Digital Narrative Framework for Red Figure Resources and Its Empirical Study

Authors: Xu Xiaojuan, Sun Aihua, Chen Tao

Date: 2023-07-20T00:00:00+00:00

Abstract

[Purpose/Significance] Digital narrative, as a form of online narrative, helps enhance the interpretability and human-computer interactivity of red figure resources, providing possibilities for optimizing the knowledge-based development of red figure resources. [Method/Process] Based on a review of research on the knowledge-based development of historical figure resources and related studies on red figures, this paper proposes a basic framework for digital narrative of red figure resources consisting of a narrative resource integration layer, a narrative element extraction layer, and a narrative work generation layer. Subsequently, taking "Red Figures in Western Anhui" as a case study, knowledge graphs are employed as a narrative tool for empirical research. [Results/Conclusion] The digital narrative framework for red figure resources can address the limitations of current approaches that primarily design entity semantic associations based on the "content" and "external" levels of red resources, forming an operational structural process that provides reference for the knowledge-based development of red resources; compared with traditional knowledge extraction methods relying on manual extraction and rule matching, the batch triple extraction model based on ontology and BERT improves the efficiency of knowledge extraction from red figure resources; through knowledge graphs, visualization of red figure resources and knowledge discovery in thematic narrative are achieved, demonstrating the application value of digital narrative in the knowledge development and organization of red figure resources.

Full Text

Digital Storytelling Framework for Red Character Resources and Its Empirical Research

Xu Xiaojuan[1], *Sun Aihua*[1], *Chen Tao*[2]
[1]School of Management, Anhui University, Hefei 230039

[2]School of Information Management, Sun Yat-sen University, Guangzhou 510006

Abstract

[**Purpose/Significance**] Digital storytelling, as an online narrative form, enhances the interpretability and human-computer interactivity of red character resources, offering new possibilities for optimizing their knowledge-based development. [**Method/Process**] Building upon a review of knowledge-based development of historical character resources and relevant research on red characters, this study proposes a foundational digital storytelling framework for red character resources comprising three layers: narrative resource integration, narrative element extraction, and narrative work generation. Subsequently, taking "Red Characters of Western Anhui" as a case study, we conduct an empirical investigation using knowledge graphs as the narrative tool. [**Results/Conclusion**] The digital storytelling framework for red character resources addresses current limitations in designing entity semantic associations that rely primarily on "content" and "external" layers of red resources, establishing an operational structural workflow that provides a reference for knowledge-based development of red resources. Compared with traditional manual extraction and rule-based matching methods, the ontology and BERT-based batch triple extraction model significantly improves the efficiency of knowledge extraction from red character resources. Through knowledge graphs, we achieve visualization of red character resources and thematic narrative-based knowledge discovery, demonstrating the application value of digital storytelling in the knowledge development and organization of red character resources.

Keywords: red character resources; digital storytelling framework; knowledge graph; thematic knowledge discovery
Classification: G350

President Xi Jinping has repeatedly emphasized the importance of developing and utilizing red literature resources, elevating this work to a national cultural strategy. He stresses the need to make good use of red resources, continue the red legacy, and strive to create new achievements worthy of history and the people [1]. In recent years, the integration and development of red literature resources have moved beyond shallow descriptions of external features such as metadata, delving deeper into the resource content to represent "knowledge units" composed of entities and relationships, thereby achieving value-added development through mining semantic associations among these knowledge units [2]. Red characters, as creators and disseminators of red culture, serve as the primary carriers for presenting and exhibiting red resource content. As Lin Yutang observed, "History originates from events, events originate from people; without people there are no events, and without events there is no history." However, constrained by limited digital technology application and incomplete multimodal feature recognition [4], traditional planar and linear narratives can no longer meet users' diverse needs. Consequently, traditional narratives of red

character resources urgently require transformation toward digital storytelling, particularly in terms of enhancing narrative interpretability within historical contexts and improving human-computer interactivity on resource platforms.

Narrative theory, as a theoretical framework for studying "the symbolic representation of a series of events connected in time and causality," provides theoretical guidance for knowledge organization through its spatial-moral-interactive-transmedia narrative systems. It can vividly reproduce relationships among revolutionary cultural historical materials, present historical facts comprehensively, avoid fragmented interpretations, and enable more thorough excavation of red literature resource values [5-7]. Therefore, this study introduces narrative theory to explore a digital storytelling framework for red character resources that encompasses resource integration, element and relationship extraction, and knowledge application, taking into account the diverse characteristics of red character resources in terms of origin, type, content, and structure. Using "Red Characters of Western Anhui" data as a case study, we explore knowledge associations among red character resources, reconstruct entity scenarios involving people-events-locations-time, complete knowledge aggregation of red character resources, and form a multi-level knowledge organization and service system that integrates narrative features with both internal and external resource characteristics.

2.1 Related Research on Knowledge-Based Development of Historical Character Resources

The development of digital intelligence technologies has advanced the knowledge organization and intelligent services of historical character resources, with relevant theories and practices continuously evolving. Research outcomes have gradually shifted from coarse-grained literature organization to fine-grained knowledge organization, significantly improving the knowledge-based development level of historical character resource organization and digital humanities research. Current domestic and international research primarily focuses on three dimensions: ontology research, knowledge graph construction, and knowledge base development.

First, regarding ontology research, scholars mainly focus on constructing historical character ontology models, such as pre-Qin figure knowledge ontology models and the Prince Jingjiang Mansion figure relationship ontology, applying these models to knowledge extraction and figure relationship inference to improve user experience in knowledge discovery [8-10]. Second, in knowledge graph research, scholars have respectively focused on multi-layered relationship networks among historical figures to identify different factions [11], constructed social network graphs to analyze kinship and academic society changes in Zhejiang's Wuzhou region during the Song Dynasty [12], developed data models for historical celebrities' study tour footprints [13], automated construction methods for figure resume knowledge graphs [14], and explored digital humanities analysis tools for mining celebrity diary resources [15], while beginning to address

hidden relationship inference [16]. This dimension emphasizes empirical testing with diverse cases. Third, concerning knowledge base construction, scholars explore construction approaches and optimization pathways for specialized figure databases [17], such as using linked data technology to develop and reorganize library special collection resources to build a genealogy knowledge service platform that supports both root-seeking searches and data analysis for researchers [18,19], and examining basic models for academic celebrity knowledge base construction with technical and operational processes illustrated through the Wu Baokang academic celebrity knowledge base case [20].

2.2 Related Research on Red Character Resources

In recent years, the integration, development, and research of red character resources have received increasing attention. Cultural memory institutions such as libraries, archives, and museums have successively organized red character resources based on local or institutional special collections and provided online services. Representative platforms include the Shaanxi-Gansu-Ningxia Border Region Red Memory · Figure Database of Shaanxi Provincial Library [21], the Chen Yi Special Collection of Fudan University Library [22], and the CPC Organizational History Special Collection of Shanghai Library's Red Literature Platform [23]. Evidently, red character resources are undergoing a transformation from "collection" to "utilization." Meanwhile, existing research primarily focuses on red character resource database construction and resource organization/application.

First, regarding red character resource database construction, Dou Peng introduced the process and experience of building the Shaanxi-Gansu-Ningxia Border Region Red Memory Figure Database, providing references for red character database construction [21]. Sun Lietao proposed a digital service construction model for red education resources using the digitalization of the Zhou Enlai Library at Meiyuan New Village Memorial Hall as an example [24]. Second, concerning resource organization and application, Zhang Yunzhong et al. proposed a knowledge graph-based question-answering service framework for red historical figures, using resources of historical figures from the old Shanghai University as a corpus to build a platform that achieves fine-grained organization and knowledge services for red historical figure resources [3]. Liu Weili constructed a knowledge graph for figures of the First National Congress of the CPC, achieving semantic associations among elements such as figures, events, locations, and organizations, providing a reference for semantic organization of red character resources [25].

In summary, while preliminary achievements have been made in data organization, digitization, and special database construction for red character resources, considerable exploration space remains. First, most research perspectives consider red literature resource development holistically, with red characters serving as important elements or themes that require deeper excavation. Second, in constructing knowledge representation systems for red character resources,

most studies position resources at the data element level, making it difficult to build complex knowledge systems or convey the profound connotations of revolutionary culture, while immersive mining such as thematic clustering, semantic association, and knowledge discovery needs strengthening. Additionally, some scholars have recognized the importance of digital storytelling for red literature resource research, exploring narrative expression of red archives, narrative development models, and optimization strategies for multimodal red collection development [4,7,26,27]. However, exploration of narrative logic and general frameworks for red character resources, particularly empirical research, remains insufficient. Therefore, this study introduces narrative theory, aligns with resource development needs, designs a digital storytelling framework for red character resources, and conducts empirical research using Western Anhui red characters as a case study, providing references for knowledge-based and scenario-based development of multimodal red character resources.

3 Framework Design for Digital Storytelling of Red Character Resources

Drawing upon narrative theory discussions, framework construction, and research methodologies by scholars such as Barber J F, Fu Yaming, and Zhang Bin [28-30], and combining red character resource development needs with user experience, this study identifies essential elements, implementation processes, and application scenarios for digital storytelling of red character resources. Simultaneously, narrative theory is employed to collaboratively improve the core vocabulary of red character resource knowledge ontology, addressing limitations in current entity semantic association designs that rely primarily on resource "content" and "external" layers. Consequently, we develop a theoretical framework for knowledge-based development of red character resources oriented toward digital storytelling, comprising three layers: narrative resource integration, narrative element extraction, and narrative work generation. This framework provides theoretical guidance for subsequent empirical research, as illustrated in Figure 1 [Figure 1: see original paper].

3.1 Narrative Resource Integration Layer

The primary objective of the narrative resource integration layer is to integrate red character resources from diverse sources, types, contents, and structures, making them digital, textual, and resource-based to provide a data foundation for digital storytelling. The specific steps are as follows: Due to historical reasons, red character data sources are relatively dispersed with diverse resource types and contents. Accordingly, based on data structural levels, red character resources are categorized into unstructured data, semi-structured data, and structured data, upon which red character data preprocessing and resource database construction are completed. Noteworthy considerations include: (1) Original red literature constitutes unstructured data that requires digitization and textualization through character recognition or manual annotation; (2) Fur-

ther integration of relevant red character entries from knowledge encyclopedias and official websites, which feature high value, wide distribution, and multimodal characteristics; (3) Red character special databases have undergone high-level digital integration with relatively structured, standardized data containing less redundancy and higher credibility, making them directly usable as analytical materials.

3.2 Narrative Element Extraction Layer

The narrative element extraction layer primarily employs ontology and knowledge extraction technologies to process and generate minimal narrative units from red character resources. Building upon integrated narrative resources and referencing knowledge extraction workflows [31-33], the narrative element extraction process is divided into narrative modeling, element extraction, element fusion, and element storage, thereby addressing limitations in entity semantic association designs based on resource "content" and "external" layers. Specifically: (1) Element modeling completes the conceptual layer design for red character digital storytelling, namely red character resource ontology construction, which serves as the basic template for red character digital storytelling. By sorting and defining relevant terms and attributes of red character resources, it achieves standardization and normalization of domain knowledge. (2) Based on the red character ontology model, narrative element extraction from red character resources includes entity extraction, attribute extraction, and relationship extraction. Entity extraction identifies real-world objects with mutual distinctions, such as persons, times, locations, and events. Attribute extraction typically supplements attributes of a specific entity; for example, for event entities, it supplements location, time, or related persons. Relationship extraction identifies associations between different entities, which have directional characteristics. (3) Considering the multi-source heterogeneity of red character data, knowledge representation forms exhibit diversity, necessitating knowledge fusion including analysis, integration, and disambiguation of extracted elements and relationships. For instance, coreference issues arise when multiple descriptions refer to the same narrative element, such as different names for the same person or place name changes due to historical evolution. Ambiguity issues occur when narrative element names can refer to multiple real entities, causing unclear meaning. (4) Element storage involves converting extracted narrative elements, attributes, and relationships into SPO triple formats, such as <element, relationship, element> and <element, attribute, attribute value>, forming minimal knowledge units describing red character resources, which are then batch-imported into databases for storage.

3.3 Narrative Work Creation Layer

The narrative work creation layer primarily utilizes graph-based narrative templates to interact with general users and domain experts through narrative tools, ultimately creating and presenting red character resource narrative works. Its

core elements include narrative elements, narrative templates, narrative tools, and interaction modes [4,30,34,35]. (1) Narrative elements serve as basic units constituting narrative works, including but not limited to time, space, persons, events, and objects, where combinations of elements from different red character resources according to certain rules can present different red character-related stories. (2) Narrative templates, also called narrative structures, are relatively stable abstract structures composed of specific backgrounds, characters, events, and other red character resource narrative elements that guide users from one node to another, enabling association and combination of different red character story materials. (3) Narrative tools serve as media for user-resource interaction, effectively rendering and presenting narrative works. Appropriate tools can be selected based on different narrative structures and plots, such as knowledge graphs, GIS, and timelines for humanities research assistance, or game software and AR/VR for entertainment. (4) Interaction modes represent different ways users employ narrative tools. Compared with traditional narratives, digital storytelling possesses stronger interactivity, advocating the transfer of narrative process control to users, who can perform knowledge retrieval, question-answering, visual browsing, gaming, and knowledge discovery according to their needs.

4.1.1 Data Introduction and Acquisition

This study selects "Red Characters of Western Anhui" resources as the case for empirical research, which demonstrates feasibility, operability, and representativeness for three reasons. First, the Western Anhui region, located at the watershed between the Yangtze and Huai River systems with a strategically important geographical position connecting east and west, has produced rich red culture and revolutionary traditions, earning it the reputation as the "three cradles" of the Red Army, generals, and the republic. Consequently, Western Anhui red character data holds significant historical research value [36]. Second, the red characters in digital resources span long time periods and wide geographical distributions, including red heroes from multiple periods and locations with complete biographical descriptions and relatively abundant data volume. Third, the digital resources have been compiled and reviewed by authoritative departments, ensuring authority and accuracy. Given the large data volume, we used Python's Scrapy framework to crawl and parse multimodal data of red characters from the aforementioned resources, involving text, images, and audio-video link addresses, ultimately obtaining 676 red character entries stored as JSON files for subsequent processing.

4.1.2 Narrative Tool Selection

This study selects knowledge graphs as the narrative tool. On one hand, knowledge graphs possess mature theoretical and technical systems that can provide reference paradigms for digital resource knowledge aggregation of red character resources. On the other hand, as a graphical and associative knowledge repre-

sentation form, knowledge graphs not only enable networked narratives of red character resources but also provide powerful knowledge retrieval and visualization functions for human-computer interaction, facilitating historical scene reconstruction and knowledge discovery of red character resources.

4.2 Narrative Modeling

Current research primarily designs entity semantic associations from "content" and "external feature" layers, which adequately describes basic resource content but exhibits weak overall narrativity. The introduction of digital storytelling transforms red character resources from unidirectional, linear narratives to networked narratives, promoting narrative, fine-grained, and associative development of resources, and providing theoretical guidance for overall design and standardized organization of narrative modeling. This study employs the practical seven-step method for red character resource ontology construction.

(1) **Determine ontology scope and identify reusable ontologies.** First, the ontology is primarily used for knowledge organization of red character resources, with red characters as the data source to describe basic information and life resumes. Second, ontology reuse refers to the process of reusing classes or properties from existing publicly available mature ontologies, which helps reduce conceptual redundancy, lower ontology construction complexity and cost, and enhance ontology extensibility. By searching the Linked Open Vocabularies (LOV), we identified 33 person-related ontology vocabularies and filtered out 5 relevant ontologies based on their characteristics: rel (describing interpersonal relationships), bio (describing biographical information), foaf (describing internet social information), whois (describing personal profiles or resumes), and shoah (describing Holocaust-related concepts). Reusable classes include "Person," "Organization," "TemporalEntity," "Place," "Event," etc., and reusable properties include "birth," "death," "gender," "otherName," "event," etc.

(2) **Summarize core concepts and classification system for red character resource ontology.** The core purpose of digital storytelling is to maximize and most finely restore historical panoramas through different narrative logics using digital technology, necessitating the summarization and extraction of multi-featured, fine-grained conceptual and classification systems. Specifically: First, according to the textual characteristics of red character resources, core concepts include: Red characters as the narrative subjects; Life resumes of red characters, which generally contain educational experiences and war experiences, with related concepts including war events, major conferences, and red locations, from which elements such as persons, organizations, time, events, locations, and resources can be extracted. Second, we adopt a top-down construction approach, performing category division on completed top-level concepts to determine parent-child relationships between concepts, forming a conceptual classification system for the red literature ontology. Top-

level concept entities include “Person,” “Organization,” “Event,” “Place,” “Resource,” “OfficialPosition,” and “TemporalEntity.” “Organization” is divided into School, MilitaryInstitution, PartyOrganization, and GovernmentInstitution based on establishment nature; “Event” is divided into WarEvent, Conference, PatrioticMovement, and ImportantOrganizationEstablishment; “Place” entities are classified according to national administrative divisions; “Resource” is divided into PhysicalResource and DigitalResource based on carrier type. The classification system is shown in Table 1 .

(3) **Define semantic descriptions for red character resources, i.e., attribute relationship definitions.** Attribute relationship definitions must reference the virtual scenarios of red character resources, analyzing the specific environments in which figures lived, key nodes influencing event development, and their interconnections. Ontology attribute relationships consist of data properties and object properties. First, data properties, as shown in Table 2 , include: Person description, including last name, first name, courtesy name, gender, birth year, death year, education level, etc., expressing relatively complete basic information and background for red character narrative 铺垫; Resource metadata, describing external resource features; Time and location, where time properties decompose “year-month-day” for combination and location properties are divided according to national geographical administrative divisions (country-province-city-county/district) to accommodate complex spatiotemporal expressions in narratives. Second, object properties, as shown in Table 3 , define relationships between classes, including: Person-related properties such as birthDate, deathDate, nativePlace, studyAt, participateIn, join, which facilitate describing red characters’ life resumes, shaping complete character images, and integrating them into narratives; Resource-related properties such as referenceFrom, collectedBy, collectionLocation; Event-related properties including occurAt (time) and locatedIn (place). These semantic relationships facilitate interconnection of key nodes in the narrative system, enabling micro-narratives among entities such as “people-events-time-places-resources.”

(4) **Construct red character resource ontology.** Based on the aforementioned core vocabulary and attribute relationships, we use the ontology construction tool Protégé for ontology building and store it in RDF/XML format. The ontology structure is shown in Figure 2 [Figure 2: see original paper].

4.3.1 Algorithm Model Structure

First, considering the large number of triple structures with red characters as Subject in red character resources, this study breaks through the common relationship extraction approach of “relationship classification based on named entity recognition.” Referencing the language model probability graph frame-

work, we better handle the problem of "multiple Predicates corresponding to the same (Subject, Object)" and optimize triple extraction from red character resources. The probability model is shown in Formula 1 (y_1~y represent predicted output sequences):

$$P(y_1, y_2, ..., y_n|x) = \prod_{i=1}^{n} P(y_i|y_1, y_2, ..., y_{i-1}, x)$$

The specific approach: First, predict the first word through Encoder output x plus sequence start tokens, then predict the second word assuming the first word is known, and so on recursively until an end token appears, as shown in Formula 2:

$$P(s, p, o) = P(s)P(o|s)P(p|s, o)$$

Second, to further handle "many-to-many" problems (multiple Subjects, Objects, Predicates), this study adopts a "semi-pointer-semi-tagging" structure for reprocessing. Specifically, we implement BERT with semi-pointer-semi-tagging for direct triple extraction. The main logic: perform fine-tuning directly on the BERT pre-trained model, i.e., first use the pre-trained model to embed input text, then use the "semi-pointer-semi-tagging" method to predict the start and end positions of the corresponding Subject; subsequently, concatenate the predicted Subject embedding with BERT's hidden states to continue predicting Predicate and Object. Notably, multiple Predicates may exist during actual prediction, so for each Predicate, we separately construct a "semi-pointer-semi-tagging" structure to predict the start and end positions of the corresponding Object, with the model directly outputting triples in (S,P,O) format, where P is given and predicted according to annotation content. The model structure is shown in Figure 3 [Figure 3: see original paper].

4.3.2 Text Annotation and Model Training

First, we determine entities and relationships for extraction. Based on current data source characteristics and red character ontology classes and properties, we identify entities, attributes, and relationships requiring manual annotation. Entities including person, birth year, death year, other names, gender, native place, organization, event, and position are represented by tags {Per, Bir, Dea, Oth, Gen, Loc, Org, Eve, Ran}, while attributes or relationships including birthYear, deathYear, otherName, gender, nativeSpace, study, join, takePartIn, undertake, and comrade are represented by corresponding tags. Second, we perform data annotation. Training data uses open-source Western Anhui red character resume resources, and we use the Doccano open-source annotation tool to label segmented character data. After labeling, we obtained over 1,000 red character data entries, dividing the labeled samples into training, test, and

validation sets at an 8:1:1 ratio. Finally, we import the training and test sets into the model for training. Model hyperparameters are shown in Table 4 .

4.3.3 Model Evaluation

This experiment uses Precision, Recall, and F1-score to evaluate BERT and its variants ALBERT and RoBERTa, with results shown in Figure 4 [Figure 4: see original paper]. Precision, Recall, and F1-score are 0.832, 0.801, and 0.814 respectively. Among the three models, BERT performs relatively poorly, while ALBERT and RoBERTa show similar performance, with RoBERTa achieving the best results, demonstrating that the model can accurately extract triples from red character resources. Therefore, this study selects RoBERTa with "semi-pointer-semi-tagging" for batch triple prediction and storage of Western Anhui red literature resources.

4.4 Element Fusion

Entity fusion methods mainly include text similarity calculation [3], knowledge linking based on standardized dictionaries or third-party databases [37], and manual screening and replacement [38]. Drawing upon these approaches, this study adopts a hybrid method for entity fusion. The specific scheme is as follows: First, perform data preprocessing by removing spaces and special symbols from entity names and replacing abbreviations and nicknames with formal names. Second, employ appropriate knowledge fusion methods according to entity types. For location entity fusion, location entities exhibit variations in expression or administrative level, such as "Jinzhai," "Jinzhai County," and "Anhui Province Jinzhai County." To address this, we use a publicly available Chinese provincial-municipal-county regional dataset as a standardized vocabulary for knowledge linking with extracted location entities. For person and organization entity fusion, the CPC Organizational History Database developed by Shanghai Library serves as a relatively complete and standardized organizational history resource, which we use as a knowledge base for knowledge linking of person and organization entities. For remaining entities, we use the short text semantic similarity model provided by the hanLP toolkit to calculate similarity between two strings, which has been trained on large-scale corpora and demonstrates high accuracy. Taking "Liu-Huo Uprising" and "Liu-Huo Peasant Uprising" as an example, Figure 5 [Figure 5: see original paper] shows the model gives a similarity of approximately 0.967. Research suggests that when string similarity exceeds 0.85, they can be considered the same entity [3]. Finally, we conduct manual evaluation and correction of entity fusion results.

4.5 Knowledge Graph Generation

(1) Through the aforementioned steps, this study obtained 3,247 entities, 2,905 attributes, and 4,035 relationships, with detailed descriptions shown in Tables 5 and 6 .

(2) Triple data storage. This study selects the mainstream graph database Neo4j and uses the py2neo third-party library to batch-import the fused triple data into Neo4j in two forms: "entity, attribute, attribute value" and "entity, relationship, entity," ultimately generating a knowledge graph containing 676 Western Anhui red character entries, with a partial knowledge graph shown in Figure 6 [Figure 6: see original paper].

5.1 Knowledge Graph Retrieval and Visualization

Retrieval and visualization functions serve as basic knowledge graph applications. The Neo4j graph database integrates a graphical interface for retrieval and visualization, allowing users to directly obtain knowledge from the graph through dragging, zooming, and clicking operations, and to construct retrieval statements using Cypher language to obtain required graphs. Taking General Hong Xuezhi as an example, the retrieval statement for querying relevant nodes and relationships to form a resume graph is "MATCH p=(n:Person{name:'洪学智'})-[]->() RETURN p", with results shown in Figure 7 [Figure 7: see original paper].

Figure 7 shows the knowledge graph centered on the "Hong Xuezhi" person node on the left, and node types with quantities and attribute information for the "Hong Xuezhi" node on the right. The graph clearly reveals that General Hong Xuezhi (1) Basic information: Born in 1913 in Jinzhai, Anhui, died in 2006. (2) Positions held: Joined the CPC, served as Red Army company commander, regiment political director, Red Fourth Army political director, deputy principal of Anti-Japanese Military and Political University Fifth Branch, deputy chief of PLA General Logistics Department, etc. (3) Military experiences: Participated in the Lixia Festival Uprising, Long March, Pingjin Campaign, Liaoshen Campaign, Yangtze River Crossing Campaign, Korean War, etc. Through networked representation of the knowledge graph, red characters' basic narrative attributes and resume information are clearly presented, enabling searchers to quickly understand knowledge associations and providing references for finding red characters' appointment and military experience data.

5.2 Knowledge Graph Retrieval and Thematic Knowledge Discovery

Utilizing graph database retrieval and visualization functions not only enables rapid retrieval of explicit knowledge about red characters but also facilitates discovery of implicit knowledge in the knowledge graph based on multi-theme associative retrieval. This study selects narrative elements of red character resources as "raw materials" to achieve thematic retrieval and diverse digital storytelling of Western Anhui red character resources in historical contexts. For example: Person-centered knowledge discovery, exploring red characters' life stories and social network relationships; Event-centered knowledge discovery, such as investigating changes in Western Anhui red character groups during the

Agrarian Revolutionary War, Anti-Japanese War, and Liberation War periods; Organization-centered knowledge discovery, such as exploring red characters with service experience in the Fourth Army of the Chinese Workers' and Peasants' Red Army and examining their superior-subordinate and succession relationships; School-centered knowledge discovery, such as querying all "studyAt" relationships in the knowledge graph to study not only red characters' educational experiences but also talent cultivation situations of various schools.

This study demonstrates data using person-centered discovery, selecting revolutionary martyr Comrade Xu Baichuan's resume data as an example. The Cypher retrieval statement is "MATCH p=(n:Person{name:'徐百川'})-[]-() RETURN p" to expand the local graph of connected nodes, where purple nodes represent persons, peach nodes represent events, green nodes represent schools, red nodes represent organizations, and blue represents positions, with results shown in Figure 8 [Figure 8: see original paper]. Figure 9 [Figure 9: see original paper] shows the character attributes of the Xu Baichuan node.

As shown in Figures 8 and 9: First, Comrade Xu Baichuan was born in 1901, died in 1931, was from Hefei, Anhui, and was an early red hero of our Party. Second, he studied at the Whampoa Military Academy and the Sixth Guangzhou Peasant Movement Institute, establishing alumni relationships with Jiang Jingtang, Mao Zhengchu, Xu Yusang, and Yu Wuxi. Third, he participated in revolutionary campaigns including the Nanchang Uprising, Guangzhou Uprising, and Liu-Huo Uprising, establishing comrade-in-arms relationships with Zhu Yunshan, Wu Zhan, and Zhou Weijiong. Notably, Xu Baichuan shares similar experiences with multiple individuals such as Jiang Jingtang and Sun Deqing. For instance, Jiang Jingtang and Xu Baichuan belonged to the same era, both studied at the Whampoa Military Academy, and participated simultaneously in the Liu-Huo Uprising, fully demonstrating content associations in historical contexts. Upon verification, Xu Baichuan and Jiang Jingtang served as division commander and political commissar of the 33rd Red Division respectively, establishing a working relationship.

6 Research Conclusions and Implications

Based on consideration of the diverse characteristics of red character resources in terms of origin, type, content, and structure, this study constructs a digital storytelling framework for red character resources grounded in narrative theory and conducts empirical research using "Red Characters of Western Anhui" data, thereby achieving knowledge graph visualization, retrieval, and discovery through graph databases. This work holds both theoretical and practical value.

At the theoretical level: First, by connecting digital storytelling with digital resource development theories, the study addresses limitations in designing entity semantic associations based on red resources' "content" and "external" layers, forming an operational digital storytelling framework for red character resources that facilitates the progressive process and logical evolution from datafication-

contextualization-storytelling to data-academia-creation. Second, through thematic association and semantic organization of Western Anhui red character resources, we can generate static-dynamic knowledge graphs revealing various features including social network relationships, community dynamics, organizational characteristics, and military education of red characters in historical periods, promoting interpretation of data in historical contexts and assisting humanities scholars' research.

At the practical level: First, compared with traditional manual extraction and rule-based matching methods, the ontology and BERT-based batch triple extraction model improves knowledge extraction efficiency for red character resources, reduces data processing costs, and promotes dissemination and utilization of red literature. Second, by transforming static, planar, and dispersed red character data into dynamic, three-dimensional, and aggregated knowledge networks, the approach enhances interconnections among content resources, reduces difficulty in red character resource retrieval and utilization, and lowers research barriers for general users and cross-domain scholars.

References

[1] Xi Jinping. Make Good Use of Red Resources, Continue the Red Legacy, and Strive to Create New Achievements Worthy of History and the People[J]. Qiushi, 2021(19).

[2] Sun Shaodan. Research on Semantic Knowledge Organization of Historical Newspaper Resources from the Perspective of Digital Humanities[D]. Changchun: Jilin University, 2022.

[3] Zhang Yunzhong, Guo Dong, Wang Yage, Sun Ping. Research on Knowledge Q&A Service Framework for Red Historical Figures Based on Knowledge Graph[J]. Library and Information Service, 2021, 65(16): 108-117.

[4] Long Jiaqing, Pei Jiajie, Zhang Qiang. Research on Development Strategies for Multimodal Red Collection Resources Oriented Toward Digital Storytelling[J]. Library Science Research, 2023, 42(5): 38-46.

[5] Richardson B. Recent Concepts of Narrative and the Narratives of Narrative Theory[J]. Style, 2000, 34(2): 168-175.

[6] Richardson B. Digital Fiction and the Unnatural: Transmedial Narrative Theory, Method, and Analysis[J]. Poetics Today, 2022, 43(4): 773-776.

[7] Zhu Lanlan, Duan Yange. Application of Narrative Theory in Red Archives Resource Development: A Case Study of the Dabie Mountain Area[J]. Archives Science Study, 2023, 37(2): 95-102.

[8] Jia Junzhi, Cui Xiyan. Interoperability and Classification System Construction Between Person Ontology Vocabularies[J]. Journal of the China Society for Scientific and Technical Information, 2019, 38(7): 731-740.

[9] Chen Yaling. Research on Pre-Qin Person Knowledge Ontology Construction Method Based on CIDOC CRM[D]. Nanjing: Nanjing Agricultural University, 2020.

[10] Wang Nan. Research on Historical Person Ontology Construction and Query Inference[D]. Guilin: Guangxi Normal University, 2017.

[11] Kim M, Lee D, Lee S, et al. Inference on Historical Factions Based on Multi-layered Network of Historical Figures[J]. Expert Systems with Applications, 2020, 161: 113703.

[12] Bol P K. From Kinship to Collegiality: Changing Literati Networks, 1100-1400[J]. Journal of Historical Network Research, 2021, 5(1): 87-113.

[13] Zhang Yunzhong, Sun Ping. Construction and Visualization of Historical Cultural Celebrity Study Tour Footprint Knowledge Graph[J]. Library Journal, 2021, 40(09): 81-96.

[14] Shen Kejie, Huang Huanting, Hua Bolin. Person Knowledge Graph Construction Based on Public Resume Data[J]. Data Analysis and Knowledge Discovery, 2021, 5(7): 81-90.

[15] Song Xueyan, Cui Haonan, Liang Ying, et al. Knowledge Discovery Research on Celebrity Diary Resources from a Digital Humanities Perspective: A Case Study of Wang Shijie's Diaries[J]. Information Studies: Theory & Application, 2021, 44(6): 105-111.

[16] Yu Juan, Huang Hengqi, Xi Yunjiang, Zhu Zhengxiang. Research on Person Relationship Knowledge Graph Inference Method Based on Graph Database[J]. Information Science, 2019, 37(10): 8-12.

[17] Zhao Xing, Li Shuning, Xiao Yanan. Person Special Database Construction Based on Multi-source Data Fusion from a Digital Humanities Perspective: A Case Study of the "Tree Planter" Work in Shanghai Library's 2018 Open Data Application Development Competition[J]. Library Journal, 2019, 38(12): 4-12.

[18] Xia Cuijuan, Liu Wei, Chen Tao, Zhang Lei. Development Practice of Genealogy Linked Data Service Platform[J]. Journal of Library Science in China, 2016, 42(03): 27-38.

[19] Cuijuan X, Wei L, Lei Z. Implementation of a Linked Data-based Genealogy Knowledge Service Platform for Digital Humanities[J]. Data and Information Management, 2018, 2(1): 15-26.

[20] Niu Li, Liu Huilin, Gao Chenxiang. Research on Academic Celebrity Knowledge Base from the Perspective of Digital Memory[J]. Information Studies: Theory & Application, 2022, 45(9): 164-172.

[21] Dou Peng. Practice and Enlightenment of Red Character Database Construction: A Case Study of the Shaanxi-Gansu-Ningxia Border Region Red Memory · Figure Database[J]. Journal of Library Science, 2015, 37(12): 36-39.

[22] Wang Le. Value Mining of University Red Special Collections: A Case Study of Fudan University Library[J]. Journal of Academic Libraries, 2022, 40(1): 97-114.

[23] Liu Qianqian, Xia Cuijuan, Zhu Wuxin. Practice and Reflection on Red Literature Service Platform Construction from the Perspective of Red Culture Inheritance[J]. Journal of Information Resources Management, 2021, 11(4): 17-32.

[24] Sun Lietao. Practice Exploration of Digital Construction of Red Literature Resources: A Case Study of the Digital Construction of Zhou Enlai Library at Meiyuan New Village Memorial Hall[J]. New Century Library, 2021(04): 48-52.

[25] Liu Weili. Research on Knowledge Graph Construction for Figures of the First National Congress of the CPC[D]. Baoding: Hebei University, 2021.

[26] Zhou Linxing, Jiang Lu. Research on Narrative Expression in Red Archives Resource Development[J]. Archives Science Study, 2022, 36(4): 4-9.

[27] Wang Qiang. Protecting and Releasing Educational Value of Yellow River Red Culture Resources under Digital Humanities Context: Telling the Yellow River Story Well[J]. Archives Management, 2022, 41(6): 122-125.

[28] Barber J F. Digital Storytelling: New Opportunities for Humanities Scholarship and Pedagogy[J]. Cogent Arts & Humanities, 2016, 3(1): 1181037.

[29] Fu Yaming, Zhang Yongjuan, Liu Wei, et al. Digital Storytelling as a Digital Humanities Method: Current Status and Possibilities[J]. Library and Information Service, 2022, 66(14): 10-19.

[30] Zhang Bin, Li Zilin. A "Digital Storytelling-driven" Collection Utilization Model for Libraries, Archives, and Museums[J]. Library Tribune, 2021, 41(5): 30-38.

[31] Xu Xiaojuan, Sun Aihua, Shi Ruju, Chen Tao. Knowledge Ontology and Its Application for Western Anhui Red Literature from a Digital Humanities Perspective[J/OL]. Library Tribune. [2023-07-10]. http://kns.cnki.net/kcms/detail/44.1306.G2.20221226.1056.002.html.

[32] Song Xueyan, Zhang Weimin, Zhang Xiangqing. Research on Knowledge Graph Construction of Qing Dynasty Ritual Vessels Based on Archival Documents[J]. Library and Information Service, 2022, 66(3): 140-151.

[33] Xia Cuijuan, Liu Wei, Zhang Lei, et al. Design of Genealogy Ontology Based on BIBFRAME[J]. Library Tribune, 2014, 34(11): 5-19.

[34] XIA C, WANG L, LIU W. Shanghai Memory as a Digital Humanities Platform to Rebuild the History of the City[J]. Digital Scholarship in the Humanities, 2021, 36(4): 841-857.

[35] Niu Li, Gao Chenxiang, Zhang Yufeng, et al. Discovery, Reconstruction, and Storytelling: Paths and Methods of Archival Research from a Digital Humanities

Perspective[J]. Journal of Library Science in China, 2021, 47(1): 88-107.

[36] Zhang Xu. Analysis of Connotation, Characteristics, and Inheritance Development of Western Anhui Red Culture Resources[EB/OL]. [2023-07-10]. https://hswx.wxc.edu.cn/2019/0326/c5005a111030/page.htm.

[37] Huang Qianqian, Yang Jianlin. Analysis of Knowledge Graph Construction and Knowledge Service Application Based on Judicial Judgment Documents[J]. Information Science, 2022, 40(2): 133-140.

[38] Song Xueyan, Zhang Weimin, Zhang Xiangqing. Research on Knowledge Graph Construction of Qing Dynasty Ritual Vessels Based on Archival Documents[J]. Library and Information Service, 2022, 66(3): 140-151.

Author Contributions: Xu Xiaojuan: conceptualization, framework design, manuscript revision; Sun Aihua: data collection, programming, initial draft writing and revision; Chen Tao: revision suggestions.

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