The Nautical Archaeology Digital Library

Carlos Monroy[1], Nicholas Parks[1], Richard Furuta[1], and Filipe Castro[2]

[1] Center for the Study of Digital Libraries, Texas A&M University,
College Station, TX 77843-3112, USA
{cmonroy, parksnj, furuta}@csdl.tamu.edu
[2] Institute of Nautical Archaeology, Texas A&M University,
College Station, TX 77843-4352, USA
fvcastro@tamu.edu

Abstract. In Nautical Archaeology, the study of components and objects creates a complex environment for scholars and researchers. Nautical archaeologists access, manipulate, study, and consult a variety of sources from different media, geographical origins, ages, and languages. Representing underwater excavations is a challenging endeavor due to the large amount of information and data in heterogeneous media and sources that must be structured, segmented, categorized, indexed, and integrated. We are creating a Nautical Archaeology Digital Library that will a) efficiently catalog, store, and manage artifacts and ship remains along with associated information from underwater archeological excavations, b) integrate heterogeneous data sources in different media to facilitate research work, c) incorporate historic sources to help in the study of current artifacts, d) provide visualization tools to help researchers manipulate, observe, study, and analyze artifacts and their relationships; and e) incorporate algorithm and visualization based mechanisms for ship reconstruction.

1 Introduction

The research methodology in Nautical Archaeology is evidence-based—data and artifacts gathered from the field provide the basis for evaluation of hypothesized relationships. The range of sources brought to bear in evaluating hypotheses is wide-ranging both in scope and also in time—sources range from hundreds-of-years-old historical treatises to digitized video indexed to data streams from modern-day satellite-based global positioning systems. In addition, data-gathering—i.e., surveying a site—is, by its nature, destructive, so the ability to validate a project's findings independently rests on the availability and completeness of the data and metadata obtained during the site survey. Clearly, relationships among these sources are complex.

We are developing a digital library framework for Nautical Archaeology that will provide a) flexible cross-linking of heterogeneous content in a dynamically-growing collection, b) flexible use of annotations to enhance community access while respecting individual information rights, c) incorporation and management of uncertain data, d) digital library replication and synchronization, and e) general applications of visualizations based on 2D grids. In our poster we will cover five major research areas in the context of Nautical Archaeology pertaining to: the excavation site, the archaeological recovery process, the artifacts collection, shipbuilding treatises, and ship modeling and reconstruction. Each area will include their corresponding source materials.
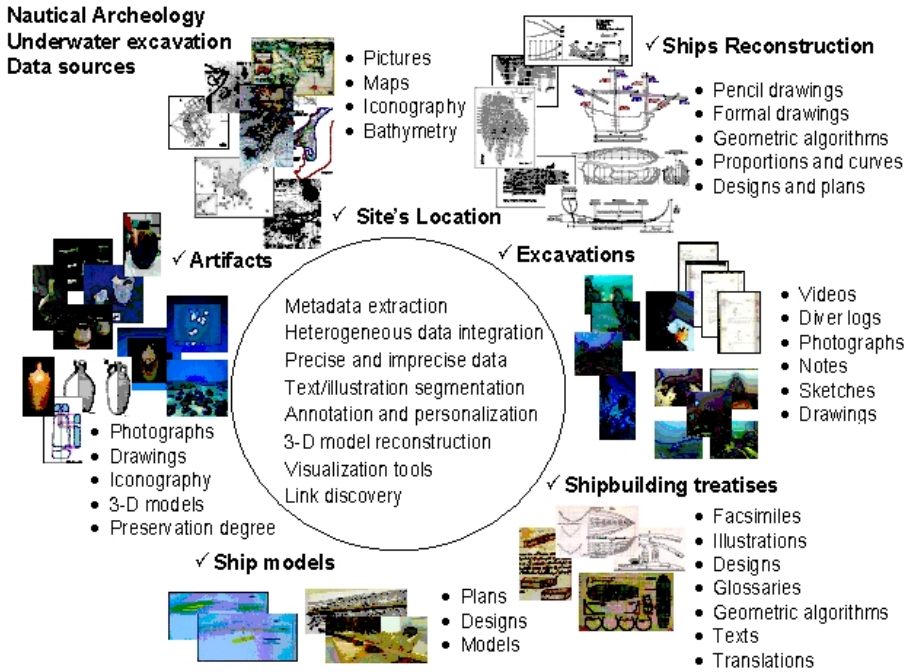
Fig. 1. The Nautical Archaeology Digital Library's data sources. The extent of the data sources involved is illustrated by the elements enumerated outside the central circle. Tasks to be addressed in the project are shown inside the circle.

The associated tasks to the aforementioned areas can be grouped as follows: a) developing a model for mapping an underwater archeological excavation site, b) establishing a protocol for storing, managing, and organizing information related to a shipwreck, c) creating a framework to enable the integration of heterogeneous data sources and media, d) developing new ways for structuring and accessing ancient shipbuilding treatises, and e) providing computational assistance for the identification and placement of ship fragments to allow ship reconstruction.

Figure 1 illustrates the scope of the proposed digital library, each area lists the data sources required. Tasks to be performed are listed in the circle at the center of the illustration.

Our project's philosophy is to investigate extensions to other's terrestrial archaeological digital libraries with the goal of addressing the unique characteristics of nautical archaeology. The result will be a resource of value to scholars and of interest to the general public.

2   Current NADL Project Activities

Our approach in the creation of this digital library is to focus not only on the collection itself but also on the work practices of the primary users. Thus, we are

developing a suite of tools to assist nautical archaeologists in their scholarly work. The major goal is to support the complete archaeological process from site discovery and excavation to conservation and publication. Therefore, we must also support the archeologist in both the connected and connectionless work environments; in essence, we propose a mobile ubiquitous digital library system.

To make a successful ubiquitous system for humanities practitioners, one needs to understand the work practices of the scholars involved. In this context, the NADL development will employ ethnographic methods in addition to traditional software usability methods. To date, we have hours of audio interviews and DVDs of captured observations of the archeologist at work; which will help us develop a tool kit that users will actually use. Presently we have already fielded a prototype tool, code-named "OnScene," to cope with the tasks performed at the excavation site. The pro-totype system will be used this summer on site in Portugal, and the experience gained will guide subsequent refinements.

Archaeological excavations have a one-time component; there exists but one chance of discovery, one instance to capture material in spatial context, and there is only one chance to perform point of capture data collection. This mobile connec-tionless environment creates the most data points and currently is where much infor-mation is lost. Additionally, incomplete, inaccurate, and subjective information makes the information technology needs of nautical archaeologists unique. In other domains, information technology tools deal with the distribution, organization, and understand-ing of content. Here, the focus is trying to assist the archeologist understand the in-formation available without the archeologist being aware of a digital library.

Further, a field excavation typically generates thousands of artifacts and other data points in a short period of time. The content types are far ranging from simple spread sheets, images, and video to the content generated by specialized archeological soft-ware. In fact, the fieldwork of an archeologist represents only a fraction of the time dedicated in an investigation. Archaeologists spend more time trying to understand the material recovered using methods familiar to other investigative scientists. Thus, our digital repository will automatically generate associations among recovered mate-rial. Also, we will assist those that use this tool kit and repository to create associa-tions themselves, similar to tagging.

Moreover, it is important to understand that archeological excavations are multi-year (if not multi-decade) endeavors where the investigators may and do change. In many instances also, objects recovered are shrouded in a concretion. Thus, these ob-jects tend to spend years as unknown entities while they endure the conservation proc-ess that is time and work intensive. Such changes will result in a discontinuity in the pace of the excavation as well as the focus.

The NADL team has been focusing on designing the product architecture to sup-port our goals. As much as possible, we seek to build from existing projects, both in the archaeological domain and also in other digital libraries activities. Projects within the archaeological domain that are influencing our initial thoughts include ETANA [2], which has developed a system that handles dissimilar content dissemination through OAI-MHP [3], and the work of the Alexandria Institute, which has success-fully articulated the need for adaptive representation of materials through their Ar-chaeoML work [1].

Our application domain of Nautical Archaeology raises an interesting set of problems. Reconstructing composite objects—such as ships—from incomplete or damaged sources requires, among other aids, the combination of algorithmic techniques and visualization tools. Ship reconstruction requires intensive querying of large amounts of dynamic information as new timbers are recovered and added into the repository. Once triaged, further visualization of the fragment in the context of a whole can help the researcher evaluate the suggested alternatives.

The repository includes timbers from other excavations as well as knowledge of the ideal characteristics, encoded from shipbuilding treatises. Shipbuilding treatises are technical manuals that include text describing the shipbuilding process, ship's proportions, illustrations of the pieces and components, and assembling instructions. The relevant treatises come from a variety of countries, kingdoms, and empires, over a span of several centuries, mainly between the 15th and 19th centuries. A challenging aspect in establishing relationships between treatises with ship remains is the capability to map them together. Further, representations and models of these relationships are important in establishing and/or validating hypotheses about the components of a ship, the construction techniques and geometric algorithms used, and the building sequences followed.

The NADL activity is still in its initial stages. Its progress can be followed at http://nadl.tamu.edu/.

Acknowledgements

References

1. Schloen, D.: Online Cultural Heritage Research Environment. The Oriental Institute of the University of Chicago http://ochre.lib.uchicago.edu/index.htm
2. Shen, R., Gonçalves, M., Fan, W., and Fox, E.: Requirements Gathering and Modeling of Domain-Specific Digital Libraries with the 5S Framework: An Archaeological Case Study with ETANA. In Proceedings ECDL2005, Vienna, 2005.
3. Van de Sompel, H., Nelson, M., Lagoze, C., Warner, S.,: Resource Harvesting within the OAI-PMH Framework. D-Lib Magazine, December 2004.