Building Digital Heritage with Teamwork Empowerment

However, building digital heritage requires substantial resources in materials, expertise, tools, and cost. Government and university projects are limited in the time and space they can devote to covering even a small part of the world’s heritage. The preservation coverage problem is most serious in areas where sources of intellectual and cultural heritage may diminish or disappear over time. A central notion that helps resolve these issues is to make it easier for digital technology to reach sources of valuable heritage. The authors propose an approach to exploit noninstitutional resources for wider participation and inclusion in digital-heritage endeavors. The approach attempts to copy the techniques of institutional digital-heritage work while bringing together noninstitutional resources and providing standard practice.

New technology advances in computers, networking, and data processing lead the way into the digital era where various digital forms are adopted for creation, preservation, and application. One of the most significant roles of digital technology is to validate the notion of “virtual heritage” by harnessing digital tools to preserve and protect the world’s intellectual and cultural legacy. To this end, major efforts across the world, especially in the developed countries, are directed toward the development of multimedia-based digital libraries and digital museums. Evidence of these efforts abounds in many national- and international-level projects, such as American Memory, Making of America, the National Science Digital Library, the European Library, the British Museum, and the National Palace Museum. Work is also underway to integrate cultural heritage with digital library framework.

The movement toward digital heritage has been strongly supported by an increase in interest and resources from governments and academics. Many research projects have pioneered and contributed to the actual development of digital libraries and museums on selected subjects. Some of them proposed overall concepts, frameworks, and models of the composition of digital libraries or museums. Others focused on tools and techniques that resulted in new and improved components. These successful results have shown exciting promise in digitally preserving the world’s valuable heritage.

Jyi-Shane Liu (jshi@cc.nctu.edu.tw) is Associate Professor and Mu-Hsi Tseng (g9111@cc.nctu.edu.tw) and Tze-Kai Huag (g9210@cc.nctu.edu.tw) are graduate students, Department of Computer Science, National Chengchi University, Taipei, Taiwan.
distributed collaboration process, team members assume predefined roles based on areas of expertise or resources they possess; perform assigned tasks and activities; collect and transform intermediary results; follow through procedures and plans; and produce complete results, as directed by the system.

Both the idea of reproducing the digital-heritage process and the necessary framework were the results of cross-disciplinary knowledge interaction in a joint project on indigenous cultural-heritage preservation. HAM was developed and used to support the digital archiving work of two teams on different subjects. One subject, called Koo theatre, pertained to a very famous and important actress (Koo Zhen Cho) in traditional Chinese opera who, fifty years ago, was recognized as the archetypal leading actress of her time. The other subject, named “history for eyes,” involved thousands of very old journalism photos. Each team was lead by a professional subject expert and was composed of a group of graduate students with partial areas of expertise in digital archiving. In both cases, HAM was considered by each team to be a highly effective platform in orchestrating amateur teams and enabling the production of high-quality heritage digitization. The embedded ontology enabled the combined efforts of amateur teams to coordinate assigned tasks and produce coherent results. The system mechanism provided a shared workspace that supported accessing and integrating distributed work results. An interesting comment by one of the team leaders was “HAM is like a sea shell that grows pearls.”

In the remainder of this paper, a conceptual framework for producing digital archiving work is proposed. Based on the requirement analysis, an ontology is developed that comprises the notions of assembling multiple knowledge and skill areas, supporting teamwork and coordination, and transcribing production plans. HAM is presented and its functions described in mediating digital-archiving production. Examples of applying HAM to conduct heritage archiving and exhibition work are provided. A discussion of related research and the contribution of HAM and its complementary position to professional digital archiving is followed by a conclusion.

# A Conceptual Framework for Digital-Archiving Production

The central purpose of this paper is to provide a framework so that institution-level digital-archiving work can be reproduced by noninstitutional resources. This involves identifying what resources are required for the work and formulating a system that follows the standard digital-archiving work process. Typical noninstitutional workers are college or graduate students with amateur expertise and tool and equipment access, as well as people possessing the interest and capability needed to contribute to the effort. Appropriately empowering amateurs would foster widespread initial endeavors in digital-heritage preservation. Another positive effect concerns promoting heritage preservation in education and in society. Hands-on experiences and actual participation enhance learning and appreciation.

## Requirement Analysis

Digital archiving of cultural heritage involves multimedia documentation about and dissemination of selected subjects. The integrated use of image, video, sound, and text provides a rich context for preserving, learning about, and appreciating the documented subjects. Multimedia materials are powerful aids for conveying information and experiences of intellectual and cultural heritage. However, the creation, manipulation, and presentation of multimedia materials require special skills and expertise. In addition, subject documentation involves selection, compilation, and interpretation of subject materials. These activities require subject-domain knowledge and access to material sources. After the subject materials are acquired, they need to be checked, categorized, annotated, and organized. Persons properly trained in library science should perform these tasks. Finally, if a digital archive is to be built as an information and database system, software development capability is needed to construct and complete the system. Therefore, the first requirement of digital-archiving production is the assembly of multiple knowledge and skill areas.

From the physical standpoint, digital archiving is achieved by producing a structured set of subject-illustrating multimedia documents, and is very similar to manufacturing a product or developing a software system. All involve a teamwork process in which members with different expertise join forces to create a final entity that satisfies both global requirements and local constraints. During the production process, team members perform assigned tasks and collaborate with each other so that their results can be integrated into a final entity. The process involves coordination issues, such as communication, task allocation, conflict resolution, and task integration. Therefore, the second requirement of digital-archiving production is the support of teamwork and coordination.

The notion of producing professional digital-archiving work with amateur teams requires another component. Just like a factory manufactures a product according to a production plan, or the cast of a play performs based on a script, an amateur team needs directions on how digital archiving is to be done. To this end, professional digital-archiving practices can be transcribed into standard operating procedures for executing heritage-archiving work.
or production plans for producing subject-illustrating multimedia documents. With appropriate guidance, many amateur teams can be taught to produce professional-quality digital-archiving work on a vast array of subjects. Therefore, the third requirement of digital-archiving production is the transcription of digital-archiving practices into production plans.

**Modeling Digital Archiving**

Based on the requirement analysis, the next step is to develop a model that provides a methodology for producing digital-archiving work. The authors employ the notion of ontology in artificial intelligence. Ontology is an agreement about shared conceptualization of a problem domain, entailed by a set of concepts, such as entities, attributes, and processes, with their definitions and relationships. It involves the development and implementation of an explicit account of shared understanding in a given domain. Such an explicit and structural representation provides a basis for dealing with the operation and coordination of problems in many types of collaborative tasks.

Thus, an ontology is developed that comprises the notions of assembling multiple knowledge and skill areas, supporting teamwork and coordination, and transcribing production plans. The ontology is then included in an information system that acts as a platform to provide necessary guidance and support for producing digital archiving work. The ontology is composed of a set of generic concepts: stage, task, activity, object, specialist, and role (see figure 1). Each concept is an entity with attributes and has dependency relationships with others.

**Specialist**

A specialist is defined by the required area of knowledge or skill and is associated with the activities specific to that position. Based on observation of standard digital-archiving practices, five types of specialists are categorized:

- **Subject-domain specialist:** A person who has sufficient knowledge of the subject domain and has access to material sources. He or she decides how the subject will be illustrated, defines the structure of the materials, and interprets the content.

- **Digital-media specialist:** A person who is familiar with digital-media equipment and tools. He or she physically captures subject materials, converts them into digital forms, and produces and manipulates digitized multimedia-subject materials with required quality and specifications.

- **Data-management specialist:** A person who has adequate librarian training. He or she collects, categorizes, and annotates subject materials and establishes authority and control information on the materials.

- **Graphic-interface specialist:** A person who specializes in artistic graphics and Web interfaces. He or she creates multimedia content-presentation format on subject materials and performs Web-page layout and interface design.

- **Software-engineering specialist:** A person who is capable of programming and software-system development. He or she designs and implements the underlying information system that provides storage and retrieval for the archived subject materials. The information system includes databases, Web-page interfaces, and system management.

These specialists may cover more than one area of knowledge if they possess multiple skills. When the team is formed, team members assume the roles of specialists in their skill areas based on consensus. Usually, the person who is the subject-domain specialist initiates the call for an archiving project. This person also acts as team leader based on his or her subject-domain knowledge and strong interest in subject-heritage preservation.

**Workflow**

At the very top level, the digital-archiving work is viewed as a process with five sequential stages. Each stage represents an aggregate work step and must produce specified intermediary results before the next stage can begin. A stage is described by its status and requirements and is divided into several tasks. Each task defines what is to be done and is further divided into a set of sequential or parallel activities. Each activity is a basic work unit that can be carried out by an individual specialist. An activity is defined by operation description, input and output objects, and expertise needed. The hierarchical division of the workflow is shown in figure 2. Partial details of stages, tasks, and activities in the model are provided.

**Stage 1. Subject planning.** The stage starts with a selected subject. Team members are assigned individual specialist roles. Digital-archiving requirements and feasibility are discussed and evaluated. At the end, a work plan and schedule are produced. The stage is composed of two tasks.

**Task 1. Feasibility evaluation.** The purpose of the task is to confirm that the basic requirements of archiving operations are met with the available resources and the subject conditions. The task consists of five parallel activities, with the specialists evaluating a recommended check list for their areas provided by the model.

**Task 2. Plan and schedule development.** In this task, a production plan and schedule are constructed based on a plan-template.

Again, the task is divided into five parallel activities, with the specialists making their subplans and partial schedules. Next, subplans and partial schedules are inte-
grated with necessary revisions and adjustments to form a complete plan and final schedule. Plan and schedule integration are performed by the subject-domain specialist as task manager.

Stage 2. Materials digitization. In this stage, subject-materials sources are physically accessed. The selected materials are digitally captured in appropriate multimedia forms according to the nature of the materials and the archiving purposes. The digitized materials are categorized and annotated for further processing. The stage is divided into two tasks.

Task 1. Material acquisition. This task involves acquiring sufficient and useful subject materials and converting them into appropriate digital formats. The task is performed as two activities:

- Subject sources are researched and targeted. Potential materials are accessed, collected, and verified.
- Subject materials are captured or converted into digital files as basic storage units.

Task 2. Subject annotation. This task concerns providing context information for acquired subject materials. The context information is documented as metadata of the material file.

- In the first activity, standard metadata formats are incorporated and revised as necessary.
- In the following two parallel activities, the subject-domain specialist and digital-media specialist annotate subsets of metadata for each material file.
- In the final activity, the data-management specialist takes responsibility for completing and verifying all subject-materials annotation.

Stage 3. Materials transformation. In this stage, raw subject materials are further processed to provide more appropriate multimedia contents. The processing includes combination, segmentation, editing, and composition that lead to retrieval and exhibition units. In addition, the content files may be reformatted for different purposes. This stage involves two tasks.

Task 1. Content production. In this task, raw subject materials are modified, refined, and pieced together to produce content units that are semantically complete. The task is performed in two activities:

- The subject-domain specialist performs complementary creation for some subject materials. For example, images are accompanied by articles and videos are augmented by oral narration.
- The digital-media specialist synthesizes correspondent pieces and produces presentable content units.

Task 2. Format conversion. In this task, content units are duplicated and converted into appropriate formats for different presentation contexts. This task is also divided into two activities:

- Presentation purposes of content units are specified by the data-management specialist.
- The digital-media specialist performs necessary file duplication and format conversion. For example, image files are duplicated and resized for previewing. Video and audio files are streamed for real-time playback.

Stage 4. Content organization. In this stage, a content structure is defined and specified in such a way that all materials are categorized according to subject views. In addition, multi-modal materials are organized by semantic relations and are linked together. A database system reflecting the content structure is implemented to provide storage and retrieval. There are two tasks in this stage.

Task 1. Structure specification. In this task, a content structure is constructed and connected to all content units. The task is performed in two activities:

- The subject-domain specialist provides a semantic structure for the contents.
- The data-management specialist elaborates the semantic structure into specific data structures and organizes all content units accordingly.

Task 2. Database development. In this task, a database for the subject contents is developed. The software-engineering specialist performs the task in two activities:

- A database schema reflecting the content data structure is devised.
- Content units are deposited into the database and are linked based on database schema.

Stage 5. Content presentation. In this stage, a set of multi-modal media contents is enhanced with artistic and graphic design layouts. The refined contents are put together with hierarchical semantic linkages and proper access to the database. Again, this stage involves two tasks.

Task 1. Interface design. This task produces final digital contents on the subject that are to be exhibited as planned and retrieved as requested by viewers. The task is to
be performed by the graphic-interface specialist in two activities:

- A presentation-interface map for the subject content is drawn up.
- Content layout is designed. Visual and sound effects are supplemented.

Task 2. Online preparation. This task completes the underlying information system that integrates the content database and the Web interfaces. The task is to be performed by the software-engineering specialist in two activities:

- An information system is constructed to provide online presentation and access to the content database.
- The usability and stability of the system are evaluated and tuned.

The workflow model provides a standard operating procedure for amateur teams to follow. Team members carry out activities in their area of responsibility.

Individual work results are received at a central repository for group access and common review. Work progress is coordinated by the specialists’ mutual recognition of their different roles in object creation, transformation, and integration.

Object

Objects are work-result entities in different forms, from raw subject materials to final contents. Objects are defined to organize individual and team results and coordinate task progress. Each activity is associated with its input and output objects. The input objects of some of the initial activities are provided as instructions or references by the model. Output objects of previous activities are used as input objects for later activities. Therefore, team members build on top of each other’s work results and contribute toward a final product.

Besides objects at the activity level, there are also aggregate objects associated with tasks and stages. These
aggregate objects combine several component objects and represent distinct intermediary results. The aggregate objects are checked and evaluated to ensure their usability and quality before subsequent tasks or stages can begin. The input-output and part-aggregate relationships among objects serve as a production plan that orchestrates all production steps and leads to the final product. A partial-object diagram is shown in figure 3, where the rectangles indicate initial objects provided by the model, and ovals represent objects created by specialists.

Role

A role describes authority or responsibility associated with objects when a person assumes that role. Three roles are defined in the model (figure 4). First, each object, except some initial objects provided by the model, has a worker who is responsible for producing the object and submitting it for review. The worker is a specialist who is assigned to the activity that has the object as output. Second, an object that is an input to an activity has a user who decides whether to accept the object. The user is a specialist who is assigned to the activity that has the object as input. A worker produces and refines an object until it is accepted for further transformation by its next user. Third, an aggregate object or a final object at the task and stage levels has a manager who imposes suitability checks and quality control before the next task or stage can begin. The manager is a specialist who is assigned to the task or stage. After the manager approves the object, it is delivered to the next task or stage.

The role-specific commitment regulates coordinated teamwork and ensures proper progress toward the final product. By assuming one of the three roles in different contexts, team members share responsibility and authority in object creation and evaluation. Role definitions resolve potential conflicts and direct appropriate interactions among those working together.

HAM

The ontology serves as operational knowledge to direct and assist an amateur team in performing the teamwork process for digital-heritage archiving and exhibition. HAM incorporates the ontology and additionally mediates amateur teams in producing professional-quality digital archiving. In particular, the use of ontology in HAM provides five aspects of mediation for enabling and facilitating digital archiving and exhibition work.

- Communication: The team is comprised of people with different backgrounds and expertise. With a shared understanding determined by the ontology, conceptual and terminological confusion are reduced or eliminated. In addition, different viewpoints and constraints are connected by the ontology so that productive interactions are possible.
- Specification: Work requirements are provided and understood by team members so that they know what to expect from each other and what to prepare for completing their work.
- Operation: The production plan—detailed in steps, procedures, objects, and relationships—is provided and tracked so that the team is assisted to follow through a joint work effort.

Figure 3. Object Transformation Process

Figure 4. Relationships between Role and Object
- Composition: Relationships among activities and objects are followed in such a way that intermediary results can be evaluated, accumulated, and integrated toward the final product.
- Reusability: The representation of important entities, attributes, processes, and relationships operationalizes the functional knowledge that goes into producing digital archiving and exhibition. This functional knowledge can be reused in recreating the process that leads to successful results.

HAM was developed as a Web-based information system running on the Windows Server 2003 operating system. The Web site and user interface were built on top of a Microsoft Server IIS (Internet Information Services) with ASP (Active Server Pages) for workflow management and user access of data and files. Considering the potentially large quantity of digital-heritage content, Microsoft SQL Server 2000 was used as the system database for its data-access efficiency and management facility. HAM was run on a Shuttle (XPC SS51G) machine with an 2.53 GHz Intel Pentium 4, 512 MB RAM, and 80GB HDD, with 100 Mbps network bandwidth.

HAM can be used as an operation platform (figure 5) to organize team efforts, accumulate results, and produce digital-archiving databases and presentation interfaces. A team accesses the system via a Web-based interface and is guided throughout the teamwork process. The work is divided into a set of hierarchical stages, tasks, and activities, which are assigned to members based on individual expertise. The system enforces a production plan for the team to follow. Message boards support team communication. The team's intermediary results are stored and organized so that each member can check on the results and provide feedback as to their adequacy. For example, while the digital-media specialist is working on the subject-materal digitization activity, other team members can access (view) these temporary results (figure 6). At any point in the work process, the team can initiate discussions on current issues (figure 7). Discussion themes may include providing comments and feedback or brainstorming for ideas and solutions.

When a specialist assumes the role of worker on an activity, HAM is used to access the activity's input objects. Some of the output objects, such as converting analog files into digital, are created offline. The completed output objects are then uploaded to HAM by the worker responsible for them. Other output objects are produced by workers directly on HAM's Web interfaces. For example, the data-management specialist completes the annotation of subject materials by filling in the necessary information on HAM's work interface (figure 8). In content organization and content presentation, the last two stages, aggregate objects, such as database and final archiving content, are produced offline and are directed to a designated and authorized server. In essence, HAM is a working environment in which teamwork on heritage archiving and exhibition is mediated and coached to produce a subject-archiving database and its presentation interface.

Examples of Archiving Projects

HAM was used in two archiving projects to conduct heritage archiving and exhibition work on different subjects. A team of graduate students and subject-domain professionals ran each of the two projects. One subject, called Koo Theatre, pertained to a very famous and important actress (Koo Zhen Cho) in traditional Chinese opera who, about fifty years ago, had been recognized as the archetypal leading actress of her time. Her daughter, who was chosen to be the subject-domain expert, led the team. Other members included graduate students with expertise in digital media, data management, graphic and interface design, and software engineering. The team used HAM as the working platform and joined forces in distributed locations on a part-time basis over a period of two months. In the end, a multimedia archiving database on the actress and a Web-based presentation interface (http://koo.theatre.nccu.edu.tw) were produced. In particular, the database contains sixty-one event descriptions, eighty-four play profiles, 294 news reports, 325 photos of stage performances,
events, and musical scores, forty minutes of theater video clips, and fifteen minutes of opera recordings. A national newspaper covered the Web site because of its uniqueness and historical significance concerning Chinese-opera documentation. Some highlights of the presentation interface are shown in figure 9.

The other subject, named “history for eyes,” involved digitizing, annotating, and documenting thousands of very old news photos, originally stored in a little-used archive room. A journalism professor led the team, made up of graduate students who had the expertise necessary for the project. HAM was used over a three-month period, and the team produced a multimedia archiving database of news photos. Now complete, the archiving database contains three thousand annotated photos and thirty-five special-topic reports, and is a valuable asset of a prestigious university. Some highlights of the presentation interface are shown in figure 10.

In both pilot projects, HAM was considered by each team to be a highly effective platform for orchestrating an interdisciplinary joint effort to produce multimedia subject-archiving databases. The embedded ontology enables the combined work efforts of amateur teams to coordinate with assigned tasks and produce coherent results. The system mechanism provides a shared workspace that supports accessing and integrating distributed work results. An interesting comment by one of the team leaders was “HAM is like a sea shell that grows pearls.”

### Discussions

The notion of imitating and producing professional digital-archiving work and the use of HAM as a mediation platform for amateur resources helped expand the scope of heritage preservation. An archiving ontology was developed to coordinate teamwork and support standard workflow. In this regard, research in computer-supported cooperative work (CSCW) has considered using computers as a facilitating tool for the coordination process of teamwork. Notions such as multimedia and multi-modal collaboration, information management for collaboration, and asynchronous collaboration, have been explored. Another related research area is project management in which computer support for dealing with the complex operational issues in cross-functional teamwork, distributed decision making, and complex forms of work integration are studied. Our approach has both flavors but emphasizes developing an ontology-driven operation platform, so that amateur teams can be coached to complete joint work process.

HAM is best used in exploratory or educational digital archiving. The steps and procedures modeled in

---

Figure 6. A View Panel to Evaluate Results

Figure 7. A Discussion Panel to Exchange Comments

Figure 8. A Work Panel to Produce Output
Figure 9. Highlights of Koo Theatre
Figure 10. Highlights of “History for Eyes”
HAM are rather simplified and may not support complex or large-scaled digital archiving subjects. For example, 3-D reconstruction and visualization of archaeological sites would require serious expense in resources and technologies. This kind of work can only be done on very selective subjects by well-equipped institutions. As a complement, HAM allows a wider participation in digital archiving and helps the discovery of obscure and indigenous heritage.

The main contribution of HAM is to provide a platform that enables the access of many more resources into digital-heritage endeavors. This has three levels of significance for the world’s heritage preservation. First, HAM helps alleviate the heritage-preservation coverage problem. More national and local heritages can be saved and digitally preserved before they disappear. Second, HAM helps expand the scope of education and research in cultural heritage. Many college-level courses can use HAM as a pedagogic tool for students’ term projects in digital library and heritage preservation. These hands-on experiences enhance students’ awareness, appreciation, learning, and discovery of indigenous heritages. Third, HAM helps create a worldwide cadre for professional archiving institutions in collecting unique and indigenous heritage. All HAM-facilitated project results can be collected into a set of linking repositories or communities. Once a promising or valuable subject work is observed, archiving institutions can invest their resources to expand its depth and upgrade it to meet professional standards. The world’s heritage-preservation endeavor can be elevated to a new level by the joint forces of digital-archiving professionals and amateurs.

## Conclusion

Multimedia subject archiving and exhibition are essential for preserving and sharing the world’s intellectual and cultural heritages. The key to overcoming limitations of time and resources is to enable wider participation by skilled amateurs who have an interest in this effort. An ontology-driven mediation approach and an implemented system (HAM) are presented. The system serves as an effective vehicle for orchestrating digital-archiving work by self-directed, self-managed teams. HAM has been used in two interdisciplinary projects as an operation platform where amateur teams were coached to join forces in producing high-quality multimedia subject archiving and exhibition Web sites. In extension as a pedagogical tool, HAM has also been used in two graduate courses in digital libraries. The authors are currently evaluating the effectiveness of the system with user studies in two digital-library graduate courses.

## References