Unified Knowledge-based content management for digital archives

Tine-Yu Hsu
Department of Information, National Museum of NaturalScience, 1 Kuan Chien Rd., Taichung, Taiwan 40453, R.O.C.

Hao-Ren Ke
National Chiao Tung University Library 1001 Ta Hsueh Rd., Hsinchu, Taiwan 30050, R.O.C.

Wei-Pang Yang
Department of Computer & Information Science, National Chiao Tung University, 1001 Ta Hsueh Rd., Hsinchu, Taiwan 30050, R.O.C.

Abstract

Worldwide museums are facing the challenge of applying knowledge management to digital archives projects to unify content acquisition, creation, organization, reusing, and publication among individuals, departments, and projects for long-term business prospective. We propose a unified knowledge-based content management approach that consists of a unified knowledge content process, a knowledge-based content management system and multi-layer reusable knowledge content structures. EER model was applied to effect integrate the global knowledge-based content management system and express multi-layer reusable knowledge content structures in digital archives of museums. A practical system was implemented to provide a unify knowledge management environment that the knowledge content creation, management, sharing, reuse, and accessing among all projects, specialists, content, technologies, and applications can be realized.

1. Introduction

Numerous digital archives projects in museums are being undertaken all over the world. The preservation of cultural heritage and the creation of value-added services are the major purpose of these projects. They are considered as the prerequisite for developing digital museums

which extend and support the construction of a traditional museum's multifunctional knowledge-based content in the fields of academic research, exhibition, education, and entertainment (Alonzo, 2000; MacDonald, 2000; Mason, 2002; Meli, 2003; Rosenberg, 2000).But most of them incur some problems in general as follows:

- Contents and resources created by separate individuals, groups, departments, or projects are dispersed which can't be managed centrally and integrally and reused by others.
- Platforms and application systems are developed by each project which are repetitious and lack integration.
- Relative applications are implemented independently so the contents and system modules can't be shared and reused.
- Digital content is managed in data or information level for preservation, but not in knowledge level for sharing and reuse.
- Lacking long-term developing perspective for the entire museum; therefore, the extension of management functionalities, system architectures and the creation of value-added services are limited.

Content silo trap (Rockley, 2003) is the major cause of the above problems, owing to the content is created by individuals working in isolation from other individuals within the institute. Walls are erected among content domains and even within the same content domain. Content silo trap leads to content being created and re-created, often with changes or differences in each iteration. The effect of silos includes poor communication, lack of

sharing, reduced awareness of other initiatives, lack of standardization and consistency, higher cost of content creation, management, and delivery, and content user suffers, too.

In this paper, we propose a unified knowledge-based approach that combines the unified content strategy and knowledge management to solve the above problems. Besides those benefits shown by Rockley such as faster time to publish, better use of resources, reduced costs, improved quality and usability of content, increased opportunity to innovate, improved workplace satisfaction, and increased customer satisfaction. Additional benefits for museums include:

- A global view of knowledge-based content creation, management and publication that tightly connects the collaboration among knowledge and IT specialists for all projects is employed.
- Provide a well-defined consistent domain knowledge system to unify and correlate all content specialists under a common conceptualization.
- Construct a standard multi-layer knowledge content structure for specialists to fully express and create explicit and implicit knowledge.
- Knowledge contents created from projects, specialists and applications can be managed and integrated centrally and shared and reused by others.
- Knowledge management concepts and techniques are applied for long-term prospective of the developing digital museum applications.

. This paper is organized as follow. We propose the unified knowledge management system framework, in the Section 2. The unified content management process, the knowledge-based content management system, and the multi-layer reusable knowledge-based content of the unified knowledge content management system framework are described in the Section 3, Section 4, and Section 5 respectively. In Section 6, we apply the Extended Entity-Relationship (EER) modeling tool to conceptually design the knowledge system that entails content across multiple domains of a museum , and multi-layer reusable knowledge content structure. We demonstrate a practical implementation of our approach in Section 7 and make a conclusion and future research directions in Section 8.

2. System framework

The components of the unified content strategy (UCS) described by Rockley (2003) entail unified process, content management system (CMS), and reusable content. We adopt Rockley's idea and extended to the considerations of knowledge management approach, we proposed the unified knowledge-based content management framework (see Figure 1) which consists unified knowledge content process, knowledge-based content management system, and multi-layer reusable knowledge content.

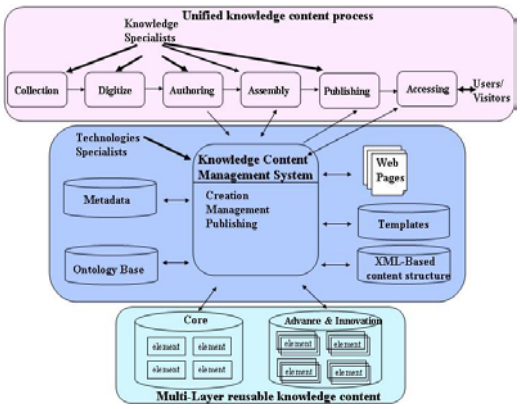

Figure 1 Framework of the unified knowledge-based content management approach

The unified knowledge content process acts as the common workflow among participants and projects that includes knowledge content collection, digitization, authoring, assembly, publishing, and accessing steps. The knowledge-based content management system works as the integration role of whole system which includes the creation subsystem for constructing metadata, contents, and ontology, the management subsystem for managing the whole knowledge content and resources for creation and

publishing, and the publishing subsystem on duty to transfer the authoring content into publishing structure and web pages for all departments and projects. The multi-layer reusable knowledge content structure defines the spectrum of whole knowledge content structure from core elements, the basic elements, to advanced and innovative elements, a multimedia document, a knowledge unit, a knowledge group, and a knowledge network for all participants to follow. A core element is organized with a multimedia object and semantic metadata which is the foundation of knowledge content. The advanced and innovative elements are further authoring manually or inference automatically from existing contents.

3. Unified knowledge content process

The unified knowledge content process provides common and collaborative workflows among knowledge specialists, IT specialists, and users. From the viewpoint of knowledge specialists, they need a standard, consistent, repeatable, sharable, and transparent environment to collect, digitize, author, and assemble knowledge content. From the viewpoint of IT specialists, they need an integrated and collaborative system developing environment to design an efficient, automatic, scalable, and interoperable system for supporting content creation, management, and publishing functions. Users can access enterprise, dynamic and systemic knowledge content via ontology browsing and metadata search.

As described in Rockley (2003), the content life cycle encompasses the creation, review, management, and delivery phases. To support a knowledge-based content manage system for a digital archives project, the content life cycle comprises knowledge content collection, digitization, authoring, assembly, management and publication phases (see Figure 2) The major tasks of each phase in the unified knowledge content processes are summarized as follows：

(1) Collection phase

- For expressing the knowledge concepts for a particular application and user group, the target version must be predefined and specified.
- The original materials such as slides, photographs, audio tapes, video tapes, or documents are prepared under a particular standard knowledge content structure.

(2) Digitization phase

- The original materials are digitized and edited into digital images, videos, audios, and text objects under standard formats.
- Some digital objects can be captured from the artifacts directly.
- The unique object identification is assigned for each multimedia object.

(3) Authoring phase

- Each multimedia object is interpreted using metadata with semantic content description to become a core knowledge element.
- Core knowledge elements are stored into enterprise core knowledge element repository that can be shared by all authors.

(4) Assembly phase

- Core knowledge elements are reused and organized into a multimedia document.
- Multimedia objects and multimedia documents are organized into a knowledge unit.
- Knowledge units are grouped to create a knowledge group.
- Multimedia documents, knowledge units, and knowledge groups are stored into advanced knowledge repository that can be reused and shared among enterprise authors and applications.

(5) Publication phase

- Every knowledge content must be approved by some reviewers before it is delivered to users.
- Each core and advanced knowledge element

will be transformed into XML-based content structure automatically.

- Authors can assign a presentation template to each knowledge element; then the Web page can be generated from the combination of the XML-based content and presentation template.

(6) Accessing phase

- All pieces of knowledge content created by the specialists and used and shared among applications are published through a unified knowledge portal.
- Users can access knowledge content through a semantic ontology-based browsing interface to have the common view with specialists.
- A metadata query interface is designed for each layer and each domain of knowledge base.

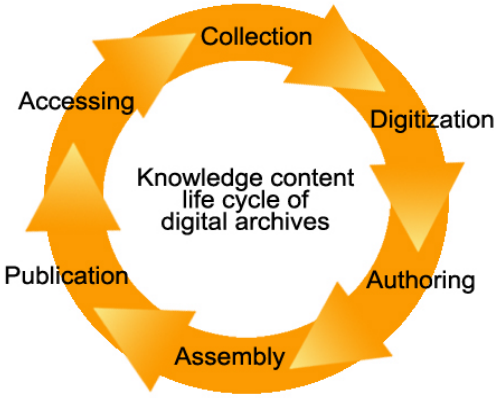

Figure 2 The Life cycle of knowledge content of digital archives

4. Knowledge-based content management system

A content management system (CMS) is responsible for the collection, management, and publication of chunks of information known as content components (Boiko, 2002). Knowledge is a kind of content that can be managed.Content management can be considered as an infrastructure to amass and distribute knowledge. If knowledge management is the collection, management, and distribution of what an organization “knows”, a CMS can prove the best platform behind a knowledge management system to build knowledge content management system (KCMS). With both knowledge content accumulation and authoring functionalities in museums, the collection system in content management can be substituted into the creation system.The functionalities of the creation system, the management system, and the publishing system of KCMS are as following：

(1) The creation system

- Assign the unique knowledge content identification (KCID)
- Edit domain knowledge ontology
- Edit metadata for describing each knowledge content element
- Author fundamental (core) knowledge content elements
- Assemble complex (advanced and innovative) knowledge content elements

(2) The management system

- Manage metadata repositories
- Manage thesaurus repositories
- Manage enterprise knowledge content repositories
- Manage a single global view of ontology base
- Manage a sharable authoring template base

(3) The publication system

- Generate the publication structure of knowledge content
- Generate ontology-based browsing interfaces
- Generate metadata-based query interfaces
- Publish into the Web from knowledge content and selected templates

5. Multi-layer reusable knowledge content

The multi-layer reusable knowledge content defines core, advanced, and innovative content structures for knowledge specialists to express both tacit and explicit knowledge content. A core knowledge element is represented by a

multimedia object of artifacts, such as a text, an image, a video, an audio, or an animation and accompanies with the semantic annotation given by specialists. A set of core knowledge elements then can be organized into multimedia documents, knowledge units, knowledge groups, and knowledge networks to form an advanced knowledge content. The innovative knowledge can be discovered or inference from core and advanced knowledge content.

In order to provide a reusable and unlimited content creation and organization model shared among applications and knowledge specialists, a multi-layer knowledge content structure is defined as follows (see Figure 3):

- Core knowledge element

  A core knowledge element is a fundamental individual multimedia object (image, audio, video, text, animation, and 3D object) associated with metadata and interpretation context.

- Advanced knowledge element

  An advanced knowledge element is made of core knowledge elements that can be reused and organized to form a multimedia document, a knowledge unit, and a knowledge group. A multimedia document is set of core knowledge elements for describing a topic relating to an artifact. A knowledge unit is used to organize all relating topics for an artifacts. A knowledge group is organized by a set of multimedia documents and/or knowledge units with same characteristics. The relative relationships between any pair of the above various structures can be specified in each structure to form a knowledge network.

- Innovative knowledge element

  Innovation knowledge elements can be generated by automatic inference or discovered from core and advanced knowledge repositories. Innovative knowledge elements have the same structures as advanced knowledge elements. The difference is that the former are created, classified and organized automatically, but the latter manually

Sets of core knowledge elements are stored in the knowledge content management system and are referenced for inclusion in a virtual advanced and innovative knowledge element. In this way, advanced and innovative elements can appear in multiple places to meet various needs, but reside in only one. In this paper, core knowledge and advanced knowledge sets are created in the authoring process to manually construct the knowledge base. Dynamic discovery mechanism will be developed to generate more extended knowledge content in the near future. This multi-layer content structure can be reused by various application packages and shared among participants transparently.

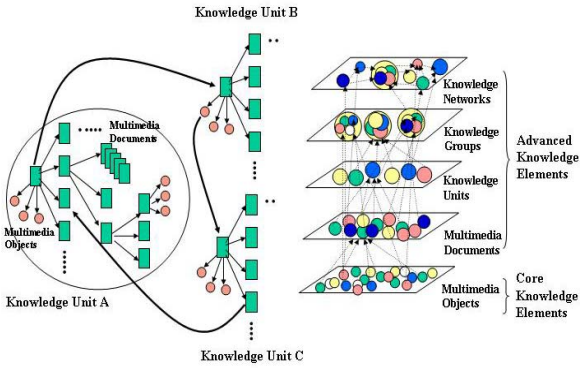

Figure 3 The multi-layer reusable knowledge content structure

6. Conceptual modeling

Conceptual modeling is the process of identifying the relevant concepts of the real world with an abstract model. Conceptual modeling aims at integrating different views of an enterprise into one global and consistent model where entities and relationships are explicitly defined. The conceptual modeling of a unified knowledge-based content management system for digital archives in our approach is designed to construct both enterprise domain knowledge system and multi-layer reusable knowledge content structures by using a thorough syntax, a semantic tool, and models that concretely express

and interpret them.

EERM owns the features of ERM and OOM, including aggregation abstraction, generalization abstraction, and association abstraction. Aggregation abstraction defines a PART-OF relation between an entity and its components. Generalization abstraction defines a subset or IS-A relationship between entities, and establishes a hierarchy from a generic entity to its subsets. Association abstraction defines a multi-valued feature of attributes. Due to the many advantages of EERM, it has been applied for conceptual modeling of database applications (Batini, 1992; Engels, 1993; Si-Said, 2002). We also successfully apply EERM for conceptually modeling the multimedia databases of museum applications in our previous research (Huang and Hsu, 1999). We we still use EERM as a conceptual modeling tool in this paper.

6.1 Conceptual modeling of enterprise domain knowledge system

The conceptual modeling of a knowledge system for across domains and applications in a museum can be summarized as the following features.

- Use a top-down design approach to construct the global knowledge system across domains.
- The relationships and constraints can be constructed between entities within a domain or across domains.
- The attributes of various knowledge contents can be specified as metadata annotated by specialists of each domain.
- For supporting efficient administration and personalized service, profiles of specialists and users must be specified.

After finishing the system conceptual design, the knowledge hierarchy and relationships in a domain or between domains could be organized to construct the global knowledge system. The knowledge systems can be constructed for the requirement of applications and also can be viewed as the knowledge ontology system to represent them. The knowledge ontology system will act as the role of common view among specialists and users. The knowledge elements entity contains a set of core and advanced knowledge elements that we have specified in Section 5. We will show their conceptual modeling process in next section. The Conceptual Modeling of the global knowledge system for a museum is shown in Figure 4.

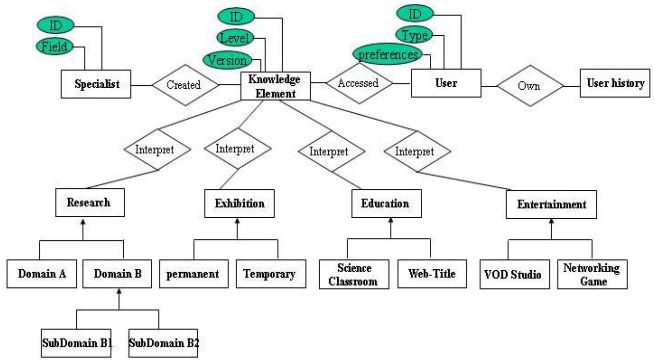

Figure 4 The Conceptual Modeling of the enterprise domain Knowledge system

6.2 Conceptual modeling of multi-layer reusable knowledge content structure

In order to provide a set of formal content structures from elementary to complex one which provide specialists to express and interpret tacit and explicit knowledge around them, in Section 3.2 we have shown the multilayer knowledge content structure that consists of core knowledge elements, advanced knowledge elements, and innovative elements. Multimedia Objects such as digital images, sounds, films, literatures, pictures, animation, and 3D models are the most basic units for constructing a core knowledge element. By reusing core knowledge elements, different scales of complex knowledge content can be organized.

A multimedia document is organized by a set of multimedia objects to represent a basic unit of a knowledge concept of a specific subject. Specialists can further organize a set of related multimedia documents to form a knowledge unit for some topics in a hierarchical structure. For the functional purpose, knowledge units can be

classified into research, exhibition, and education types. A set of multimedia objects, multimedia documents, and knowledge units with the same characteristics and properties can be clustered to form a knowledge group for applications.

Any multimedia object, multimedia document, knowledge unit, or knowledge group can have reference links to each other within a domain or across domains to meet the various applications' and user's need; thus a knowledge network is generated. The multimedia document, knowledge unit, knowledge group, and knowledge network with the characteristics to be organized and reuse the core knowledge element are called advanced knowledge elements. The core knowledge elements and advanced knowledge elements can be dynamically and systematically excavated to lead to an unlimited extension of larger knowledge units, groups and networks. We will develop this mechanism in the next stage. The multi-layer reused knowledge content structure in EERM model is shown in Figure 5.

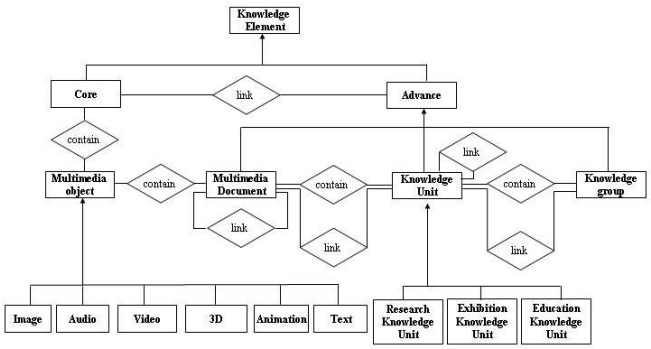

Figure 5 The conceptual modeling of Multilayer knowledge elements

6. A practical implementation of digital archives project

A unified knowledge content management system for a digital archives project has been practiced in National Museum of Natural Science (NMNS) of Taiwan. NMNS has been opened since 1987. Near 650 thousand specimens are collected by NMNS.

NMNS is one of the institutional projects of the National Digital Archives Program of Taiwan, from 2002 to 2006.The project is executed in a global view of entire museum resources including applications (academic research, exhibition, education, entertainment, and so on), participants (knowledge specialists, IT specialists, and users), content (core and advanced knowledge element repository), and technologies (content creation, management, publishing) for long-term prospect. The content experts of all collection and research departments (Zoology, Botany, Geology, Anthropology, Exhibition, and Education), and the system developer of the information department are cooperative in a collaborative way based on a unified processes and knowledge-based content management.

The system architecture is organized into three layers (see Figure 6). The bottom layer is designed for creating and managing the core knowledge element repository with vast multimedia objects. The middle layer is the knowledge creation and management layer where advanced knowledge elements are created and managed. Those elements must be organized by specialists to express the knowledge concept. Therefore these organized content can be assigned to a presentation format and applications. The top layer is the knowledge publishing and accessing layer where knowledge content with presentation format will be transformed into web pages. Users can access the content through ontology-based browsing and metadata search query interface.

The KCMS includes the creation system, the management system, and the publishing system. Each system has been designed to include the functions that are specified in Section 3.1. Figures 7 – 14 show the screen of major steps during the process of multi-layer knowledge content construction.

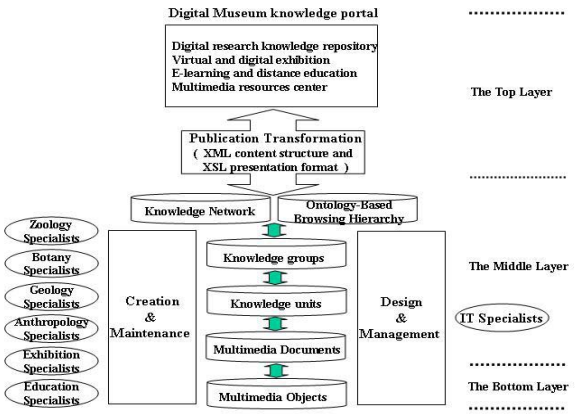

Figure 6 The implementation architecture of knowledge content creation, management, and publishing

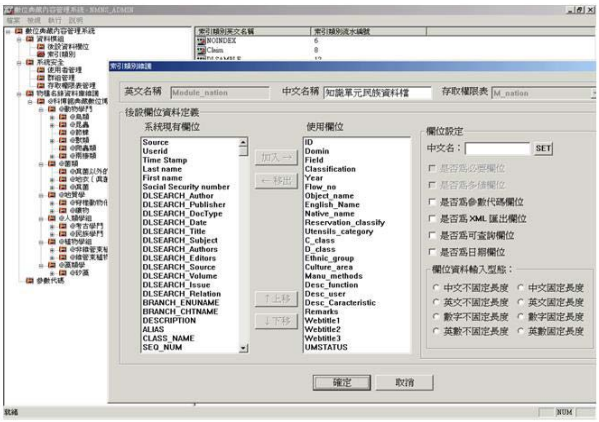

Figure 7 Domain ontology creation and maintenance

Figure 8 The management of XSL files for presentation formats

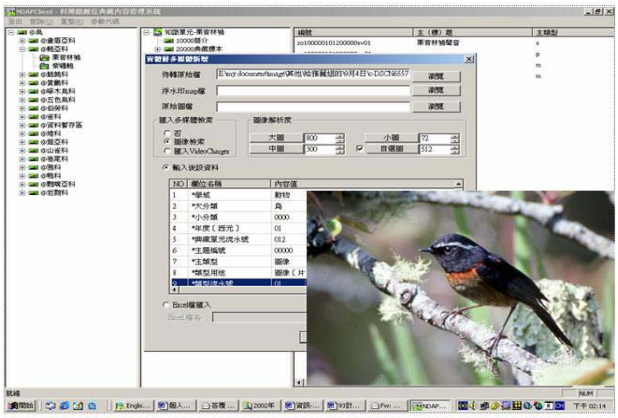

Figure 9 The creation of a core knowledge element

Figure 10 The creation of multimedia documents

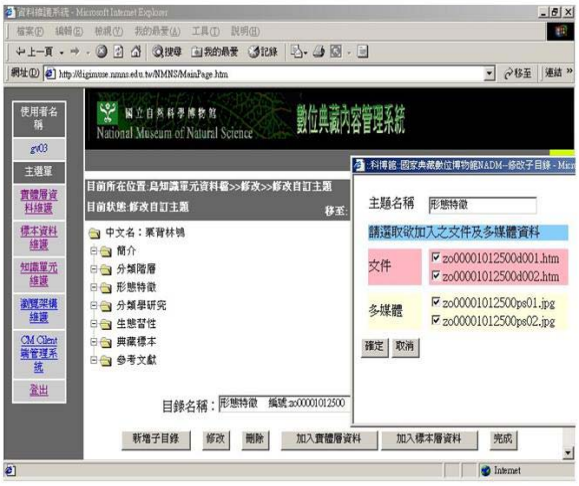

Figure 11 The creation of a knowledge unit

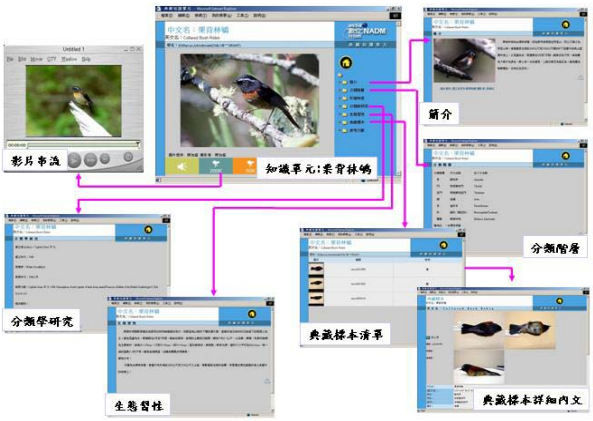

Figure 12 An example of knowledge unit

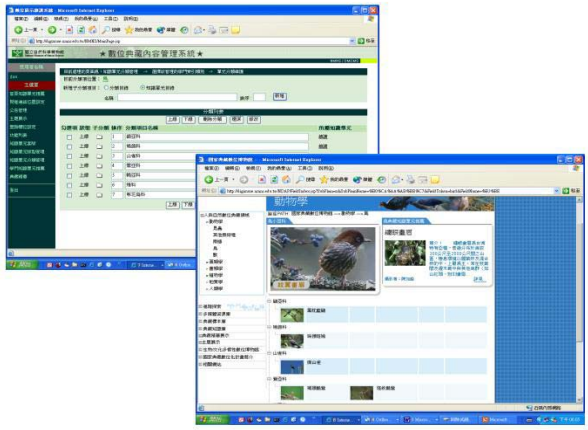

Figure 13 The creation of a knowledge group

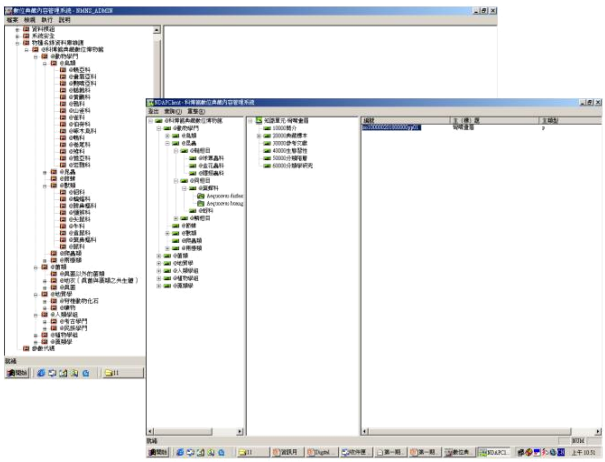

Figure 14 The creation of a ontology-based browsing hierarchy

6. Conclusion and Future work

For long-term perspective, worldwide museums are not only seeking for solutions to manage massive dynamic tacit and implicit knowledge content from all specialists and activities, but also thinking to transform themselves into value-added services for audiences with the help of novel information technologies.

In this paper, we give a solution to preserve, create, and manage the dynamic and massively growing enterprise knowledge content efficiently through a global view of knowledge-based content creation, management and publication processes to tightly connect the collaboration among participants. We provide a well-defined common domain knowledge system view among enterprise content specialists and construct a flexible multi-layer knowledge content structure for specialists to express and create explicit and implicit knowledge. By building a single global view of enterprise content, specialists and applications can share and reuse them transparently. Therefore, the plenty amount of knowledge content can be created repeatedly under the standard and consistent model. This approach not only provides a new communication environment between museum specialists and audiences but also creates an unlimited chance for an innovation museum.

In the near future, we will continue our research on discovering, classifying, and organizing core knowledge elements and advanced knowledge elements to generate innovation knowledge elements automatically from internal or Web knowledge content. The ontology-based knowledge system will also be dynamically extended and constructed according to the knowledge discovery process. Finally, the intelligent personalization services and Semantic Web mining mechanism will be developed for providing adapted and on-demand knowledge content for various users.

Reference

Ann Macintosh, Ian Filby and John Kingston (1999), ”Knowledge management techniques: teaching and dissemination concepts”, Int. J. Hum.-Comput. Stud., Vol.51, No.3, pp.549-566.

Ann Rockley (2003), “Managing Enterprise

Content : A unified content strategy", New Riders Publishing.

Alonzo, C. Addison (2000), "Emerging Trends in virtual heritage", IEEE multimedia.

Bob Boiko (2002), "Content Management Bible", Hungry Minds, Inc.

Carlo Batini, Stefano Ceri, and Shamkant B. Navathe (1992), "Conceptual database design," The Benjamin/Cummings Publishing Company, Inc..Chen, P. P. (1976), The entity-relationship model toward a unified view of data, ACM Transaction on Database Systems,Vol.1, No.1, pp9-36.

Gregor Engles, Martin Gogolla, Uwe Hohenstein, Klaus Hulsmann, Predita Lohr Richter, Gunter Saake, and Hans-Dieter Ehrich (1993), "Conceptual modeling of database applications using an extended ER model," Data & Knowledge Engineering, Vol. 9, pp. 157-204.

Jungpil Hahn and Mani R. Subramani (2000), "A framework of knowledge management systems :Issues and challenges for theory and practice", The 21st International conference on information systems, Brisbane, Australia.

Kavakli, E. (2003), "Museum Communication and New Technologies", Museology Journal.

Ken Sakamura (2000), "Digital Museum Distributed Museum Concept for the 21st Century", The University of Tokyo Digital Museum.

George F. MacDonald and Alsford, Stephen (1997): The Digital Museum. Paper published in The Wired Museum: Emerging Technology and Changing Paradigms, Washington: American Association of Museums, 1997.

George F. MacDonald, (2000), " Digital Visionary " , Museum news March/April

Macro Meli (2003), "Knowledge Management: a new challenge for science museums", Cultivate interactive Issue. 9.

Marwick, A.D. (2001), Knowledge management technology ,IBM system Journal, Vol. 40, No. 4.

Mason, I. (2002), "Knowledge management and cultural institutions", In Common Threads, MAD Conference.

Nenad Stojanovic and Ljiljana Stojanovic (2002), "Usage-Oriented Evolution of Ontology-Based Knowledge Management Systems", pp.1186-1204

Reagan W. Moore (2001), "Knowledge-Based Persistent Archives", SDSC, SDSC TR-2001-7, January 18, 2001

Ronald Maier (2001), "Towards a framework for knowledge management strategies: Process orientation as starting point", Proceedings of the 34th Hawaii international conference on system science.

Rosenberg, M. J. (2000), "E-learning: strategies for delivering knowledge in the digital age", McGraw-Hill Professional Publishing.

Samira Si-Said Cherfi and Nadira Lammari (2002), "Towards an Assisted Reorganization of Is_A Hierarchies", The proceedings of Object-Oriented. Information Systems : 8th International Conference.

Suzanne Keene (1997), "Becoming Digital", Museum Management and Curatorship, Vol. 15, No. 3, pp.299-313.

Sandra Beheim, Stephan Cradock, Eva Dächert and Sonja Pfurr (1999) "The learning Museum : How Museums use Information Technology to present value-added Collection Information for Lifelong Learning", Proc. Of the 7th International BOBCATSSS Symposium Learning Society-Learning Organization-Lifelong Learning, Bratislava, Slovak Republic, Jan. 25-27, 1999

Yin-Fu Huang,Tien-Yu Hsu (1999), "A Design and Implementation of Museum Multimedia Information System", Journal of Information Science and engineering 15, pp. 273-286.