ethnic groups in China, including some ancient ethnic groups.

Since this book was published, the Center for Studies in China Ethnic Culture and Language, which is affiliated to Chinese Academy of Social Sciences, has set up Digitizing Documentation Working Team for "Wang's Motif-index of China Mythologies Digital Archiving Project" to further promote the research of folkloristics, folk literature and ethnic literature, as well as related parallel disciplines and to facilitate specialized academic discussion, dialogue and to disseminate academic results. Dr. Wang Xianzhao heads the team, assistant researcher Guo Cuixiao is responsible for researching database and website requirements, program design, customizing metadata standards, data entry and related work, and Senior Research Fellow, Dr. Bamo Qubumo acts as academic advisor.

With technical support from the Beijing Zhongyan Technology Co., Ltd., this research group launched the project "Wang's Motif-index of China Mythologies Database" (WMCMDB) in March 2014. By August, the research group had completed the design plan and started inputting test data, improving functionality and adjusting the webpage structure. By November 2014, the database was online.

WMCMDB systematically organizes extensive numbers of mythological texts through motif-index. It allows multi-dimensional searches and is available via open access on the internet, for the convenience of researchers. The database enables the systematic display of common and individual features for each ethnic group's mythologies and examines how these mythologies have spread and changed. It aids the macro research and comparative analysis of China mythologies, and promotes dialogue and communication about China mythologies and other world mythologies. According to scholars, with this indexing system "constructing a motif-index of mythologies database in the age of modern information technology...in terms of Chinese mythological motif research, is a technical revolution. [3]

WMCMDB is not limited to mythology researchers, it is also available for mythology enthusiasts. It can be used for the research of mythology, motif, and literature, as well as fields such as ethnology, study of religion, folkloristics and communications, providing the researcher with a means to look for patterns and carry out in-depth research and analysis.

II. DESIGN PRINCIPLES AND CHARACTERISTICS

The *Wmotif* and WMCMDB were created in line with the following design principles, and have the following corresponding characteristics:

A. Professionalism

Created by mythologist and database experts, WMCMDB contains reliable data sources and provides high quality data to professional researchers while also acting as a specialized academic exchange platform.

B. Systematicity

The mythology motif-index is essentially a coding system based on ten designated motif groups, and seeks to categorize these motifs groups in terms of form or content.

C. Interoperability

All motifs in WMCMDB are compared with Stith Thompson's *Motif-index of Folk-literature* [4], which is taken as the international standard. If a motif appears in Thompson, the Thompson motif code is added to the reference column of the according motif. This allows researchers to use data comparatively, and facilitates the linking and sharing with other motif-index systems that use Thompson's motif-index.

D. Usability

WMCMDB and its website have been designed with the target users in mind, focusing on user experience, and providing a wide range of browsing and retrieval methods to allow users to find the information they require as quickly as possible.

E. Expandability

The motif-index system and metadata system can be expanded, leaving room for more motifs and refined metadata elements in the future. WMCMDB is open to users, who can add new motifs where appropriate, according to their experience and judgment, thereby enriching the number of Chinese mythological motifs and creating a more rational motif structure.

III. DATA SOURCE AND CONSTRUCTION PROCESS

The data from WMCMDB comes from the book *Wmotif*, and the motifs indexed in the book *Wmotif* come from four sources: (1) relevant domestic and international publications; (2) unpublished yet authoritative publications; (3) academic journals; (4) material collected through personal fieldwork.

Embarking on the research of indexing mythological motifs, Dr. Wang Xianzhao's aim has been to provide extensive mythological motifs from ethnic groups in China to a large number of users through a computer and online retrieval system, also enabling interactive dialogue between these users. The processes of researching mythological motifs and database construction have therefore been integrated. The whole process can be roughly divided into the following intertwined phases: (1) Collecting myths or mythology related texts; (2) Transforming the mythological texts into electronic texts; (3) Extracting the core or basic motif from a large selection of mythologies; (4) Using statistics, calculus, topology and other methodologies to predict the order of motifs, developing various motif groups and sub-groups, and constantly readjust the balance between different groups; (5) Having established a certain number of motif sources, translating Stith Thompson's six volumes of *Motif-index of Folk-literature*[4] and comparing them with each motif, look for missing parts, adjusting or correcting motif groups and motif descriptions; (6) Using an Excel worksheet to create a natural order of motifs within each group, further revising and adjusting the order, making cross-category adjustments for each motif sub-group and deleting repeated motif codes; (7) Improving the layout and description of motif groups, improving the standardization of motif groups and the scientific presentation of motifs; (8) Customizing the metadata elements from WMCMDB, amending the Excel worksheet outlined in step (6) in line with these metadata elements, and converting it into a standardized form to import into the database; (9) Developing the database and website, entering data from the Excel worksheet and making it available online.

Step (3) was particularly important for the construction of the WMCMDB. When extracting motifs from the mythological texts, it is important to retain the "context" of the motif, namely the mythological motif source name, the narrator, the narrator's ethnic group, the collector, the translator, the date of creation, the circulation area, the language, the name of the publication, the publisher, the date of publication, the page number on which the motif appears, and so on. This information is important in providing value judgment and

verifying vitality for the motif, allows the motif to become structural data, and providing basic metadata information for the database.

IV. TECHNOLOGY ARCHITECTURE

WMCMDB and the website content management system is provided by Beijing Zhongyan Technology Co., Ltd. The framework of the technology structure is as follows: The system is based on B/S structure. (B/S = Browser / Server)；Server Operation System: Debian Linux 6.0.10; Web Server Application: Apache/2.2.16 (Debian) mod_fastcgi/2.4.6; Programming Language: PHP 5.3.3; Database: MySQL 5.1.73-1+deb6u1; User Interface Structure: HTML + CSS + Javascript + PHP.

V. OUTCOMES AT THIS STAGE

A. The WMCMDB and website have been constructed and data is gradually being entered

The WMCMDB categorizes motifs into ten groups. There are 33,469 motifs in total and these motifs are divided into three levels. Specific statistics for motif numbers are shown in Table I:

TABLE I. STATISTICS FOR THE TEN GROUPS AND THREE LEVELS OF MOTIFS IN WMCMDB

Group of Motifs			Quantity of Motifs			
Code	Name	Coding Range	Motifs at the first level	Motifs at the second level	Motifs at the third level	Total
W0	God and god-like figures	W0-W999	566	1989	2142	4687
W1	World and Natural Objects	W1000-W1999	398	1603	2606	4607
W2	Human and Man-Kind	W2000-W2999	421	1488	1448	3357
W3	Animals and Plants	W3000-W3999	510	1180	2281	4681
W4	Natural Phenomena and Natural Order	W4000-W4999	290	1010	1179	2479
W5	Social Organization and Social Order	W5000-W5999	244	877	1111	2232
W6	Tangible culture and Intangible culture	W6000-W6999	443	1484	1439	3366
W7	Marriage and Sex	W7000-W7999	347	956	1008	2311
W8	Disaster and War	W8000-W8999	376	1143	1136	2655
W9	Other Motifs	W9000-W9999	403	1393	1298	3094
Total			3998	13823	15648	33469

The development of WMCMDB is now complete and data is gradually being uploaded and updated, with close to 10,000 motifs so far. A specialized WMCMDB website has also been developed and is accessible at: http://myth.ethnicliterature.org/

B. Implementing a variety of methods for motif retrieval

WMCMDB motifs are displayed in the "Motif chart", which includes codes, motif descriptions and references. References include Thompson motif codes, the ethnic origin of the myth from which the motif originates, related motifs and examples, and so on.

Users can browse or search for motifs in various ways, including:

1) Search according to motif group and sub-groups.

2) Finding a motif directly from its order in the "Motif chart." This method allows the user to see the information from all motifs at three levels as well as the natural structure of the motif-index.

3) Search via related motifs. The motif code table contains related motifs. By clicking onto the related motif, it is possible to expand its narrative elements or structure to provide a multi-faceted understanding.

4) Using a key word search to find a precise motif. It is possible to define a motif level and use key word searches including motif codes to find motifs. Searches can also be carried out using Thompson codes or by ethnic group. For more precise searches, a key word can be added as well. For example, to find all "gods and spirits" related Tibetan mythological motifs at the first level, the key word shown in Fig. 1 are filled in to obtain the following results:

Fig. 1. Defining Motif level and adding key word to search for a motif

The webpage shows users the total number of motifs in the current database, the number of motifs in each level, the ten ethnic minorities with most mythologies and total number of motifs for each. These figures can be clicked on, and the corresponding results will be displayed. Clicking on the ethnic group names under references on the motif code table will also provide quick retrieval of all mythological motifs of that ethnic group.

VI. POTENTIAL TYPICAL APPLICATIONS

The essence of the motif index lies in its practical application within mythology research. As well as providing accurate searches for relevant motifs within mythologies, it also has many other applications.

A. The application of the mythology motif-index within the specific comparison of mythologies.

The most direct way of using the motif-index in mythology research is for quantitative measurements and comparisons on the same type of texts from different ethnic groups or different regions. Extracting and arranging motifs from each text enables comprehensive analysis on the numbers and changes of a motif, thus discovering common patterns and differences.

B. The application of the mythology motif index within the structural analysis of mythologies.

Many motifs can be combined into different mythological narrative types, in what is known as the "mythology narrative structure model", which is significant for general analysis. Theoretically, with the help with these models, quantitative and qualitative analysis or comparative studies can be carried out on any mythology. The model divides mythologies into the following combinations: (1) chain narrative structure model; (2) divergent narrative structure model; (3) embedded narrative structure model; (4) parallel narrative structure model; (5) composite narrative structure model; and (6) other forms of narrative structure models. [5]

C. The application of the mythological motif index for combined pattern analysis of mythology types.

The extraction and display of motifs appears to be randomized, but this is actually an expression of the type of structure of narrative literature within mythology. The settings and arrangement of motifs within the WMCMDB will not only provide general rules for motif arrangements, but also determine the motif groups arising from different motif combination.

VII. FUTURE PLANS

By the end of 2015 the project team will have completed all 33,469 mythological motifs entries which will be available on the internet. Using the WMCMDB, the project team plans to add mythological texts, pictures, audios and videos, as well as mythology research books and articles, thus establishing a "Database on Mythologies of China". Embedding motif codes into these mythological texts, pictures, audio and video

materials, and linking them via the WMCMDB will connect all this data, enabling interactive multimedia retrieval, automatically generating extensive research information, creating characteristic data, and ultimately a convenient channel for the comparative study of mythologies in the digital age. Researchers can take the motif-index as a basis to construct new mythology research databases with individual characteristics [6].

VIII. SUMMARY AND EXPERIENCE SHARING

From the first stage of writing the academic reference book, preparations should be made for the construction of the database. Constructing the database alongside the book will facilitate the search for repeated entries and mistakes, speed up the work progress and reduce errors. Ideally, this book should be made directly into an online database, while giving open access to public.

Researchers in the humanities and social sciences should acquire a certain level of computer technology skills, or at least have awareness in this regard and work with database professionals to this end.

Mythological motifs have vitality and can appear repeatedly across time and space within various cultures and cultural products. Mythological motif-index database should therefore be freely expandable, and allowed to develop in line with the ideas of the project team.

REFERENCES

[1] J.-X. Chen, *Mythology: Exploration of Motif Analysis Methods*, Wuhan: Hubei Education Press, 1997, pp. 34-35

[2] X.-Z. Wang, *Wang's Catalogue: Motif in China's Mythology*, Beijing: China Social Science Press, 2013.

[3] S.-H. Hou, "Wang Xianzhao's Mythological Motifs Research and Humanities," *Chinese Social Sciences Today*, Nov.24, 2014 (B01).

[4] Stith Thompson, *Motif-index of Folk-literature*: *A Classification of Narrative Elements in Folktales, Ballads, Myths, Fables, Mediaeval, Romances, Exempla, Fabliaux, Jest-books, and Local Legends* (Vol. 1-6). Bloomington: Indiana University Press, 1989.

[5] X.-Z. Wang, *"*On the Implementation and Application of Motif-index in Mythology Research: Taking *Wang's Catalogue: Motif in China's Mythology* as Example," *Journal of Guizhou Minzu University*, No.1, 2015, pp. 66-72

[6] X.-Z. Wang, "The Need to Accelerate the Database Construction of Mythologies in China,*" China Culture Daily*, Aug. 23, 2011.