Empirical Analysis of Digital Resource Integration for Intangible Cultural Heritage Based on Topic Maps: Postprint

Authors: Shi Yi, Xiong Huixiang, Lu Yingying

Date: 2023-08-26T00:00:00+00:00

Abstract

[Purpose/Significance] Intangible cultural heritage resources hold significant social value. The application of topic map technology to the integration of intangible cultural heritage digital resources can better facilitate their inheritance and preservation. [Method/Process] Based on an analysis of the themes, relationships, and resource indexing of intangible cultural heritage digital resources, this study constructs a topic map for these resources and implements their visualization. [Results/Conclusion] The findings demonstrate that topic map technology can provide semantic-based integration for intangible cultural heritage digital resources, visually display the relationships between topics to users, and offer visual navigation.

Full Text

Preamble

Vol. 62 No. 7 April 2018
ChinaXiv Cooperative Journal

An Empirical Analysis of Digital Resource Integration for Intangible Cultural Heritage Based on Topic Maps

Shi Yi[1], Xiong Huixiang[2], Lu Yingying[2]
[1] Wuhan University Library, Wuhan 430072
[2] School of Information Management, Central China Normal University, Wuhan 430079

Abstract

[Purpose/Significance] Intangible cultural heritage (ICH) resources hold significant social value. Applying topic map technology to the integration of

ICH digital resources can better facilitate their preservation and inheritance. [**Method/Process**] This study analyzes the themes, relationships, and resource indexing of ICH digital resources, constructs a topic map for ICH digital resources, and implements its visualization. [**Result/Conclusion**] The findings demonstrate that topic map technology can provide semantic-based integration for ICH digital resources, visually present inter-topic associations to users, and offer visual navigation.

Keywords: topic map; intangible cultural heritage; digital resources; visualization
Classification Number: G254
DOI: 10.13266/j.issn.0252-3116.2018.07.012

Both intangible cultural heritage (hereinafter referred to as "ICH") and tangible cultural heritage represent great crystallizations of human social civilization with significant social value. Unlike the visibility of tangible cultural heritage, ICH is expressed through oral traditions, physical movements, behaviors, and skills, possessing intangible and living characteristics [1]. From its creation to its development, ICH does not exist in isolation within the spiritual world of individuals but is closely connected to its ecological space and cultural environment, maintaining its vitality only when it adapts to these contexts [2]. This presents greater challenges for ICH protection and inheritance. Currently, scholars both domestically and internationally have begun exploring the semantic organization of ICH using knowledge organization technologies such as metadata, ontologies, topic maps, and linked data, including the establishment and comparison of ICH metadata standards, ICH ontology construction, ICH topic map development, and the building of ICH resource aggregation platforms based on linked data. However, among these various methods for organizing ICH resources, research on the application of topic maps to ICH remains relatively limited. A search of CNKI, Wanfang, and Web of Science databases using the keywords "intangible cultural heritage," "ICH," and "topic maps" yielded only two articles examining the combination of topic maps and ICH [3-4]. Furthermore, while certain research exists on the semantic organization of ICH resources, several issues persist. For instance, metadata focuses on describing thematic concepts but struggles to represent relationships between topics; ontologies emphasize describing relationships between concepts but cannot describe relationships between topics and resource entities, and their construction is relatively complex; linked data focuses on describing relationships between topics and resource entities but requires all resources to be converted into a unified data format as a prerequisite. Topic maps, as a new digital knowledge organization technology that incorporates ideas from the Semantic Web and ontologies, can create rich semantic models, browse and retrieve various types of information, integrate heterogeneous data resources, and offer strong flexibility with functions such as knowledge navigation and knowledge positioning. Therefore, this study applies topic map technology to the organization of ICH digital resources, constructs a topic map for ICH digital resources, and conducts an empirical study using "Huangmei Opera," which originated in Hubei and Anhui provinces. The re-

search finds that applying topic maps to ICH digital resource integration not only more fully realizes the utilization value of ICH and facilitates its preservation and inheritance but also visually presents information such as historical origins, inheritance lineages, and school development in an intuitive manner, providing users with more rational and effective pathways for resource utilization.

2 Overview of Topic Maps and ICH Digital Resource Integration

2.1 Topic Map Technology

ISO/IEC 13250 introduces topic maps as follows: a topic map defines a multi-dimensional space containing numerous nodes, each representing a topic. The distance between topics can be calculated by the number of intermediate topics that must be traversed from one topic to another, and these intermediate topics constitute the path for establishing relationships between the two topics [5]. Topic maps have three fundamental elements: topics, associations, and resource guidance [6]. "Topics" can be anything—whether existing or not—and a topic may have one or more base names as well as alternative names. "Associations" represent relationships between two or more topics and can be categorized into different types, i.e., association types. A topic can connect to one or multiple information resources related to it, and these connected information resources are called resource guidance (also referred to as "resource entities" or "resource sources" by some scholars). Resources within the topic map are called internal resource guidance, while resources linked via Uniform Resource Identifiers (URIs) are called external resource guidance.

The conceptual model of topic maps comprises two layers: the topic map layer and the information resource layer. The topic map layer consists of associations between various topics, as shown in Figure 1 [Figure 1: see original paper]. The circles in Figure 1 represent topics, and the connecting lines represent associations between topics. The information resource layer contains various information resources, typically referring to digital information resources such as articles, web pages, and videos. These resources connect to topics in the topic map layer while independently existing on the Internet.

2.2 ICH Digital Resource Integration

UNESCO's Convention for the Safeguarding of Intangible Cultural Heritage defines ICH as the practices, representations, expressions, knowledge, and skills that communities, groups, and individuals recognize as part of their cultural heritage, including instruments, objects, artifacts, and cultural spaces associated with them. ICH digital resources refer to ICH resources published, accessed, and utilized in digital form through computer technology, communication technology, and multimedia technology, such as ICH audio resources, video resources, and databases. ICH digital resource integration involves representing and order-

ing the internal and external characteristics of ICH digital resources according to ICH classification systems, thereby transforming them from disorder to order and facilitating scientific and systematic integration. ICH digital resources are not isolated entities; they represent unique cultural characteristics of a region and must be promoted by cultural inheritors. Thus, ICH is intimately connected to its ecological space and cultural environment. To comprehensively display ICH digital resources and their interrelationships, semantic organization is required—extracting themes from ICH digital resources according to ICH classification systems, determining relationships between themes, and standardizing descriptions of relationships among various ICH digital resources to form a semantic-based ICH digital resource integration model.

3 A Topic Map-Based Integration Model for ICH Digital Resources

3.1 Framework for Topic Map-Based ICH Digital Resource Integration

Based on the fundamental principles of topic maps, a framework for ICH digital resource integration can be constructed, as shown in Figure 2 [Figure 2: see original paper]. The framework primarily consists of four components: the resource layer, generation layer, application layer, and user layer [7].

3.1.1 Resource Layer

ICH-related resources form the foundation for comprehensively displaying the semantic structure of ICH digital resources. Located at the bottommost level of the topic map-based ICH digital resource integration framework, the resource layer serves as the starting point for the entire integration process, providing substantial heterogeneous information resources related to ICH for topic map construction, including digital resources, personnel resources, institutional resources, knowledge resources, and project resources.

3.1.2 Generation Layer

The topic map generation layer extracts thematic concepts from collected and organized ICH digital resources, classifies topic types and associations, and defines resource guidance. It then uses the Ontopoly component to edit the ontology and instances of the topic map, performs semantic annotation on the ICH digital resource topic map ontology, verifies the correctness of the topic map's representation of ICH digital resources, analyzes the association relationships, merges topic maps with identical associations, and finally stores the validated topic maps in the ICH digital resource knowledge base.

(1) **Define Topic Types.** ICH digital resources involve a wide range and diverse categories, necessitating thematic classification. From a hierarchical perspective, this study divides ICH digital resources into six first-level topics: "Ethnicity," "Region," "ICH Project," "Level," "Inheritor," and "Applicant," with sub-topics further divided according to each topic's

specific circumstances. Among these, "Ethnicity" represents the ethnic characteristics of ICH projects; "Region" indicates geographical divisions; "ICH Project" identifies the name of the ICH project; "Level" reflects the importance of the ICH; "Inheritor" concerns the inheritance history and main inheritors of the ICH; and "Applicant" is the entity that applied for the ICH project designation. ICH digital resources encompass numerous types with rich content, and significant differences exist between different domains. Constructing a complete ICH digital resource topic map requires substantial effort and cost. Therefore, this study selects "Huangmei Opera," a national-level ICH project from Hubei Province, to construct a topic map for Huangmei Opera digital resources. As one type of ICH project, Huangmei Opera has its own second-level, third-level, and further sub-topics based on the first-level topics.

(2) **Define Association Types.** After defining topic types, associations between topics must be defined. These associations link different topics to form a knowledge network. This study defines six association types corresponding to the six topic types: "Ethnicity and ICH," "ICH Project, Applicant, and Level," "Master-Apprentice," "Inheritance," "Located In," and "Belongs To." Specifically, "Ethnicity and ICH" represents the relationship between ethnicity and ICH projects; "ICH Project, Applicant, and Level" denotes the complex many-to-many relationship among ICH projects, applicants, and levels; "Master-Apprentice" is the transmission relationship between inheritors; "Inheritance" is the relationship between inheritors and ICH projects; "Located In" connects ICH projects with regions; and "Belongs To" represents the relationship between first-level topics and sub-topics. Second-level topic associations primarily include "Contains/Belongs To," "Synonymous," "Collaborative," "Performative," "Master-Apprentice," and "Creative" relationships.

(3) **Resource Guidance.** Resource guidance primarily involves referencing and locating resources for topics. Based on different resource content types, resources are categorized into four types on the foundation of ICH-related resources: ICH digital resources, institutional resources, personnel resources, and project resources. Among these, digital resources are further divided by format into text, images, audio, video, and models.

3.1.3 Application Layer

The application layer provides services such as retrieval, reorganization, and querying of ICH digital resource knowledge content according to user needs during the knowledge service process. It primarily offers search, navigation, recommendation, and maintenance functions. The search function allows users to locate topics of interest and their corresponding resource guidance through topic map query languages. The navigation function provides navigational assistance through associations between topics in the topic map, enabling users to conduct expanded retrieval. The recommendation function leverages the semantic association characteristics of topic maps to automatically recommend

resources related to the user's search topic, improving recall rates. The maintenance function includes basic operations of addition, deletion, and updating for the three fundamental elements—topics, topic types, and associations—allowing continuous improvement and maintenance of the ICH topic map.

3.1.4 User Layer

The user layer, situated at the topmost level of the ICH digital resource topic map architecture, provides functions such as topic map browsing and retrieval. Users can utilize the Omnigator component of the topic map to browse topic knowledge pages and adjust various modules through the user interface.

3.2 Advantages of Topic Map-Based ICH Digital Resource Integration

Topic maps have been termed the "GPS of the information age" [8] and represent a technology capable of achieving knowledge management, offering unique advantages in ICH digital resource integration. These advantages manifest in several aspects:

First, **semantic description.** Topic maps originate from traditional query assistance tools such as back-of-book indexes and thesauri, incorporating the strengths of traditional indexing and scientific knowledge representation from libraries to describe complex knowledge structures [9]. Topics and associations in topic maps allow for flexible definition of complex semantic relationships, providing a semantic-based interface for organizing ICH digital resources and enabling semantic-based query and retrieval.

Second, **integration of distributed resources.** Topic maps utilize resource guidance to establish links between ICH digital resources and topics. Through this mapping relationship, they can index widely distributed heterogeneous ICH digital resources, effectively addressing the current information silo phenomenon in ICH digital resources.

Third, **knowledge organization.** Topic maps classify knowledge based on topics and structure the knowledge system of ICH digital resources through inter-topic relationships, allowing organization at the knowledge level and establishing an ICH digital resource knowledge base.

Fourth, **result visualization.** Topic maps provide visual presentation methods, with graphical representations enabling users to understand topics and their relationships in ICH digital resources at a glance. Topic maps can display both overall effects for a comprehensive understanding and local effects by adjusting control parameters. Additionally, when users search for specific keywords, the retrieved keywords change color for easy clicking.

4 Empirical Analysis of Topic Map-Based ICH Digital Resource Integration

ICH resources encompass numerous types with rich content, and significant differences exist between domains. Therefore, constructing a complete ICH digital resource topic map requires substantial effort and cost. This study selects the national-level ICH project "Huangmei Opera" from Hubei Province as a case study to construct a topic map for Huangmei Opera digital resources.

Huangmei Opera, one of China's five major opera genres, has far-reaching influence. Originating in Huangmei County, Hubei Province, it was formerly known as Huangmei tune or tea-picking opera and is now widely popular in Anqing City, Anhui Province, and Huangmei County, Hubei Province. Initially, Huangmei Opera evolved from tea-picking tunes from late Qing Dynasty Hubei that spread to neighboring Anhui regions, where they incorporated local folk arts and were performed using the Anqing dialect, gradually developing into a distinct opera genre then known as Huai tune or Huai melody. Subsequently, Huangmei Opera borrowed elements from Qingyang tune and Hui melody to perform "full-length plays." Over a century centered in Anqing City, Huangmei Opera gradually evolved into Anhui's primary local opera genre and a nationally renowned major opera form [10].

Currently, the dissemination scope of Huangmei Opera is diminishing, and survival challenges face Huangmei Opera troupes at all levels, particularly county-level troupes. Consequently, urgent protection is needed for this ICH resource.

4.1 Theme Analysis

The Huangmei Opera domain contains numerous terms and vocabulary. This study draws thematic vocabulary from four primary sources: (1) existing thesauri for organizing professional terms in the Huangmei Opera domain, such as the Chinese Library Classification, Chinese Thesaurus, and Chinese Classified Thesaurus; (2) authoritative literature on Huangmei Opera, such as the *Chinese Opera Annals—Anhui Volume* [11]; (3) extraction of keywords and subject terms from Huangmei Opera-related literature; and (4) relevant online resources such as authoritative websites like China Opera Network and Baidu Baike. Through collection and analysis of Huangmei Opera-related thematic terms, a hierarchical classification of Huangmei Opera themes was developed, as shown in Table 1 .

4.2 Theme Association Analysis

Based on the above theme analysis, associations between themes can be identified. This study defines the following association types for Huangmei Opera themes: "Contains/Belongs To," "Synonymous," "Collaborative," "Performative," "Master-Apprentice," and "Creative."

Contains/Belongs To Association. This represents hierarchical relation-

ships between Huangmei Opera themes, existing between topic types and their sub-topics or between sub-topics at different levels. Examples include the "Repertoire" topic type containing "Aria" topics, different "Repertoires" containing their respective "Role Types," and "Aria" topics belonging to different "Music Styles."

Synonymous Association. Some Huangmei Opera themes have synonymous relationships, such as between "Mournful" and "Grief-stricken," "Melancholic" and "Depressed," or "Joyful" and "Lively" in Huangmei Opera music styles.

Collaborative Association. Different performers in the same repertoire, as well as directors, playwrights, and actors in the same performance, share collaborative relationships.

Performative Association. Performers have performative relationships with their works and role types, enabling the construction of performative associations among "Performing Artists," "Repertoires," "Films," "TV Dramas," "Stage Plays," and "Role Types."

Master-Apprentice Association. Huangmei Opera inheritance follows traditional master-apprentice learning methods, primarily through "master leading apprentice" transmission. Consequently, many Huangmei Opera artists have master-apprentice relationships; for example, Yan Yungui was the only master for whom Yan Fengying formally apprenticed, while Yan Fengying herself took disciples such as Tian Yulian, Qin Qixia, and Wang Fengzhi.

Creative Association. Directors, playwrights, and composers have creative relationships with their papers, articles, books, films, and music.

4.3 Resource Guidance Analysis

This study's resource links for Huangmei Opera themes primarily derive from the China Intangible Cultural Heritage Network, China Quyi Network, CNKI, Wanfang Database, ICH Inventory Database, and search engines. Describing topic attributes and indexing related resources helps users better understand topic meanings and connotations. Topic maps can use resource guidance to indicate different topic attributes and direct users to locations of relevant digital or other resource formats. Following the previous classification of Huangmei Opera themes, Table 2 describes attributes and resource types for selected topic types.

4.4 Implementation of Topic Map-Based ICH Digital Resource Integration

Topic map development tools play a crucial role in topic map development, providing functions for creation, storage, display, and maintenance. These tools typically exist as topic map engines or TMAPI (Topic Map API). OKS, TM4J, tinyTIM, and XTM4XMLDB are all TMAPI-compliant topic map engines. This study primarily introduces the OKS knowledge development suite, which is also

the development tool used in this research. OKS (Ontopia Knowledge Suite), developed by Ontopia, supports complete topic map construction, is fully Java-based, and can provide topic map functionality for any application with stability, efficiency, and advancement [12]. The OKS architecture comprises three layers: storage layer, engine layer, and application layer (from bottom to top). The storage layer supports both relational database backend (RDBMS backend) and in-memory storage technologies; the engine layer includes the topic map engine and full-text query engine; the application layer primarily includes the topic map Navigator, topic map browser Omnigator, and topic map visualization tool Vizigator.

After constructing the Huangmei Opera topic repository, Ontopoly in OKS can be used to edit the Huangmei Opera digital resource topic map, Omnigator to browse it, and Vizigator to visualize it. Ontopoly contains four modules: Description, Export, Instance, and Ontology. Description provides overall description of the topic map, requiring explanation of the construction purpose, significance, and resource sources; Export enables topic map output, allowing the constructed topic map to be saved locally in formats such as XTM or RDF according to requirements; Ontology provides functions for defining topic types, resource types, association types, role types, and name types; and the Instance module defines instances of topic types.

4.4.1 Generation and Browsing of Huangmei Opera Digital Resource Topic Map Ontopoly provides functions for defining topic types, resource types, association types, role types, and name types. Based on the above analysis, Huangmei Opera digital resource information was sequentially input into the Ontopoly tool.

Omnigator is a free tool in OKS for browsing topic map navigation frameworks, outputting topic maps in HTML format. In the Omnigator interface, users can directly view association relationships between topics and resource guidance. Figures 3 [Figure 3: see original paper] and 4 [Figure 4: see original paper] respectively show browsing results for the "Huangmei Opera" topic map and the "Yan Fengying" topic type using Omnigator.

4.4.2 Visualization of Huangmei Opera Digital Resource Topic Map Navigator is a visualization component in OKS for browsing constructed topic maps. Topic map visualization displays the semantic network structure formed by inter-topic relationships [13]. Figures 5 [Figure 5: see original paper] and 6 [Figure 6: see original paper] respectively show visualization results for the "Yan Fengying" topic instance and the "Female Imperial Son-in-Law" film topic instance, where different colors represent different association types. Users can select topics of interest during browsing to track additional knowledge and obtain a good interactive experience.

4.4.3 Analysis of Huangmei Opera Digital Resource Topic Map Results Through topic map visualization, users can browse knowledge maps of topics in intuitive graphical formats. Using the visualization results for "Yan Fengying" as an example, the interface displays that "Yan Fengying" is an instance of "Performing Artist," shows the repertoires, films, and TV dramas performed by Yan Fengying, and reveals master-apprentice and collaborative relationships between Yan Fengying and other artists.

This analysis demonstrates that constructing ICH digital resource topic maps can improve current ICH digital resource organization in China in several ways:

First, China's ICH digital resources currently remain at the stage of digital preservation and information integration, without addressing semantic relationships between resources. The constructed ICH digital resource topic map provides semantic-based navigation, presenting all topics related to a given topic and their semantic relationships when users browse, enabling convenient discovery of other relevant topics and understanding of inter-topic relationships.

Second, China has vast quantities of ICH resources from diverse sources, making it difficult for users to locate needed resources. In topic maps, since topics and resources are separated, users need not consider resource locations but can directly link to relevant text, image, audio, and video resources through external resource guidance URIs.

Third, using visualization plugins provided by topic maps, users can browse visualization results for all topics in the topic map, with relationships between different topics displayed intuitively.

This study used the OKS Samplers' Ontopoly function to construct a Huangmei Opera topic map, employed Omnigator to browse relevant topics, and implemented Huangmei Opera topic map visualization through the Vizigator tool, thereby validating the feasibility of applying topic map technology to ICH digital resources. Combining topic maps with ICH digital resources offers dual benefits: on one hand, topics and associations in topic maps allow flexible definition of complex semantic relationships, providing a semantic-based interface for ICH digital resource integration that enables semantic query and retrieval, while resource guidance functions establish links between ICH digital resources and topics to index widely distributed heterogeneous ICH digital resources and effectively address information silo problems; on the other hand, topic maps separate the resource domain from the topic domain, allowing the same topic to point to different ICH digital resources and different topics to point to the same ICH digital resources, enabling users to view perspectives from different angles to meet diverse utilization needs.

However, the combination of ICH digital resources and topic maps remains a new research field. Current research primarily focuses on small-scale data studies and theoretical research stages. Future research should implement topic map-based ICH digital resource information services on large-scale data foundations, such

as dynamic navigation and personalized information retrieval for ICH digital resources.

References

[1] Wang Yugui. The Living Characteristics of Intangible Cultural Heritage and Its Protection [EB/OL]. [2016-12-07]. http://www.mzb.com.cn/html/report/141120591-1.htm.
[2] Wang Jushan. Reconsideration of the Characteristics and Protection of Intangible Cultural Heritage [J]. Social Sciences Journal, 2006(5): 165-167.
[3] Wang Meng, Xu Xin. Application of Topic Map Technology in Intangible Cultural Heritage Information Resource Organization [J]. Library and Information Service, 2015, 59(14): 15-21.
[4] Hu Jian. Construction and Implementation of a Virtual Cultural Corridor for Tujia Music [D]. Wuhan: Central China Normal University, 2016.
[5] ISO/IEC 13250:2003 Information Technology—SGML Applications—Topic Maps [EB/OL]. [2016-12-17]. https://www.iso.org/obp/ui/#iso:std:iso-iec:13250:ed-2:v1:en.
[6] Pepper S. The TAO of Topic Maps [EB/OL]. [2016-12-22]. http://www.ontopia.net/topicmaps/materials/tao
[7] Deng Min. Research on Tag Semantic Mining Based on Topic Maps [D]. Wuhan: Central China Normal University, 2014: 37-44.
[8] Qin Tiehui, Guo Yanji, Sun Lin. The "GPS" of the Information Age—Topic Maps [J]. Journal of Jiangxi Library Science, 2005(1): 1-3, 7.
[9] Zhang Yutao, Xia Lixin. Research on E-government Information Resource Integration Model Based on Topic Maps [J]. Journal of Intelligence, 2009(7): 161-165.
[10] Huangmei Opera [EB/OL]. [2016-12-30]. http://www.ihchina.cn/5/10821.html.
[11] Chinese Opera Annals Editorial Committee. Chinese Opera Annals: Anhui Volume [M]. Beijing: China ISBN Center, 1993: 366-367.
[12] Product White Paper: The Ontopia Knowledge Suite [EB/OL]. [2016-11-18]. http://xml.coverpages.org/Ontopia-O-K-Sv13.pdf.
[13] Ding Y, Jacob E K, Fried M, et al. Upper Tag Ontology (UTO) for Integrating Social Tagging Data [J]. Journal of the American Society for Information Science & Technology, 2010, 61(3): 505-521.

Author Contributions

Shi Yi: Drafted the manuscript.
Xiong Huixiang: Supervised the research and guided the manuscript.
Lu Yingying: Revised and polished the manuscript.

Note: Figure translations are in progress. See original paper for figures.