Research Progress on Semantic Organization of Cultural Heritage: Postprint

Authors: Li Zhangchao, He Lin

Date: 2023-04-01T00:00:00+00:00

Abstract

[Purpose/Significance] This study investigates the current state of development in cultural heritage semantic organization, which holds significant reference value for cultural heritage research in China. [Method/Process] Employing systematic survey, case analysis, and statistical analysis methods, and based on summarized survey data, this paper reviews the current research status of semantic organization in cultural heritage projects from two perspectives: semantic organization approaches and knowledge services/tools, and analyzes key technologies for cultural heritage semantic organization across three dimensions: knowledge modeling, knowledge extraction, and knowledge mining and utilization. [Results/Conclusion] The study reveals that data interoperability, domain ontology standardization, personalized semantics, automated tools, and data copyright constitute the key factors for future development in cultural heritage semantic organization.

Full Text

Research Progress in Semantic Organization of Cultural Heritage

Li Zhangchao, He Lin
Department of Information Management, Nanjing Agricultural University, Nanjing 210095

Abstract: *[Purpose/Significance]* This study examines the current state of semantic organization development in cultural heritage, offering important reference value for cultural heritage research in China. *[Method/Process]* Employing systematic investigation, case analysis, and statistical methods, this paper synthesizes survey data to review the research status of semantic organization in cultural heritage projects from two perspectives: semantic organization methods and knowledge services/tools, while analyzing key technologies from three dimensions: knowledge modeling, knowledge extraction, and knowledge mining/utilization. *[Result/Conclusion]* The study identifies data interoperability, domain ontology standardization, personalized semantics, automated tools, and data copyright as critical factors for future development.

Keywords: cultural heritage; semantic organization; ontology technology; knowledge services
Classification Number: G255
DOI: 10.13266/j.issn.0252-3116.2020.07.001

Cultural heritage represents the cultural wealth created by humanity during social and historical development, possessing historical, artistic, and scientific value. In-depth research and excavation of cultural heritage help elucidate cultural essence, protect and inherit cultural heritage, and promote cultural exchange and mutual learning between China and other countries. With the development of information technology, an increasing amount of cultural heritage has been digitized, resulting in ever-growing resources of diverse types that provide a solid foundation for protection, research, and inheritance. However, due to varied data storage formats and the lack of unified description standards, these resources exhibit strong heterogeneity. How to organize massive heterogeneous cultural heritage resources in a computer-processable and understandable manner has become an urgent research problem. This study provides reference for the semantic organization of cultural heritage resources.

2. Research Status of Semantic Organization in Cultural Heritage Projects

With the development of humanities big data, cultural heritage data has evolved through digitization, structuring, and semanticization, gradually exploring deeper semantic relationships among knowledge points embedded in cultural heritage. Based on survey data summaries, this paper examines the current state of semantic organization in representative cultural heritage projects from technical, normative, and systematic perspectives, focusing on semantic organization methods and semantic knowledge services/tools. Drawing on platform materials from various projects, we summarize research themes, source resource types, semantic organization patterns, and semantic services provided by typical cultural heritage projects .

2.1 Survey Data Sources Overview

In recent years, digital humanities and humanities computing research has attracted widespread attention from scholars in computer science, library and information science, and history. Researchers have attempted to integrate intelligent information processing technologies with cultural heritage resource studies, conducting theoretical and practical research on semantic organization. Leveraging semantic web and machine learning technologies, scholars have performed knowledge extraction, organization, and service research on collections from libraries, museums, and archives, gradually accumulating relevant theories, methods, and technologies for cultural heritage information resource management. Against this background, this study investigates the current state of domestic and international cultural heritage semantic organization research based on typical projects.

(1) Main Research Focuses of Cultural Heritage Projects. Typical domestic and international cultural heritage projects primarily conduct semantic organization research around historical figures, historical events, and cultural collections. Some scholars design semantic description models for historical figures, such as CBDB (China Biographical Database), which focuses on Chinese historical biographical data covering kinship and social relationships [1]. The Song Dynasty Academic Semantic Network Platform organizes and reconstructs academic inheritance and kinship relationships among Song figures, while the Jinshi Database studies biographical data of over 100,000 successful imperial examination candidates since the 6th century. Some semantic portals focus on historical events, such as the Finnish WarSampo project, which studies World War II events related to Finland using war diaries, photo albums, and memoirs [2]. Other portals center on cultural collections, including artifacts, paintings, music, and architecture. For example, MuseumFinland studies collections from Finnish museums and companies like Nokia [3], while Europeana examines books, music, and artworks from European libraries, museums, and archives [4].

(2) Data Sources for Cultural Heritage Projects. Cultural heritage resources form the foundation and key to these projects. Investigation of project platforms reveals that current domestic and international projects primarily base their work on collections or digital resources from public cultural institutions, third-party organizations, universities, and research institutes, mainly categorized as self-built relational databases and multi-source heterogeneous databases.

Self-built relational databases refer to projects that construct databases by collecting, reading, and compiling original historical materials or utilizing secondary sources. These ensure data originality and authenticity but require extensive manual processing with low efficiency. For instance, CBDB' s early stage involved manual processing of official biographies, epitaphs, tomb inscriptions, local gazetteers, and biographical indexes [5], while the Jinshi Database' s data was largely compiled manually by Professor Gong Yanming and his team.

Multi-source heterogeneous databases integrate databases constructed by universities, research institutes, public cultural institutions, and publishers. While data acquisition is convenient, these databases exhibit strong multi-source heterogeneity, with the key challenge being compatibility between different data types and organization methods. For example, WarSampo' s data primarily originates from Finnish museums and archives [2], while CBDB includes data from published biographical indexes, chronologies, McGill University' s Ming-Qing Women' s Writings Database, and Academia Sinica' s Ming-Qing Archives Authority Database [5].

2.2 Semantic Organization Methods

Semantic organization methods achieve ordering, regularization, or systematization of cultural heritage resources according to their characteristics, evolving through four stages: unstructured, structured, linked, and intelligent, transforming cultural heritage data from documentary to quantitative and then to intelligent data.

(1) Unstructured Stage. Initially, cultural heritage data existed primarily in unstructured forms within various information resources, showing extreme irregularity and incompleteness. Early organization mainly relied on OCR recognition and manual entry for digitization, requiring substantial time and effort to read and organize resources while depending on historians' subjective induction and deduction. For example, the Jinshi Database used manual methods to collect and enter literature, constructing indexes such as "General Records of Chinese Imperial Examinations," "Song Dynasty Imperial Examinations Records," and "Ming Dynasty Imperial Examinations Records" [6].

(2) Structured Stage. The structured stage employs natural language processing and metadata technologies to perform word segmentation, part-of-speech tagging, and named entity recognition, achieving datafication, scaling, ordering, and regularization. For instance, CBDB uses computer text mining techniques like regular expressions to accurately extract historical figure data from digital cultural resources [7]. This stage enables word-level text extraction, generating numerous thematic indexes such as "Thirteen Classics Index" and "Twenty-Four Histories Place Name Index," as well as specialized dictionaries like "Chinese Historical Place Names Dictionary" and "Official Positions Dictionary."

(3) Linked Stage. The linked stage builds upon structured data by constructing ontology framework models to achieve concept relationship extraction and semantic association. As the goal of most current cultural heritage projects, this stage involves building thematic vocabularies and domain ontologies based on different research fields, questions, contexts, and historical material characteristics, forming linked data that enables association analysis across time, space, and themes. For example, WarSampo constructs a WWII ontology framework and linked open data cloud, allowing organizations and individuals to access, share, and semantically interlink structured cultural heritage data [2]. MuseumFinland builds an ontology model conforming to museum collection characteristics to achieve semantic linking of Finnish museum collections [3].

(4) Intelligent Stage. The intelligent stage utilizes data from previous stages for deeper semantic analysis from temporal, spatial, thematic, and network dimensions, gradually forming a new data-driven digital humanities research paradigm. This stage represents the target for some current projects. The Song Dynasty Academic Semantic Network Platform performs temporal sequence analysis, spatial geographic analysis, and social network analysis on Song figures based on CBDB data. Chen Pei-shih analyzed Qing dynasty administrative processing efficiency and procedures by examining citation patterns of administrative documents across different government departments [8].

2.3 Semantic Knowledge Services and Tools

Semantic knowledge services leverage data from cultural heritage projects according to user needs, employing tools such as precise and hierarchical knowledge retrieval and browsing, automated or manual annotation and segmentation tools, statistical analysis tools for character and word frequency based on time, place, and person features, and multi-dimensional visualization tools for temporal, GIS-based spatial, and relationship analysis.

In terms of semantic knowledge services, MuseumFinland provides faceted search and semantic linking/browsing services. The Jinshi Database offers surname statistics, dynasty statistics, and other multi-dimensional statistical functions [11]. WarSampo provides different types of war information perspectives for user retrieval and offers data recommendations including global database downloads, URI redirection, linked data browsing, SPARQL queries, data production, editing, authentication, and information visualization [2].

Regarding semantic tools, CBDB developed the CBDBRegexMachine tool based on regular expressions to facilitate large-scale knowledge mining from the database, helping researchers extract and visualize Chinese historical figure biographical data. Europeana provides precise search and filtering tools and opens REST API interfaces, allowing developers to use database data for application development. The Song Dynasty Academic Semantic Network Platform builds "knowledge graph structure exploration" and "relationship discovery" tools, helping users understand the structure and relationships of Song academic networks through visualization while supporting autonomous exploration and entity relationship discovery [9].

3. Key Technology Analysis for Cultural Heritage Semantic Organization

The development of semantic web and machine learning technologies provides effective solutions for knowledge system construction, fusion, and application in cultural heritage resources, elevating services from resource-level to knowledge-level that computers can process and understand. Specifically, this involves knowledge modeling, knowledge extraction, and knowledge mining.

3.1 Knowledge Modeling

Cultural heritage resources are massive, structurally chaotic, and highly heterogeneous. Knowledge modeling provides structured, model-based knowledge representation that enables structuring, semanticization, and sharing of cultural heritage knowledge, providing crucial support for knowledge services. Traditional knowledge modeling primarily used classification methods and thesauri, but with semantic web development, more projects now employ metadata and ontology technologies.

3.1.1 Thesauri. Thesauri developed rapidly from the 1960s, covering various domains. In cultural heritage, the Getty Institute systematically constructed humanities vocabularies based on international standards, covering art, architecture, bibliography, and archives, including the Art & Architecture Thesaurus (AAT), Conservation Thesaurus (CT), Cultural Objects Name Authority (CONA), Getty Thesaurus of Geographic Names (TGN), Union List of Artist Names (ULAN), and Getty Iconography Authority (IA) [12]. Other projects build specialized thesauruses, such as CBDB' s ancient Chinese official position and address tables, the Finnish National Library' s General Finnish Thesaurus (YSA) covering cultural heritage and other fields [13], the Library of Congress Subject Headings (LCSH), and Wuhan University Digital Humanities Research Center' s Dunhuang Mural Thesaurus (DMT) [14].

3.1.2 Ontology Modeling. Ontology models mainly include event ontology, person ontology, and person-event ontology. Event ontology uses event class structures as the main thread, comprising object, action, time, environment, assertion, and language expression elements [15], such as MuseumFinland' s historical event ontology based on chronologies that enables interoperability of metadata about collections, materials, people, locations, time, and institutions while describing different states of entities [3]. Person ontology uses people as the knowledge organization main thread to reveal social relationships and achieve formalization and structuring [16], such as CBDB' s ancient Chinese person ontology and BiographySampo's classified person ontology. Person-event ontology links people and events as the main thread, such as WarSampo adding event classes to person ontology to coordinate disparate person information into a series of events, enhancing model extensibility.

Commonly used semantic models include CIDOC CRM [17], EDM (Europeana Data Model) [18], BIBO [19], HOP [20], and SEM [21] as resource content description rules. These map different metadata forms to a common underlying ontology model, building domain ontologies that reveal relationships between digital heritage concepts and using minimal metadata patterns for precise event identification, disambiguation, and representation. Among them, CIDOC CRM, developed by the International Committee for Documentation, is one of the largest and most standardized ontology frameworks in cultural heritage, effectively promoting information source integration, transfer, and exchange. It is widely used in projects like WarSampo, Getty, World Heritage Foundation' s Arches [22], British Museum' s ResearchSpace [23], and classical art heritage portal CLAROS [24].

Ontology framework models worldwide are numerous and self-contained, while cultural heritage content, sources, languages, formats, and standards are diverse, inevitably causing difficulties for semantic integration. In China, most ontologies are project- or dynasty-specific, lacking standardized, universal, systematic frameworks like CIDOC CRM or EDM, particularly for traditional Chinese culture—an issue requiring attention in future digital humanities research.

3.1.3 Metadata. Metadata is the universal solution for most cultural heritage projects to transform historical texts (unstructured) into data (structured), enabling sharing, interaction, and integration of heterogeneous data from different sources [27]. Current focus centers on metadata semantic interoperability, including metadata extension and alignment.

Metadata extension expands existing metadata. For example, WarSampo manages and extends metadata based on CIDOC CRM, while Italy' s Protocollo Informatico extends DC metadata elements and maps project models to DC elements. Metadata alignment [28], such as MuseumFinland aligning heterogeneous data to DC-standard metadata, enables resource integration.

Metadata standards are fundamental [29]. Familiar, user-friendly standards facilitate data alignment and merging. The cultural heritage field commonly uses Dublin Core (DC) namespace and its extension principles to define element sets, terms, and attributes for collected objects [30], widely applied in libraries and government websites. Europeana' s EDM model defines classes and properties, revealing aggregation structure, resource relationships, event contexts, and thematic associations. Based on EDM, MIMO integrates European musical instruments [25], CARARE integrates archaeological and heritage digital content [26], and King' s College London' s SPQR integrates over 68,000 ancient Greek and Roman inscriptions.

Framework models for cultural heritage resources typically use W3C-recommended semantic web standards like RDF(S) and OWL, employing SKOS to build controlled vocabularies for resource description [31].

3.2 Knowledge Extraction Technology

Knowledge extraction in cultural heritage identifies knowledge points and their semantic relationships. Current advances have seen these technologies applied to cultural heritage information resources. Rule matching and machine learning are widely used methods, selected based on the information resources being processed.

3.2.1 Rule Matching Methods. Rule matching uses domain knowledge to analyze text features, build corresponding rules, and write regular expressions for pattern matching. This method suits texts with inherent syntactic regularities, such as classics and local gazetteers. In practice, CBDB' s historians and computer scientists design regular expressions for historical figure characteristics, with editorial teams verifying matched texts [7]. They develop targeted extraction rules based on local gazetteer and epitaph writing patterns. While mature and widely applied, rule matching suffers from subjectivity in rule formulation, potentially compromising logical systematicity and completeness, and shows poor portability across texts and domains. Some projects attempt automatic rule learning to address this.

3.2.2 Machine Learning Methods. Common machine learning methods include feature-based and neural network approaches. Historians annotate training texts according to agreed rules to create corpora for model training. The system can then process new texts. Feature vector-based methods achieve good results in historical research. Natural language processing enables text classification, automatic segmentation, named entity recognition, dependency parsing, and event extraction, using algorithms like SVM, decision trees, random forests, CRF, and deep learning techniques such as CNN and RNN. For example, Italy' s natural language processing laboratory created tools like LinguA, READ-IT, and T2K for the CHROME project, enabling annotation, segmentation, entity recognition, relation extraction, and visualization [32].

3.3 Knowledge Mining and Utilization

Linked (open) data, knowledge graphs, and multi-dimensional visualization methods like spatiotemporal, relationship, and cluster analysis enable semantic knowledge services, continuously improving service effectiveness and levels [10].

3.3.1 Linked (Open) Data. In practice, linked and open data technologies are most applied in libraries, museums, archives, and galleries ("four institutions"). CultureSampo, BookSampo, and WarSampo use the Sampo linked data publishing model integrating semantic models, metadata alignment, and domain ontologies to integrate collections from four institutions, universities, and publishers across three dimensions: culture (paintings, novels, comics), bibliography (authors, editors, publishers), and war (diaries, albums, memoirs). Europeana, British Museum' s ResearchSpace, Yale Center for British Art, and American Art Collaborative (AAC) also use CIDOC CRM to link collections across EU institutions.

Linked (open) data essentially uses the RDF data model, employing tools like OWL and SKOS to convert unstructured or differently-standardized structured data from content providers into standardized, structured data [33-34], enabling modeling, creation, coordination, and aggregation. The principle framework is shown in Figure 1 [Figure 1: see original paper].

3.3.2 Knowledge Graphs. Knowledge graphs represent concepts and semantic relationships, forming an interconnected, decentralized global knowledge network [35] that links large corpora and supports digital information systems, navigation, visualization, discovery, and semantic retrieval. Their significance lies not only in improving recall and precision but also in revealing hierarchical relationships for multi-dimensional visualization.

In practice, Italy' s ArCo project built 7 controlled vocabularies and 169 million triples from 820,000 cultural heritage entities' event and location information, using SPARQL to process and publish RDF data [36]. CultureSampo built a Finnish historical figure knowledge graph, while the Song Dynasty platform built a "Song academic lineage" knowledge graph and developed the RelFinder relationship discovery tool, converting knowledge graphs into relationship instances explainable in natural language for precise semantic retrieval of <X,Y,Z> triple relationships [9].

3.3.3 Multi-dimensional Visualization Technology. Visualization technology uses computer graphics and image processing to convert digital data into graphical representations across different dimensions. Applications in cultural heritage include cluster analysis, spatiotemporal analysis, and entity relationship visualization.

Cluster Analysis Visualization. Resources are divided into categories based on intrinsic similarities for thematic clustering. Europeana exemplifies this by revealing European history from themes like archaeology, art, fashion, industrial heritage, music, sports, and war [4]. Some projects organize resources by characteristics for visualization, such as MuseumFinland by material features and Europeana by color features.

Spatiotemporal Visualization. This includes temporal visualization (timeline-based organization of time-series data, such as WarSampo' s visualization of war events and biographies), spatial visualization (GIS-based geographic organization, like distribution maps of ancient Zhejiang pagodas and route maps of Tang Xianzu' s travels), and spatiotemporal integration (combining temporal and spatial analysis). Finland' s BiographySampo [37] visualizes person-events on maps with timelines, showing events' temporal and geographic distributions simultaneously, as illustrated in Figure 2 [Figure 2: see original paper] for Eliel Saarinen' s life events.

Entity Relationship Visualization. This reveals social relationships in cultural heritage resources, where nodes represent entities (people, places) and edges represent relationships, forming complex social networks. The Song Dynasty Academic Semantic Network Platform exemplifies this, integrating 22 entity relationships from CBDB data including student relationships (kv:hasStudent), children relationships (schema:children), and native place relationships (kv:nativePlace). Figure 3 [Figure 3: see original paper] shows relationship discovery among Wang Anshi, Sima Guang, and Su Shi using the platform' s RelFinder tool [9].

4. Existing Problems and Future Development Trends

4.1 Data Interoperability Issues

Due to the multidisciplinary and heterogeneous nature of cultural heritage data, current organization mainly achieves simple merging of concepts and URIs in ontologies. With diverse data sources, formats, and standards, semantic interoperability remains difficult. Semantic web technologies involving ontologies and metadata provide universal standards and frameworks for data interoperability. Future goals should shift toward true concept integration, ontology matching, and semantic entity association, considering whether data was created by different methods or users with different scientific backgrounds. Semantic technology rules for fuzzy reasoning using multidisciplinary information and probabilistic description logic, as well as reasoning for dynamically evolving data, are necessary for deep semantic mining and multi-dimensional concept fusion.

4.2 Ontology Standardization Issues

Ontology construction remains a hot topic in cultural heritage. While most projects extend or adapt CIDOC CRM to improve semantic organization efficiency, this can lead to syntactic defects, conceptual inconsistencies, and semantic problems. Additionally, different institutions or domains building ontologies with different thesauri or standards creates isolated, incompatible semantic issues. Solutions require expanding ontology construction participants to include more individuals and organizations while enhancing open access levels to strengthen ontology standardization. Ontology mapping, alignment, and sharing based on logical principles can also improve semantic interoperability.

4.3 Multilingual Knowledge Expression Issues

Cultural heritage exhibits obvious multilingual characteristics, with most institutions storing data in local languages. While the open data movement encourages public data provision, ensuring publication in languages beyond local ones remains challenging. For worldwide sharing and use, multilingual issues must be addressed from system and model development origins, including ontology multilingualization, metadata standard multilingualization, and multilingual knowledge representation, acquisition, and translation. Future projects will see further development of localization, language processing, translation, representation, and natural multilingual data management technologies.

4.4 Personalized Semantic Issues

Personalized semantics are necessary for expressing subjective opinions and reasoning. Cultural heritage objects with specific attributes often hold different aesthetic or interpretive meanings for people with different backgrounds, interests, and purposes. Controversial heritage may receive mixed evaluations. The key lies in providing personalized semantic services including personalized homepages, semantic hyperlinks, personalized retrieval, and navigation [38]. Future trends include semantic reasoning technology that derives different semantics and provides personalized knowledge services based on semantic relationships, whether using description logic or semantic knowledge extraction.

4.5 Automated Tool Construction Issues

Currently, only a few projects develop automated tools during semantic organization. Cultural heritage institutions should use semantic technologies to process data as linked data while developing user-friendly automated tools that integrate technical workflows and handle large data volumes, such as automatic annotation, segmentation, and visualization tools. Although some tools exist, their automation level remains low. Enhancing utilization of semantic web in cultural heritage organization requires improving tool automation and convenience.

4.6 Data Copyright Issues

The popularity of digital humanities and open data movements has expanded digital humanities resources, but unresolved data copyright issues hinder information distribution and sharing by museums and libraries. Solutions can draw from existing intellectual property protection measures at legal, institutional, and licensing levels. Metadata technologies and standards, as well as blockchain technology, can also protect data copyright.

References

[1] FULLER M. An account of the structure of the China biographical database [EB/OL]. [2019-11-10]. https://projects.iq.harvard.edu/chinesecbdb/%E8%B3%87%E6%96%99%E5%BA%AE

[2] WarSampo: Finnish world war ii on the semantic Web [EB/OL]. [2019-12-03]. https://seco.cs.aalto.fi/projects/sotasampo/.

[3] HYYONEN E, JUNNILA M, KETTULAS S, et al. Finnish museums on the semantic Web: the user' s perspective on museumfinland [C]//Proceedings of Museums and the Web 2004. Arlington, 2004.

[4] About EUROPEANA [EB/OL]. [2020-01-30]. https://www.europeana.eu/portal/en/about.html.

[5] CBDB 资料来源与涵盖范围 [EB/OL]. [2020-01-30]. https://projects.iq.harvard.edu/chinesecbdb/%E8%B3%87%

[6] Gong Yanming. Reconstructing 40,000 Song dynasty jinshi archives—introduction to "Song Dynasty Imperial Examinations Records"(14 volumes) by Gong Yanming and Zu Hui [EB/OL]. [2020-01-30]. http://rwsk.zju.edu.cn/2015/1112/c2039a170378/page.htm.

[7] Fu Junmai, FULLER M A. Chinese biographical database user guide (Chinese version) [EB/OL]. [2019-10-25]. http://172.16.20.58/cache/2/03/projects.iq.harvard.edu/6ed82dd42b5705

[8] Chen Pei-shih. Information technology and historical document analysis [D]. Taipei: National Taiwan University, 2011.

[9] KVISION. Song dynasty academic semantic network—relationship discovery (RELFINDER) [EB/OL]. [2019-10-25]. http://dh.kvlab.org/cbdb_{kg}/.

[10] Zuo Dan, Ou Shiyan. Review of semantic description and organization of humanities information resources [J]. Library Tribune, 2019, 39(8): 21-31.

[11]《历代进士登科数据库》产品介绍 [EB/OL]. [2019-10-15]. http://examination.ancientbooks.cn/docDengke/dkHelp

[12] Getty vocabularies [EB/OL]. [2019-10-20]. http://www.getty.edu/research/tools/vocabularies/index.html.

[13] YSA—General Finnish thesaurus [EB/OL]. [2019-10-20]. https://www.kansalliskirjasto.fi/en/node/167.

[14] 敦煌壁画主题词表项目介绍 [EB/OL]. [2019-10-20]. http://dh.whu.edu.cn/dhvocab/dhresource/html/intro.html

[15] Liu Zongtian. Event ontology architecture [J]. Computer Science, 2012, 39(2): 45.

[16] Wang Nan. Research on historical figure ontology construction and query inference [D]. Guilin: Guangxi Normal University, 2017.

[17] CIDOC CRM [EB/OL]. [2019-10-20]. http://www.cidoc-crm.org/.

[18] General EDM factsheet and presentation [EB/OL]. [2019-10-20]. https://pro.europeana.eu/resources/standardization-tools/edm-documentation.

[19] Bibliographic ontology [EB/OL]. [2019-10-20]. https://en.m.wikipedia.org/wiki/Bibliographic_{Ontology}

[20] STIVERS C. The ontology of HOP in dark times [J]. Administrative theory & praxis, 2008, 30(2): 225-239.

[21] VAN H W R, MALAISI V, SEGERS R, et al. Design and use of the simple event model [J]. Social science electronic publishing, 2011, 9(2): 128-136.

[22] What is arches [EB/OL]. [2019-10-20]. https://www.archesproject.org/what-is-arches/.

[23] ResearchSpace [EB/OL]. [2019-10-20]. https://www.researchspace.org.

[24] About CLAROS [EB/OL]. [2019-10-20]. http://explore.clarosnet.org/XDB/ASP/clarosHome/about.html.

[25] About MIMO [EB/OL]. [2019-10-20]. http://www.mimo-international.com/MIMO/about-mimo.aspx.

[26] Carare association [EB/OL]. [2019-10-20]. https://www.carare.eu/about/.

[27] HYYONEN E. Cultural heritage linked data on the semantic web: three case studies using the sampo model [C]//VIII Encounter of integral management of information in cultural centres. Vitoria-Gasteiz: Artium, 2016: 19-20.

[28] RUOTSALO T, HYYONEN E. An event-based approach for semantic metadata interoperability [C]//The international semantic web conference. Heidelberg: Springer, 2007: 409-422.

[29] Wang Hongxia, Su Xinning. E-government information resource organization model based on metadata [J]. Information Studies: Theory & Application, 2007, 30(1): 116-121.

[30] Wang Ping, Huang Xinping. Research on semantic fusion methods for digital cultural resources based on linked data [J]. Library and Information Service, 2016, 60(12): 29-37.

[31] HYYONEN E, LINDQUIST T, TORNROORS J, et al. History on the semantic web as linked data–an event gazetteer and timeline for WWI [EB/OL]. [2019-10-25]. https://seco.cs.aalto.fi/publications/2012/hyvonen-et-al-ww1-cidoc-2012.pdf.

[32] T2K (Text-To-Knowledge) [EB/OL]. [2019-10-25]. http://www.italianlp.it/demo/t2k-text-to-knowledge/.

[33] Xia Cuijuan, Liu Wei, Zhao Liang, et al. Linked data publishing technology and implementation—taking Drupal as an example [J]. Journal of Library Science in China, 2012(1): 51-59.

[34] BIZER C, HEATH T, BERNERS-LEE T. Linked data: the story so far [C]//Semantic services, interoperability and Web applications: emerging concepts. Hershey: IGI Global, 2011: 205-227.

[35] HASLHOFER B, ISAAC A, SIMON R. Knowledge graphs in the libraries and digital humanities domain [EB/OL]. [2019-10-25]. https://arxiv.org/pdf/1803.03198.pdf.

[36] GROUNDA I. ArCo: the Italian cultural heritage knowledge graph [EB/OL]. [2019-10-25]. https://www.groundai.com/project/arco-the-italian-cultural-heritage-knowledge-graph/1.

[37] BiographySampo [EB/OL]. [2019-10-25]. http://biografiasampo.fi/henkilo/p992/kartat.

[38] Ju Yanhui. Research on personalized semantic portal knowledge organization [J]. Information Science, 2014(3): 53-56, 67.

Author Contributions: Li Zhangchao: paper writing and revision; He Lin: topic selection and framework design, providing revision suggestions.

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