A semantic framework for digital archives of ancient Chinese villages aimed at humanistic discovery: a case study of four villages in Henan

Tianjiao Qi, Yiduo Wang, Meiman Wong, Tsamchu Basang, and Linglu Li

Abstract

Introduction. Ancient Chinese villages are crucial cultural legacies facing preservation challenges. Current archiving has drawbacks. This research aims to create a semantic framework for digital archives to boost humanistic knowledge discovery.

Method. We examined three common archiving rules, extracted cultural elements, and designed a semantic framework by reusing international ontologies CIDOC-CRM. Then, we applied it to four villages in Henan for dataization.

Analysis. After comparing the archiving rules and integrating cultural elements, we evaluated the framework. We also processed the villages' archives, including data annotation and Protege knowledge base construction.

Results. The framework consists of 6 entities, 29 relations, and 94 attributes. The semantic structures of the four villages' digital archives were presented, confirming the framework's practicality and applicability.

Conclusion(s). The semantic framework offers a feasible solution for ancient village archive dataization and knowledge discovery. Future efforts should center on enhancing knowledge presentation and utilization for better cultural heritage protection.

Introduction

Ancient Chinese villages, typically referring to settlements established before the Republic of China era, are also known as *'traditional villages'* and *'constitute a significant legacy of Chinese agrarian civilization'* (Feng Jicai, 2012). Renowned for their historical depth and abundance, these villages preserve a wealth of historical information and cultural heritage. As President Xi Jinping has emphasized, *'rural civilization is the cornerstone of Chinese civilization, villages are its carriers, and agrarian and scholarly traditions are our soft power'* (Xin Huashe, 2023). However, during China's rapid modernization and urbanization, these ancient villages and their associated traditional cultures have faced critical challenges.

To ensure their preservation, the China Federation of Literary and Art Circles initiated a nationwide survey and archiving of ancient villages in 2003. Since 2012, multiple government departments have jointly launched the Traditional Village Survey and established an expert committee to compile the Traditional Village Directory. By 2023, six batches comprising 8,155 villages of significant conservation value had been included in the directory. Inspired by these national efforts, several provinces, including Anhui and Zhejiang, have completed local *'thousand-village archiving'* projects. These initiatives have yielded substantial progress in documenting ancient villages, laying the groundwork for digital archiving, which has become a focal area in recent years.

> By the end of 2024, the 'Chinese Traditional Village Museum' had launched digital archives for 1,458 village-specific exhibits' (Traditional Chinese Village Digital Museum, 2024), 'while Sichuan Province plans to establish a comprehensive village archive database by 2025' (Luo Zhiyang, 2019).

Similarly, the Taizhou Archives has included digital archiving of traditional villages in its 14th Five-Year Plan.

Motivated by the progress in archiving practices, Chinese academia has conducted extensive studies on the role of ancient village archives in cultural heritage preservation, resource development, cultural discovery, and historical tracing. Scholars generally agree that village archives should encompass tangible cultural heritage, intangible cultural heritage, historical documents, notable historical events, and influential figures. Researchers such as An Hongqingn (An, H.Q.2017, p.29-30), Xu Xinyun (Xu, X.Y.2017, p.43-48), and Ren Yue (Ren, Y.2018, p.79-84) emphasize the importance of documenting historical culture, material forms, and their transformations, while Man Yi (Man, Y. and Wang, P.2017, p.12-15) advocates for the standardization of archiving practices. He Siyuan (He, S.Y.2017, p.49-53) and Zheng Jinxia (Zheng, J.X et al.2023, p.194-198) further explore the institutional mechanisms and roles of archival departments in traditional village conservation. Tang Qi (Tang, Q.2023, p.66-68) argues that archival preservation should go beyond static records to emphasize the 'activation' of traditional villages.

Despite these efforts, significant challenges remain. The vast scale but relatively low perceived value of ancient village archives, coupled with the lack of attention in both practice and theory to these issues, hinder progress. In practice, the diversity of archiving stakeholders and the absence of unified standards lead to severe issues of multi-source heterogeneity, impeding the digitization of archives. Multi-source heterogeneity refers to the complex origins of village archives, which include government departments, cultural institutions, private organizations, and local villagers, resulting in diverse content structures and formats such as text, images, and audiovisual materials. Additionally, differing understandings and expressions of cultural elements among stakeholders further exacerbate semantic discrepancies. These challenges significantly hinder data mining and the effective utilization of village archives. Theoretically, existing studies rarely focus on how the public leverages cultural knowledge in village archives, nor do they propose a systematic

framework or tools to support the structured storage, semantic analysis, and relational mining of unstructured data.

Against this backdrop, this study leverages existing archiving standards and ontology construction techniques commonly used in cultural heritage research. By integrating the internationally recognized CIDOC-CRM semantic framework (FAN, Q. et al.2023, p.469-482), Event Ontology (Raimond, Y., Abdallah, S,2007), and related narrative ontology studies, this research aims to construct a formalized semantic ontology centered on the cultural connotations and core elements of ancient villages. The goal is to explore a feasible pathway for the digitization, knowledge integration, and discovery of village archives. Using archives from four ancient villages in Henan Province as a case study, this study applies the proposed ontology to data digitization, providing a model for the structured archiving and relational mining of ancient village resources. This work seeks to advance the discovery, understanding, dissemination, and preservation of the humanistic value of ancient Chinese villages.

Cultural elements in existing archiving rules for ancient villages

Before constructing a semantic framework for digital archives of ancient villages, this study analyzed common cultural elements in existing archiving rules widely used in China. This analysis provides a foundation for developing and aligning the vocabulary of the proposed ontology. The selected archiving rules include: the Village Information Form for the Digital Museum of Chinese Traditional Villages issued by the Ministry of Housing and Urban-Rural Development of China (Digital Museum of Chinese Traditional Villages,2018), the Text-Based Archiving Form for Traditional Village Surveys and the Image-Based Registration Form for Traditional Village Surveys from the Chinese Traditional Village Network, and the Classification Scheme for Digital Resources of Gaoqian Houku Village developed by the Digital Memory team at Renmin University of China(National Archives Administration of China,2021). These rules represent key approaches at the levels of national registry, grassroots survey, and digital database construction.

Characteristics and differences in existing archiving rules

The Village Information Form for the digital museum of Chinese traditional villages, published in 2018, is a foundational document for digital archiving of ancient villages. Created by the Ministry of Housing and Urban-Rural Development with support from the China Academy of Urban Planning and Design and local administrative departments, it provides a unified template for documenting traditional villages nationwide. This form emphasizes the systematic recording of basic village information and architectural culture, covering geographic location, historical development, architectural characteristics, and cultural heritage. It highlights detailed documentation of village history, migration processes, significant historical figures, and major events. For cultural heritage, it systematically captures data on architectural elements (names, materials, styles, and functional layouts) and includes non-material aspects such as folklore, festivals, rituals, and local dialects, along with information on cultural activities and inheritors of intangible heritage. Notably, the form also includes data on local cuisine, products, and tourism resources, supporting economic development and cultural tourism initiatives.

The Text-Based Archiving Form and Image-Based Registration Form from the Chinese Traditional Village Network were developed as part of the '*Preserving Nostalgia—Traditional Village Survey and Archiving*' project, launched in 2014 with support from the Ministry of Housing and Urban-Rural Development and various cultural organizations. These guidelines focus on comprehensive documentation of cultural and social life, including social structure, population distribution, customs, production methods, natural and cultural heritage, and historical changes, supplemented by extensive visual materials. The rules emphasize combining textual and visual data to document the natural environment, geological conditions, village morphology, architectural relics, and street information, aiming to capture the interconnectedness of human and natural elements. For

cultural heritage, the guidelines prioritize recording items listed in heritage registries, such as historical artifacts, family histories, and manuscript collections, often supported by photographs to vividly depict traditional customs, production activities, and even village inhabitants.

The classification scheme for digital resources of Gaoqian Houku Village, developed by Renmin University of China, was created under a major National Social Science Fund project focusing on the digital preservation and transmission of historical and cultural villages. Starting in 2016 with Gaoqian Village in Xianju County, Zhejiang Province, the project has emphasized the use of digital technologies for resource classification and association. This framework is technically oriented, facilitating the digitization and dissemination of cultural resources. It incorporates oral history recordings, such as interviews with elder villagers and cultural inheritors, to preserve the living memories of ancient villages. Additionally, it provides guidelines for 3D modeling and digital collection of spatial layouts, natural environments, and architectural structures. The framework is characterized by its focus on historical records, family genealogies, manuscripts, and multimedia documentation, showcasing a professional archival approach.

Cultural elements in existing archiving rules

A comparative analysis of the three archiving rules in terms of data structures, content coverage, and technological applications revealed distinctive characteristics. The Ministry of Housing and Urban-Rural Development's rule emphasizes standardization and comprehensiveness, focusing on detailed basic data for unified national management and systematic conservation. The Chinese traditional village network's guidelines prioritize cultural and social life, combining textual and visual documentation to highlight the depth and breadth of cultural transmission and social structures. Renmin University's framework centers on the structured management and association of digital resources, leveraging modern technologies to enhance the efficiency and dissemination of archival materials.

Despite these strengths, existing rules primarily offer checklist-style guidelines for data collection without fine-grained annotation standards for the cultural, heritage, and memory elements within the archives. This limitation makes it difficult to identify distinct cultural characteristics from archived data, leading to the phenomenon of *'standardized archiving with little distinction'*.

To address this issue, this study identified a core set of cultural elements shared across the three rules, forming the foundational vocabulary for the proposed ontology framework. As illustrated in Figure 1, the study integrates key concepts related to village history, spatial relationships, and human elements. Cultural elements concerning historical narratives include the development history of villages, family lineages, heritage evolution, and architectural history. Spatial elements emphasize natural landscapes, cultural landmarks, spatial layouts, and architectural imagery. Human elements focus on identifying key figures who have significantly influenced the historical and spatial development of ancient villages.
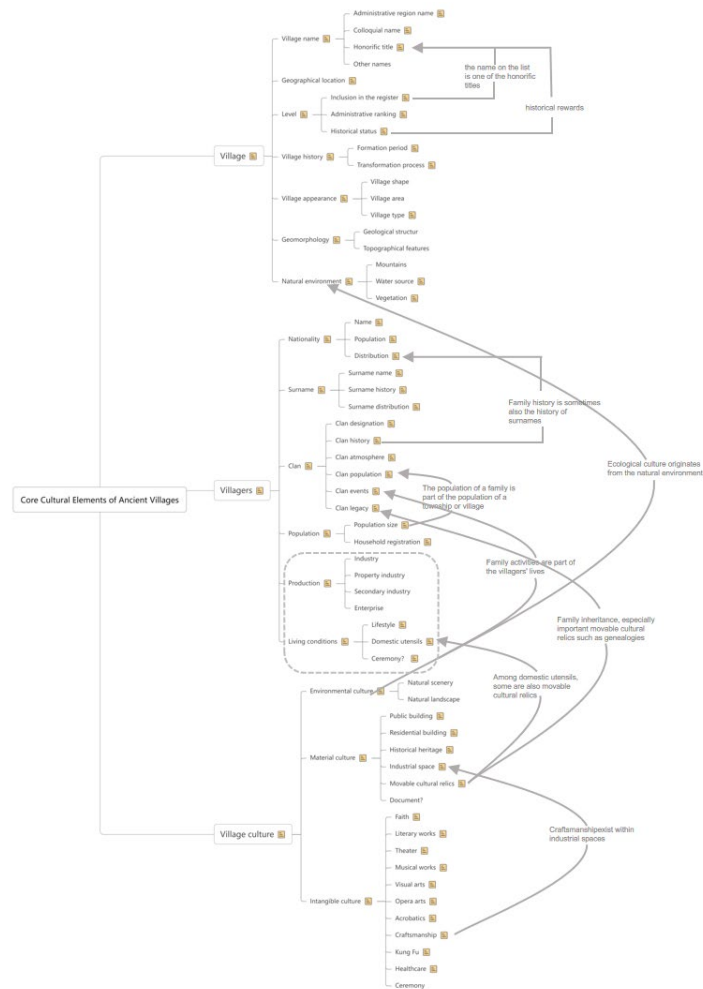
Figure 1. Core cultural elements of ancient villages in existing archiving rules

The systematic identification of core cultural elements of ancient villages addresses the limitations in data annotation within the three major archiving rules. However, it lacks the establishment of diversified relationships between elements, failing to support deeper exploration and discovery of cultural knowledge associated with ancient villages. Moreover, the identified core elements exhibit semantic redundancy or ambiguity when aligned with common semantic frameworks in cultural heritage or historical research. This necessitates further work on semantic disambiguation and alignment to build a systematic and clear semantic model. Such a model would facilitate uncovering humanistic connections within ancient village archives and establish linkages between these archives and other related historical and cultural heritage resources.

Design of a semantic framework for digital archives of ancient villages based on general ontology reuse

After extracting core cultural elements from existing archiving rules, this study further developed a semantic framework for digital archives of ancient villages by reusing general ontologies commonly employed in international cultural heritage and historical research. The CIDOC conceptual reference model (CIDOC-CRM) (Raimond, Y., Abdallah, S.2007) was the primary ontology reused, while the event ontology was applied to describe historical events related to ancient villages. Research on narrative ontologies (Ciotti, F.2016, p.29-44) was utilized to describe the stories of historical figures associated with ancient villages.

Adaptability of general ontologies

The CIDOC-CRM is an internationally recognized conceptual model developed by the International Committee for Documentation of the International Council of Museums (ICOM). It establishes a universal and extensible semantic framework for integrating cultural heritage information, enabling semantic definitions and structured descriptions of concepts and relationships in cultural heritage resources. The model's comprehensive descriptions of entities (e.g., entity) and properties (e.g., property) allow accurate depiction and relational mapping of key aspects such as people, places, and events associated with ancient villages. This makes it well-suited for resource integration and organization.

For this study, the CIDOC-CRM model was adapted to represent the geographical location of villages using the E53 Place category, individuals of historical significance as E21 Person, and relationships such as P8 took place on (linking events to relevant objects) and P74 had current or former residence (indicating places of residence over time). However, the generic nature of CIDOC-CRM presents limitations in addressing the specific characteristics of ancient villages. For example, the E41 Appellation category does not adequately represent the unique cultural designations of villages, which include architectural styles, traditional crafts, natural landscapes, and historical artifacts. Therefore, this study extended the *'appellation'* attribute to describe distinct cultural elements based on traditional archiving rules.

The event ontology provides a structured framework for describing and categorizing events, offering concepts such as event: event, event: factor (elements within an event), and event: product (event outcomes). However, it lacks specificity for relationships commonly seen in the historical events of Chinese ancient villages, such as indirect causal links between village founding or migration and family-related events, or implicit correlations between customs or production habits and wartime events. This study first categorized common historical event types of ancient villages and plans to further analyze complex event relationships in future research.

Existing narrative ontology research, while lacking uniform semantic frameworks, typically describes characters, their attributes and actions, story locations, and related events. These narrative elements complement the CIDOC-CRM and event ontology by filling gaps in the description of narrative aspects, particularly character-related attributes.

Based on the reuse of general ontologies, this study constructed a preliminary semantic framework for ancient villages. The framework was integrated with digital archive data from Gaoqian village in Zhejiang Province, which the research team has been developing since 2016. The framework was evaluated against a classification directory and conceptual model of cultural resources agreed upon by an expert team specializing in ancient villages. Expert feedback indicated that the current model provides high coverage of cultural elements, accurately infers relationships between cultural heritage and historical events, and offers good extensibility for unique features in the natural and cultural environments of ancient villages.

Basic components of the semantic framework for digital archives of ancient villages

Building upon the classification and alignment of core cultural elements, this study expanded descriptions to capture the unique cultural phenomena and attributes of ancient villages. The resulting digital archiving system centers on ancient villages and encompasses seven major entities, 29 relationships, and 94 properties. As shown in the partial entity-relationship (ER) diagram in Figure 2, the semantic framework includes six main categories: basic village information, natural heritage, tangible cultural heritage, intangible cultural heritage, historical documents, historical events, and notable figures.

This framework aims to support in-depth exploration of the historical context, spatial relationships, and human networks of ancient villages, providing a foundation for uncovering and analyzing their rich cultural heritage.
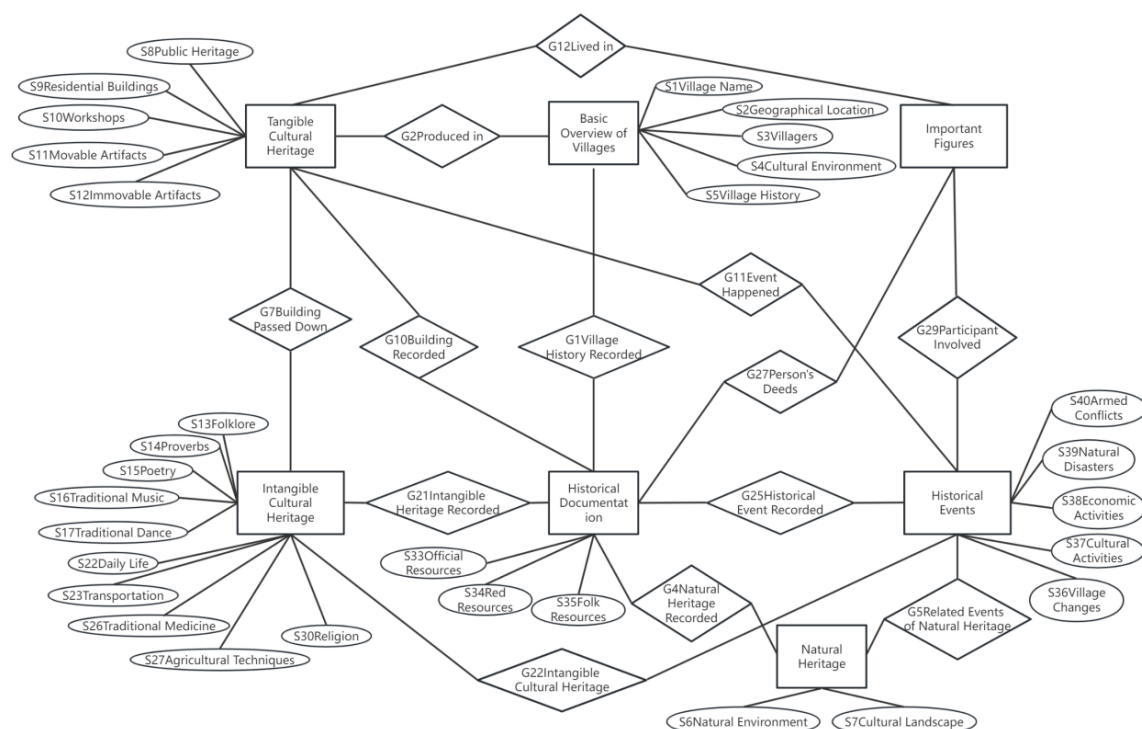

Figure 2. Sketch of digital archiving ER for ancient villages (only some relationships are presented)

Entity types in the semantic famework for digital archives of ancient villages

By analyzing existing rules for archiving ancient villages, it was found that aspects such as village overview, natural environment, and historical and cultural heritage are universally present in all the rules. Specifically, the *'registration form for Chinese traditional villages'* provides the most concise summary, the *'Information Table for Museumization of Chinese Traditional Villages'* offers the most comprehensive classification and content for tangible and intangible cultural heritage, The archiving standards developed by the Digital Memory of Ancient Villages Team at Renmin University of China provide the most detailed descriptions of village overviews and residents. In summary, this study excludes descriptions of tourism guidance in the *'information table for museumization of Chinese traditional villages'*, integrating culinary products and agricultural specialties into sections on folk customs, daily life, and knowledge skills. It adopts the maximum union of these three sets of rules and reuses the CIDOC-CRM standards suitable for describing ancient village resources to create a framework encompassing six main entities. (See Table 1).

Entity	Description	Example	Resource Type	Entity Code	Reuse Standard	Remarks

Basic Overview of Villages	**Village Name**		Refers to the administrative name of the village, typically given by the government based on geography, history, and culture	Haoshangqiao Village, Dayu Town, Gongyi City, Zhengzhou, Henan Province		S1		
	Geographical Location		The administrative area where the village is located, including province, city, township, etc.	Dayu Town, Gongyi City, Zhengzhou, Henan Province		S2	E53 Place	
	Villagers		Residents living in the village, including social, cultural, economic, and			S3		
	Cultural Environment		The social and cultural conditions in the village			S4		
	Village History		Historical descriptions of the village, including the formation era, reasons for its		Text records such as village chronicles	S5		
Natural Heritage	**Natural Environment**		The natural geographical conditions of the village			S6		
	Cultural Landscape		Landscapes that combine nature and human activities, holding cultural			S7		
		Public Heritage	Cultural heritage jointly owned, used, and maintained by the community, holding significant historical, cultural, social, or economic value	Temples, ancestral halls, stages, pagodas, historical sites, ancient bridges, wells, tombs, walls		S8		Organized by time, in the format "Building-Time"

Tangible Cultural Heritage	**Architecture**	**Residential Buildings**	Buildings used for daily living purposes	Cave dwellings, courtyards, roundhouses, bamboo houses, stilt houses	Text records such as village chronicles, photos, architectural drawings, and other artifacts	S9	
		Workshops	Places where people engage in handcrafted production and processing	Oil mills, papermaking workshops, tofu factories, dye houses, ceramic workshops, blacksmiths		S10	
	Artifacts	**Movable Artifacts**	Items of cultural heritage that can be transported or moved	Artworks, manuscripts, books, representative objects	Text records such as village chronicles, photos, drawings, and other artifacts	S11	First indicate storage location, then specify
		Immovable Artifacts	Cultural heritage that cannot or should not be moved due to its scale, nature, or association with a	Steles, murals, ruins, temples, city walls		S12	
Intangible Cultural Heritage	**Oral Culture**	**Folklore**	Cultural elements transmitted and preserved orally rather than in written form	Heroic legends, mythological stories		S13	
		Proverbs	Agricultural proverbs, life wisdom, observations of nature, social			S14	
		Poetry	Pastoral poems, labor songs, festival poetry, lyrical poems			S15	
		Traditional	Forms of art expressed through physical actions,	Folk songs, guqin (zither), erhu (two-		S16	

		Music	voice, expressions, and language	stringed fiddle)			
		Traditional	Festive dances, ethnic minority			S17	
		Traditional Drama	Peking opera, Yue opera, Huangmei opera, shadow puppetry			S18	
	Performing Arts	**Quyi (Chinese Folk Art Forms)**	Xiangsheng (comic dialogue), storytelling, drum songs		Text records such as village chronicles, photos, audio files, video files	S19	
		Acrobatics	Tightrope walking, face-changing, contortion			S20	
		Traditional Martial Arts, Sports, and Compet	Taijiquan, Baduanjin (Chinese exercise), dragon boat races, wrestling, archery			S21	
Intangible Cultural Heritage	**Folk Life**	**Daily Life**	Traditional lifestyle of the village, an important part of cultural identity	Farming, folk games, folk arts		S22	
		Transportatio	Horse carriages, water transportation			S23	
		Rituals and Custom	Wedding and funeral customs, sacrificial rituals, banquet			S24	
		Household Beliefs	Ancestor worship, deity worship, feng shui beliefs			S25	
	Knowledge and Skills	**Traditional Medicine**	Knowledge and skills gained through time, experience, and learning in the village	Herbal medicine, acupuncture, cupping therapy, massage therapy		S26	

		Agricultural Techniques	Sowing, fertilizing, irrigation			S27		
		Traditional Handicrafts	Weaving, embroidery, pottery making, carpentry			S28		
		Knowledge of Natural and Astronomical Phenomena	Seasonal changes, star observation, feng shui			S29		
	Spiritual Beliefs	Religion	Belief systems collectively followed in the village that influence values, behavior, and	Taoism, Buddhism, Confucianism		S30		
		National Spirit	Solidarity, hard work, frugality, respect for nature, including special regional or wartime			S31		
		Moralit	Filial piety, integrity			S32		
Historical Documentation	Official Resources		Documentation, description, or analysis of figures or events related to the village	Village chronicles, local records, land deeds, ration tickets	Text records, photo archives	S33		
	Red Resources		Revolutionary documents, stories of significant figures during revolutionary times			S34		
	Folk Resources		Maps, objects collected from non-official sources, including ancestral portraits,			S35		
Historica	Village		Significant events in the village that have	Foundation of the village,	Text records,	S36		

l Events	Changes	shaped its development	relocation process	photo archives			
	Cultural Activities	Festivals, religious events, educational activities, cultural exchanges			S37		
	Economic Activities	Agricultural production, handicrafts, commercial trade			S38		
	Natural Disasters	Floods, droughts, earthquakes			S39		
	Armed Conflicts	Events related to defense and resistance during			S40		
Importan t Figures		Individuals who have made significant contributions or have had a major influence on the	Cultural inheritors, historians		S41	E21 Person	

Table 1. Entity table of the semantic framework for digital archives of ancient villages

Entity relationships in the semantic framework for digital archives of ancient villages

The relationships between entities in the semantic framework for digital archives of ancient villages primarily occur between various types of cultural heritage and historical narratives. Tangible cultural heritage, intangible cultural heritage, and documentary heritage each record, reflect, and reveal the historical events, figures, activities, and phenomena of ancient villages. As shown in Table 2, these relationships not only construct a multidimensional network of cultural resources but also, through interconnected nodes, demonstrate the dynamic evolution of ancient village histories. For example, tangible cultural heritage, such as ancient buildings and sites, is often closely linked to specific historical events or figures. Through the documentation of documentary heritage, the social context and cultural significance behind these events or figures can be further revealed. Intangible cultural heritage, such as traditional crafts and folk activities, when associated with documentary and tangible cultural heritage, reflects the lifestyles and spiritual aspects of the residents of ancient villages. These multidimensional entity relationships not only help to comprehensively understand the cultural significance of ancient villages but also provide a scientific basis for their preservation and inheritance.

Code	Relationship Name	Description	Entity 1	Entity 2	Reuse Standard
G1	Village History Records (Recorded)	Historical documents record the history of the village	Village History	Historical Documents	
G2	Produce in	The place where production takes place	Production	Workshop	
G3	Used for Production (Used for)	Tools used in production	Production	Movable Artifact	

G4	Natural Heritage Records (Recorded)	Historical documents record the natural heritage	Natural Heritage	Historical Documents	
G5	Related Events of Natural Heritage	Historical events related to the natural heritage	Natural Heritage	Historical Event	
G6	Buildings based on skills	The craftsmanship used in building construction	Building	Traditional Handicrafts	
G7	Building Passed Down	Oral traditions related to the building	Building	Oral Culture	
G8	Building Used for (Required)	The application of buildings in folk life	Building	Folk Life	
G9	Building Carries (Recorded in)	The spiritual beliefs carried by the building	Building	Spiritual Belief	
G10	Building Recorded (Recorded)	Historical documents record the building	Building	Historical Documents	
G11	Event Happened (Happened in)	Events that happened in the building	Building	Historical Event	P8 - took place on or within (witnessed)
G12	Lived in (Residence)	The residence where the person lived	Building	Important Person	
G13	Artifacts based on skills	The craftsmanship used in the creation of artifacts	Artifact	Traditional Handicrafts	
G14	Artifact Passed Down	Oral traditions related to the artifact	Artifact	Oral Culture	
G15	Props Used for (Used in)	Props used in performing arts	Artifact	Performing Arts	
G16	Artifact Used for (Required)	The application of artifacts in folk life	Folk Life	Artifact	
G17	Artifact Carries (Recorded in)	The spiritual beliefs carried by the artifact	Artifact	Spiritual Belief	
G18	Artifact Records (Recorded)	Historical documents record the artifact	Artifact	Historical Documents	
G19	Artifact Related Event	Historical events related to the artifact	Artifact	Historical Event	
G20	Records Person (Recorded in)	The artifact records the person's life and contributions	Artifact	Important Person	
G21	Intangible Heritage Recorded (Recorded)	Historical documents record the intangible cultural heritage	Intangible Cultural Heritage	Historical Documents	

Code	Attribute Name	Entity	Attribute Description	Example	Reuse Standard
G22	Intangible Heritage Related Event	Historical events related to intangible cultural heritage	Intangible Cultural Heritage	Historical Event	
G23	Intangible Heritage Passed Down	Oral traditions related to intangible cultural heritage	Intangible Cultural Heritage	Oral Culture	
G24	Tools Used (Used in)	Tools, decorations used in intangible cultural heritage	Intangible Cultural Heritage	Movable Artifact	
G25	Historical Event Records (Recorded)	Historical documents record the historical event	Historical Event	Historical Documents	
G26	Historical Event Passed Down	Oral traditions related to the historical event	Historical Event	Oral Culture	
G27	Person's Deeds	The main deeds and contributions of the person	Important Person	Historical Documents	
G28	Person Passed Down	Oral traditions related to the important person	Important Person	Oral Culture	
G29	Participants have (Participates in)	Participants involved in the historical event	Important Person	Historical Event	

Table 2. Entity relationships in the semantic framework for digital archives of ancient villages

Attribute descriptions in the semantic framework for digital archives of ancient villages

During the definition of the basic intensions and extensions of the six major entities and their subcategories, this study conducted a general description of the attributes commonly found in ancient villages across various regions and types in China. A basic attribute table was created, listing the attribute descriptions of ancient village entities in detail, including but not limited to geographical location, historical development, architectural style, cultural value. By accurately describing these attributes, it becomes possible to reflect the multidimensional information of ancient villages and provide abundant data support for subsequent mining of digital archive resources more comprehensively. In addition, the definition and classification of attributes refer to the standards of CIDOC-CRM, ensuring consistency and interoperability of data. Through this approach, the semantic analysis of ancient village digital archives can go beyond the level of entities and relationships to delve into the details of attributes, thus providing a solid foundation for future research and applications.

Code	Attribute Name	Entity	Attribute Description	Example	Reuse Standard
D1	Common Name		The informal name of the village commonly used by local residents, possibly derived from the village's history, geography, or famous stories	For example, Haishangqiao Village is commonly known as "Haiyan"	
D2	Honorary Title	Village Name	Honorary titles awarded to the village by the government or other authorities in recognition of its unique cultural, historical, economic, or social contributions	For example, listed in the fifth batch of Chinese traditional villages	

D3	Ethnicity		The ethnic distribution in the village and related historical activities	For example, Han ethnicity	
D4	Surname		The distribution of surnames in the village and related historical activities	For example, Wang family in Haishangqiao Village	
D5	Clan	Villager	The distribution of clans in the village and related historical activities	For example, the Hua family clan in Youdian Village	
D6	Population		The population count and household registration in the village		
D7	Production		The main industries in the village and the villagers' way of life	For example, agriculture, livestock, fishery, forestry, handicrafts, tourism	
D8	Streets and Alleys		The distribution of functional areas in the village	For example, residential, commercial facilities, entertainment facilities, village hospital	
D9	Defense Structure	Cultural Environment	Structures built to protect the village from natural disasters or military attacks	For example, city walls, fortresses, air-raid shelters, dikes	
D10	Roads		The distribution of roads in the village		
D11	Formed Time Is		Important event nodes in the historical development of the village	For example, Haishangqiao Village was established in the fifth year of the Qianlong reign (1740)	P4 - has time-span (is time-span of)
D12	Formation Reason	Village History	The reason for the formation of the village	For example, a natural spring in the western valley led to the construction of a bridge for residents to cross	
D13	Evolution Process		The process of village evolution, including formation, migration, etc.		
D14	Area	Natural Environment	The area of the village	For example, 2865 acres of arable land, approximately 6.7 square kilometers	

			The type and terrain of the village	For example, mountain, hill, plateau	
D15	Shape		The type and terrain of the village	For example, mountain, hill, plateau	
D16	Vegetation and Animals		The biological coverage of the village	For example, crops like wheat, corn, millet, sweet potatoes, sesame, peanuts, sunflowers, rapeseed	
D17	Soil Type		The type of soil in the village, such as red soil, black soil, yellow soil	For example, red soil, black soil, yellow soil	
D18	Hydrology		The hydrological conditions of the village, such as rivers and lakes	For example, river basins, wetlands	
D19	Mountain and Terrain		The mountains and terrain of the village	For example, backed by the Taihang Mountains, flowing through the Jinsha River	
D20	Climate Type		The climate zone of the village	For example, subtropical climate, temperate monsoon climate	
D21	Natural Disaster		The types of natural disasters experienced by the village	For example, mudslides, floods	P4 - has time-span (is time-span of)
D22	Cultural Landscape Type Is		The type of cultural landscape in the village	For example, ancient trees, ancient wells	
D23	Formation Time Is		The formation time of the cultural landscape		
D24	Maintenance Time Is		The time when the cultural landscape was maintained		
D25	Disappearance Time Is	Cultural Landscape	The time when the cultural landscape disappeared		
D26	Formation Place Is		The place where the cultural landscape was formed		
D27	Cultural Landscape Recorded Has		The records related to the cultural landscape		
D28	Function Has		The main function of the cultural landscape	For example, ancient wells providing freshwater	P101 - had as general use (was use of)

D29	Formation Time Is		The formation time of the building		P4 - has time-span (is time-span of)
D30	Maintenance Time Is		The time when the building was maintained		
D31	Disappearance Time Is		The time when the building was destroyed		
D32	Formation Place Is		The place where the building was formed	For example, the Wang family courtyard in Haishangqiao Village in Zhengzhou, Henan	
D33	Events Happened		Events that took place in the building	For example, France and Germany signed armistice agreements in the same railway carriage in 1918 and 1940	P8 - took place on or within (witnessed)
D34	Area Is	Building	The area occupied by the building		
D35	Number of Floors Is		The number of floors in the building		
D36	Creator Is		The designer or creator of the building		
D37	Maintainer Is		The person who maintains the building		
D38	Resident Is		The past and present residents of the building		
D39	Owner Is		The past and present owners of the building	P52 - Current owner (owns it now)	
D40	Building Type Is		The type of building	For example, ancestral hall, temple, academy, drum tower, opera stage, guild hall	
D41	Protection Level Is		Whether the building is included in various protection lists	For example, national protection, provincial protection, municipal protection, county protection	

Table 3. Attribute table in the semantic framework for digital archives of ancient villages (Partial Illustration)

The humanistic discovery capabilities of the semantic framework for digital archives of ancient villages: a case study of four villages in Henan province

Overview of the four villages in Henan province and their digital archives

Henan Province boasts abundant political, military, artistic, and cultural resources, nurturing ancient cities and preserving numerous historically rich ancient villages. This study selects four villages in Henan: Haishangqiao, Yuanqiao, Youdian, and Dawangmiao —for integrated analysis of their digital archive resources. By conducting semantic analysis based on the semantic framework proposed in this study, the research aims to verify the applicability of the semantic framework to the analysis of actual digital archives of ancient villages and explore innovative paths for constructing and presenting humanistic knowledge associations within ancient villages.

These four villages, all listed in the catalogue of Chinese traditional villages, possess rich archival resources covering almost every aspect of the archiving system, with unique characteristics for each village: Haishangqiao preserves intact archives of notable residences and biographies, featuring a wealth of personal records; Yuanqiao retains plaques from the Ming and Qing dynasties and historic buildings, offering substantial tangible archives; Youdian, known for its distinctive cave dwellings and sunken courtyards, holds abundant property-related archives and Dawangmiao served as the site of Henan University during the anti-Japanese war period and contains significant manuscript archives related to red cultural resources. From the perspective of archival resource types, the digital archives of these four villages are primarily composed of text and image resources, with limited digitalized audio and video resources. In terms of archival content, most of the information concentrates on the past 50 years, with some records of legends and background concerning the villages' founding but lacking detailed textual and physical documentation. The specific conditions of the digital archives of these four villages are presented in Table 4 (see Table 4).

Resource Status		Dayugou Town, Haishangqiao Village, Gongyi City, Zhengzhou, Henan Province	Dajindian Town, Yuanqiao Village, Dengfeng City, Zhengzhou, Henan Province	Shanhua Town, Youdian Village, Yanshi District, Luoyang, Henan Province	Tantou Town, Dawangmiao Village, Luanchuan County, Luoyang, Henan Province
Basic types of digital resources	**Text Resources**	'Haishangqiao Village Chronicles', Wang Family Genealogy, 'Da Jie Stele Inscription' etc.	'Speaking of Yuanqiao', List of attendees at the first Dengfeng County Party Congress, 'Dengfeng County Chronicles', stele inscriptions, etc.	'Hua Family Genealogy', 'Youdian Town's Restoration of the Bull King Temple', 'Hong Xing Guan's Patching Stele', 'Establishment of Jade Pavilion', 'Yan County Chronicles', 'Youdian Village Chronicles', etc.	'Sun Family Genealogy', 'Imperial Edict Stele', 'Dawangmiao Village', Sun Family Genealogy, etc.

Image Resources	Real estate certificates from the Republic of China period, deeds, food coupons from the early PRC period, historical buildings like Guang ye Wang's residence, etc.	Deeds, property rights, tax payment certificates from the Republic of China period, Yuan Family Genealogy, historical buildings like tunnels, earthen ramparts, San Xian Temple, etc.	Donation lists, historical buildings like cave dwellings, ancient wells, ancient north, ancient plaques, etc.	Route maps and classroom design drawings of Henan University Agricultural College, student IDs, photos of famous scholar Jing run Chen speaking at Henan University, historical buildings like Wen Qu Bridge, Sun Family Mansion, Li Family Mansion, etc.
Video Resources		Yuanqiao 1979		
Data Resources			Statistics of scholars from the Ming and Qing dynasties, statistics of college students before the founding of New China, statistics of junior college students after the liberation of Youdian Village, statistics of political, military, educational, and business situations after the liberation of Youdian Village, etc.	
Online Resources	Traditional villages on the 'Track together' website	Traditional villages on the 'Track together' website	Traditional villages on the 'Track together' website	Traditional villages on the 'Track together' website
Features — **Spatiotemporal coherence**	From the Qianlong era of the Qing Dynasty to the present, the village history	From the Hongwu era of the Ming Dynasty to the present, the village history	From the Northern Song Dynasty to the present, the village history has been documented	From the late Ming and early Qing dynasties to the present, the village history has

	has been recorded in the village chronicles.	has been documented in the book 'Speaking of Yuanqiao'.	in the 'Youdian Village Chronicles'.	been documented in the 'Dawangmiao Village'.
cultural heritage preservation	Well-preserved historical buildings and village civilization	Well-preserved historical buildings and village civilization		The village's red gene from the period of the war of resistance against Japan continues to this day
Storability and replicability	primarily text and image resources	primarily text and image resources	primarily text and image resources	primarily text and image resources
modality	Static	Static	Static	Static

Table 4. Overview of digital archive resources for the four villages in Henan province

Data processing workflow for digital archive resources of the four villages in Henan province

Based on resources such as village chronicles, museum collections, oral accounts from villagers, and field visit photographs, this study first standardized and annotated the data using Excel, following the semantic framework for digital archives of ancient villages described in Section 3.2. Subsequently:

Hierarchical entities were created using Protégé, with relationships and attributes added to incorporate the resources of the four villages. A knowledge base was built in Protege. To ensure clarity, flexibility, and distinctiveness, the OntoGraf visualization plugin was used to present the framework. For instance, Haishangqiao village highlights residential buildings associated with the Wang family, while Yuanqiao emphasizes the relational network of the Yuan family.

Semantic structure presentation of digital archives for the four villages in Henan province

Haishangqiao village, located in Dayugou town, Gongyi city, Zhengzhou, Henan province, is renowned for its historically significant Wang family genealogy and the *'Haishangqiao village chronicle'*. The history of the village's resources can be traced back to the Qianlong era of the Qing dynasty, with textual and image data as the primary materials. The village's historic architectural complex is a remarkable sight, with its overall layout well-preserved. The Wang family residence, as an iconic building, highlights the village's unique cultural characteristics. Within the resource framework, *'residential buildings'* emerge as a central element, radiating outward to form a distinct spatial layout. Detailed records of folk customs and daily life are also fully reflected in the framework, offering subsequent researchers an abundant and meticulous collection of materials.
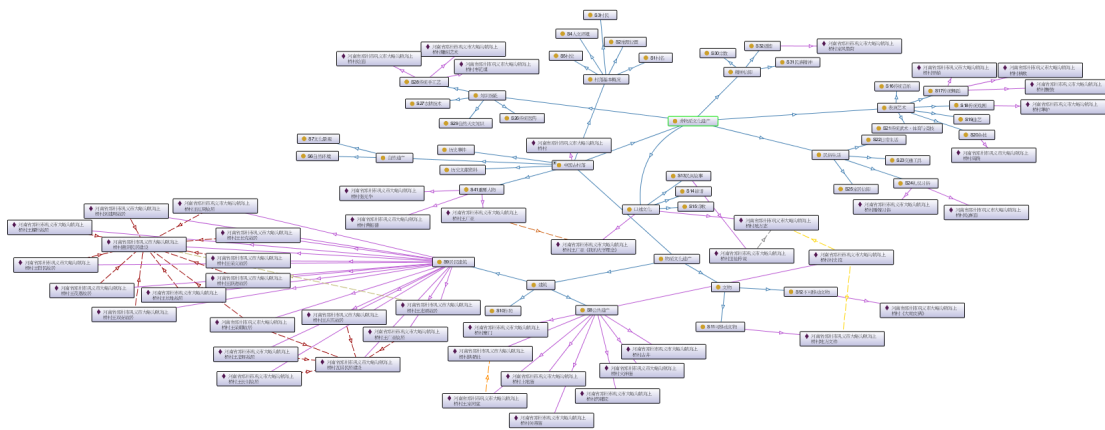
Figure 3. Semantic structures presented in the digital archives of Haishangqiao village

Yuanqiao village, located in Dajindian town, Dengfeng city, Henan province, possesses abundant resources. Centered on the village chronicle '*talking about Yuanqiao*', the archives provide detailed records of biographies, historical relics, events, and folk customs, with its history tracing back to the Hongwu era of the Ming dynasty. Numerous immovable cultural relics, including a significant number of inscribed stone tablets, have been well preserved. In the resource framework diagram, the Yuan family lineage serves as the main thread, clearly illustrating the continuity and evolution of Yuanqiao village in terms of folk culture, village development, historical progression, and economic growth.
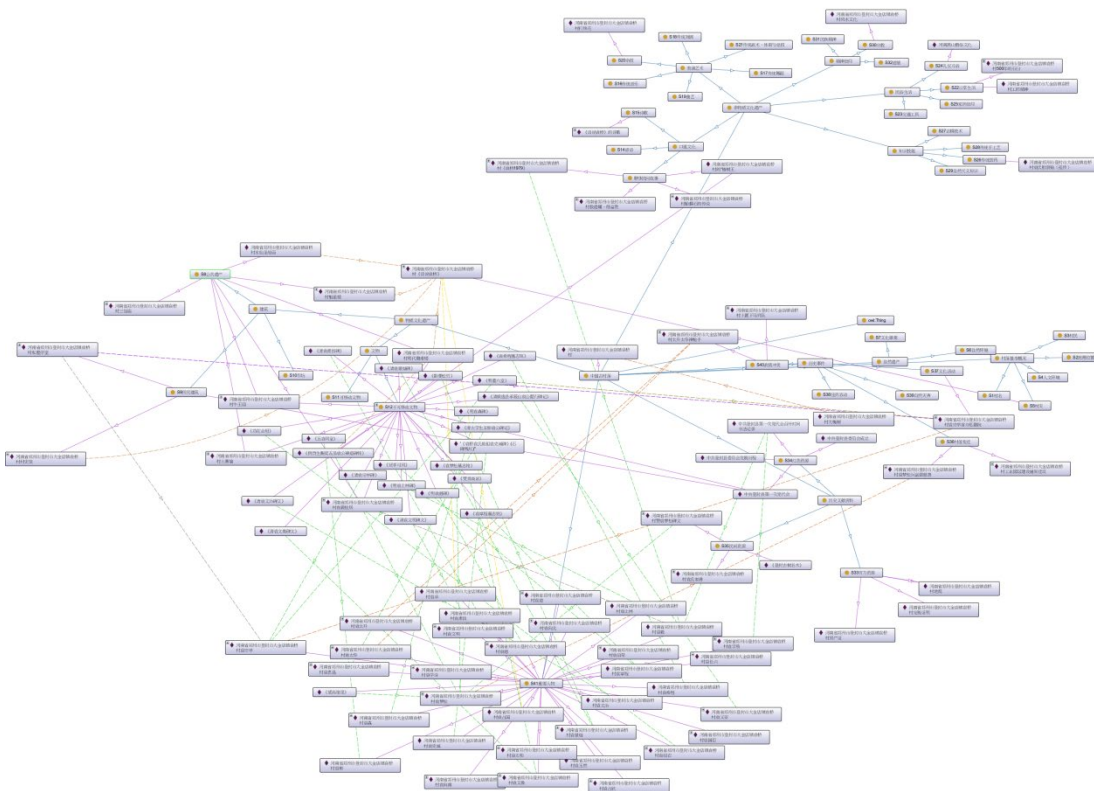
Figure 4. Semantic structures presented in the digital archives of Yuanqiao village

Youdian village, situated in Shanhuo Town, Yanshi District, Luoyang City, Henan province, is notable for its distinctive sunken courtyards and well-preserved immovable cultural relics such as stone tablets, plaques, ancestral halls, and temples. The Hua family, with its 22-generation genealogy, is central to the village's history. The village chronicle 'Hua family genealogy' documents the village's history, biographies, public heritage, and the unique historical narratives of its stone tablets in great detail. In the resource framework diagram, 'Youdian village chronicle' forms the core, supported by immovable relics and public heritage. The diagram illustrates the village's historical development, traces of its prominent figures, and the literary corpus centered on its numerous stone tablets, presenting a vivid and multidimensional portrayal of Youdian Village.
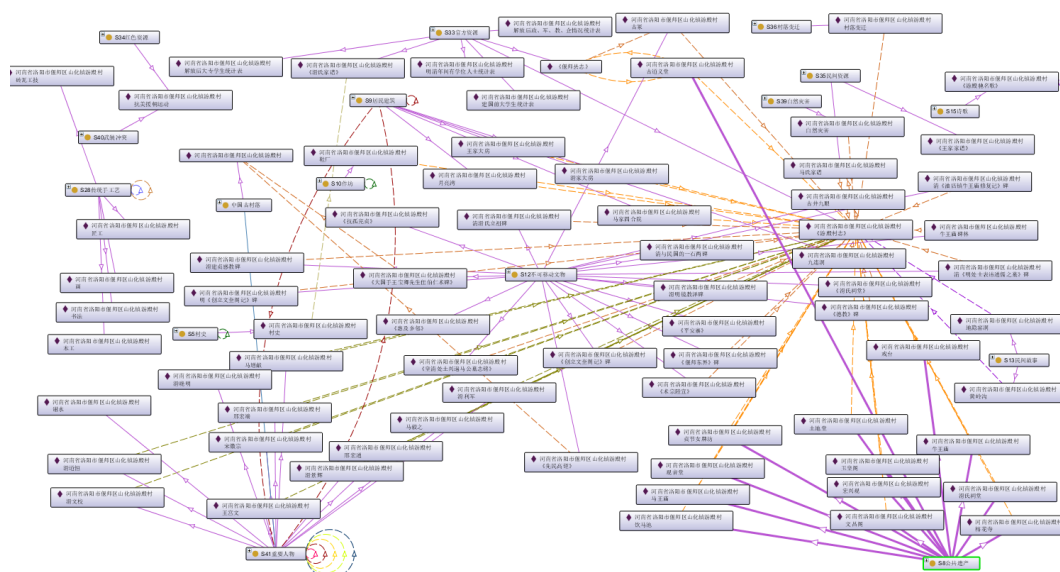

Figure 5. Semantic structures presented in the digital archives of Youdian village

Dawangmiao village, located in Tantou town, Luanchuan county, Luoyang city, Henan province, was once a campus of Henan university. The site still preserves the former women's dormitory of the university and holds rich red cultural resources, although some public heritage has suffered damage. The village's archives focus on its wartime history and Henan university's educational legacy. In the resource framework diagram, prominent figures are mainly educators and teachers, while most immovable cultural relics are associated with Henan university, including the women's dormitory and teaching tools. Folk stories center on red narratives of resistance against foreign invaders during the war. These elements collectively highlight Dawangmiao village as a red cultural site rooted in patriotism and national sentiment.

Figure 6. Semantic structures presented in the digital archives of Dawangmiao village
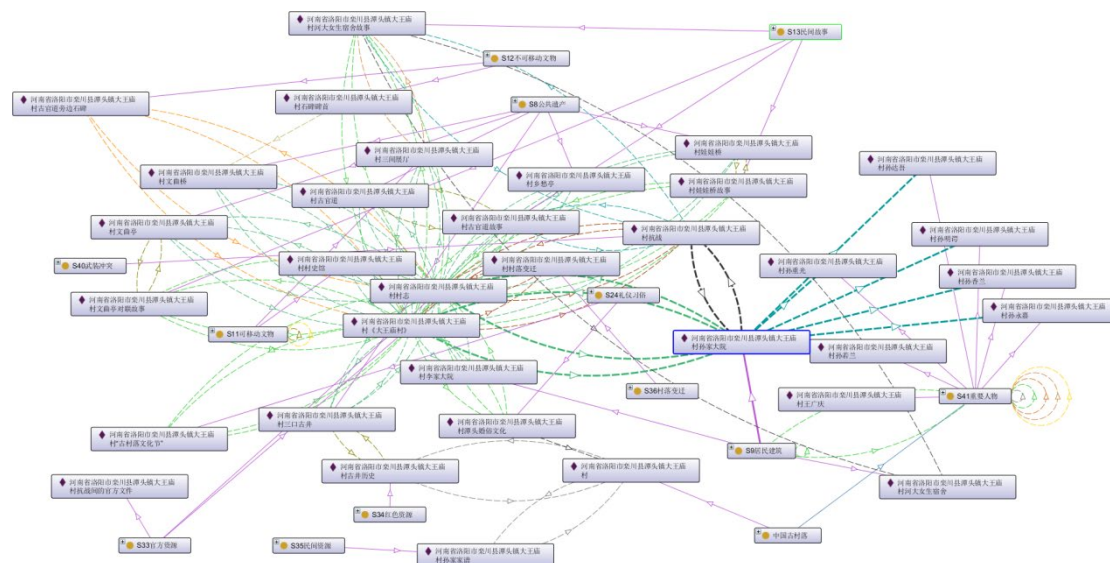
Future presentation of digital archives of ancient villages for cultural narratives

Based on the extraction of core cultural elements from the archival standards of Chinese ancient villages, this study employs internationally recognized semantic frameworks and cultural analysis models to construct a semantic analysis framework for digital archives of ancient villages. Using the digital archival resources of four villages in Henan Province as a case study, it explores a pathway for cultural knowledge discovery rooted in the digitization, semantic annotation, and presentation of ancient village archives.

However, cultural knowledge discovery is only the initial step. The challenge lies in effectively presenting this knowledge within digital spaces. Research teams must focus on designing immersive and interactive narrative scenarios grounded in the spatial storytelling of ancient villages. Future directions may include developing an interactive knowledge-based Q&A system that integrates online and offline experiences, educational competitions, and personalized digital guides for ancient villages. By leveraging technologies such as VR and AR, immersive exhibition halls for ancient villages could be created, along with interactive cultural games aimed at decoding historical elements. As digital technologies continue to advance, this framework has the potential to be further optimized, making significant contributions to the preservation and promotion of ancient village cultural heritage.

Acknowledgements

About the author(s)

Tianjiao Qi is associate Professor in School of Information Resource Management and one of researchers in Research Center of Digital Humanities of Renmin University of China. She received her Ph.D. from Renmin university of China, and her research interests are semantic organization of digital archives, digital cultural heritage, and digital memory. She can be contacted at qtjjoy@163.com.

Yiduo Wang is a bachelor student in the School of Information Resource Management at Renmin University of China. She can be contacted at 2022200983@ruc.edu.cn.

Meiman Wong is a bachelor student in the School of Information Resource Management at Renmin University of China. She can be contacted at wongmeiman@qq.com.

Tsamchu Basang is a bachelor student in the School of Information Resource Management at Renmin University of China. She can be contacted at bsc2022202829@163.com.

Linglu Li is a bachelor student in the School of Information Resource Management at Renmin University of China. She can be contacted at 18888525801@163.com.

References

An H.Q., (2017). A preliminary study of traditional village archives management work. Beijing Archives,2017(05):29-30. https://doi.org/10.3969/j.issn.1002-1051.2017.05.009

CIDOC. what is the CIDOC CRM?   https://cidoc-crm.org/

Ciotti, F., (2016). Toward a Formal Ontology for Narrative[J]. Matlit Revista do Programa de Doutoramento em Materialidades da Literatura ,20164(1):29-44. https://doi.org/10.14195/2182-8830_4-1_2

Digital Museum of Chinese Traditional Villages. Chinese traditional villages digital museum village building filling instructions. https://wj.dmctv.cn/index/中国传统村落数字博物馆村落建馆填报说明.pdf

FAN,Q.,TAN,G.X.,SUN,C.M.,(2023) Research on video resource description and knowledge organization of intangible cultural heritage. Journal of Central China Normal University(Natural Sciences) ,2023,57(03):469-482. https://doi.org/10.19603/j.cnki.1000-1190.2023.03.017

HU,Y., CHEN,S.,CAO,W., CAO,C,Z., (2014) The Concept and Cultural Connotation of Traditional Villages. Urban Development Studies,2014,21(01):10-13. https://doi.org/10.3969/j.issn.1006-3862.2014.01.021

Luo Z.Y.,(2019).Unified implementation of filing and listing protection for ancient towns, vill ages, and residential buildings throughout the province.Headline. https://www.toutiao.com/article/6711169247831130638/?upstream_biz=doubao&source=m_redirect&wid=1736261058380

He,S.Y., (2017). Guarding the Rvyal Memory: Study on the Archiving of Traditional Village.Archives Science Study ,2017(05):49-53. https://doi.org/10.16065/j.cnki.issn1002-1620.2017.05.009

Liu,H.Z.,Bao,H.,Yu,J.H., (2006). A Semantic Web Architecture of Virtual Museum Based on CIDOC CRM. Application Research of Computers ,2006,(04):50-53. https://doi.org/10.3969/j.issn.1001-3695.2006.04.017

Man,Y.,Wang,P., (2017). Present Situation and Strategy Research of Traditional Village Archiv aI Protection. Archives & Construction,2017(07):12-15. https://doi.org/10.3969/j.issn.1003-709 8.2017.07.004

National Archives Administration of China. (2021). Research on the Digital Memory Construction of "Ancient Villages in Taizhou". National Archives Administration of China. https://www.saac.gov.cn/daj/kjcgtg/202102/1a43b031ef3d4ef08bb12e3c66fb5a5f/files/0bb22f 3af6844ae7b5fd9f815154ec66.pdf

Raimond, Y., Abdallah, S., (2007). The Event Ontology.https://motools.sourceforge.net/event /event.html

Tang, Q., (2023) Research on archival protection and governance of traditional village cultural heritage based on living heritage. China Archives,2023, (10):66-68. https://kns.cnki.net/kcms2/article/abstract?v=YK1b_MowDz8P0ikNU4izkQyVvLor_QCrpq7m uznj7QSDXAN6j9mtqsCE_oqrq3UB_q5yshaFdyoWSVHVHFaJQ6RdPQLPQq-OGZFIIxBd-Q6Eg2jOiJjijxlfVStJ_HGunHqA8nSk_zkN2ez9Fd036nH9Q3Re6UvzNQQTGjfWxl4=&uniplatform= NZKPT

Traditional Chinese Village Digital Museum.Traditional Village Weekly Information Essence (November 2, 2024- November 8, 2024).https://www.dmctv.cn/zxShow.aspx?id=3454

Traditional Villages Network. Traditional village archiving survey (text) filing form. http://www.chuantongcunluo.com/index.php/Home/Lddc/bgxz_wz.html

Traditional Villages Network. Traditional village archiving survey (picture) archive table. http://www.chuantongcunluo.com/index.php/Home/Lddc/bgxz_tupian.html

Xinhuashe, (2023) First Watch|Traditional village protection, general secretary emphasizes. https://news.cyol.com/gb/articles/2023-04/28/content_4wlPRVIWlN.html

XU, X.Y., LIU, X.X., (2017) Interpretation of Connotation and Differences between Ancient Villages Archives and Rural Archives.Archives Science Study.https://doi.org/10.16065/j.cnki.issn1002-1620.2017.04.008

Zheng, J.X., Luo, Y.G., Han, Y.H., (2023). Research on the Document System of Multi-Subjec t Collaborative Management of Traditional Village Cultural Heritage. Comparative Study of Cultural Innovation,2023,7(34):194-198. https://kns.cnki.net/kcms2/article/abstract?v=YK1b_ MowDz9k0ew_GX_xiEqk27ItNsGTR9RBq4zR7hvQ8yll0zfKrMtERm2_v9Smlqb0FE6RrvQwHW--I S4joZ5qyN9be17y9dq3dW9oIit-C7Yzd2Ad1a-Zpg4BZ1_xQk4hUuqv3Jy6j75HqQaqylzwcFjRnUuB_C Dy26k_Ulo=&uniplatform=NZKPT