

Preliminary Survey of Image Management and Access Systems

A Report submitted to the
Andrew W. Mellon Foundation

Christie D. Stephenson
Brian Sheppard

Digital Library Initiatives
University of Michigan Library
Ann Arbor, MI 48109

October 11, 2000

I. Introduction—Framing the Investigation

In August 2000, we were approached by the staff of the Andrew W. Mellon Foundation to undertake an “environmental scan” of the current state of image management and delivery systems. The Foundation’s interest in this arena springs from their longstanding commitment to supporting the arts and museums and their more recent support of projects involving the intersection of information technology and humanities. More specifically, they have recently supported several projects to create large image-rich digital collections, such as the Architecture and Design Collection of the Museum of Modern Art and the Dunhuang Archive. These interests, coupled with the success of JSTOR, have led the Foundation to consider the potential for creating a related entity, currently referred to as ARTSTOR, to manage and distribute high-quality, scholarly, image-rich collections, such as those they are currently supporting through the creation phase. Though they had knowledge of some specific delivery systems, they were interested in having a better sense of the range of commercial products capable of managing and delivering this kind of content to specialist users. The primary goal of the survey was to help the Foundation evaluate whether or not their service could be built on one or more existing products or would require a major development effort, as was the case with JSTOR.

Further clarification of the vision for ARTSTOR and requirements for this study were carried out through several phone calls and draft documents, resulting in a scope of work statement. In those discussions, a number of assumptions about the requirements for the system were clarified, including:

1. While the majority of the content would likely be still images and accompanying structured metadata, the collections could include other data types (e.g. sound files, video) and unstructured text.
2. Some content might be created centrally and, therefore, the system would need to be able to provide cataloging and workflow management tools.
3. Some content might be federated from different sources, so the system would need to provide for data import and mechanisms for mapping between divergent metadata schemes.
4. End-users needed functionality that would facilitate use of the content in teaching and research.

These requirements helped shape the conduct of the subsequent investigation as well as guiding the process of describing individual products.

II. Methodology

The list of potential products to be investigated in detail was populated using a variety of methods. The principal investigator has had a long-standing involvement in the realm of image database systems¹ and was able to identify several potential products from personal knowledge. Research in journal indexes and on the web broadened the pool, adding a large number of digital asset management systems that have been emerging to serve the need of media-rich businesses and media-dependent industries. A particularly extensive listing of such products appeared in the journal **Publish** in 1999.² All of the one hundred and eleven products included in this list were reviewed on the web to further refine the pool. A variety of criteria were applied to narrow the list to those companies that appeared to have particularly promising products. In this rapidly growing and

¹ Christie Stephenson developed an early web-based image delivery system as Coordinator of the Digital Image Center in the Fine Arts Library at the University of Virginia. She served as Project Director for the Museum Educational Site Licensing Project in its final year and continues to monitor developments in this arena.

² Nichoson, Mark, "Master your media: a complete guide to digital asset management," **Publish** 14:3 (60-7) March 1999.

changing environment, a number of companies included in the **Publish** list had vanished, merged with others, or changed their emphasis. Other products were eliminated if they appeared to have a very narrow focus on a particular vertical market, such as the pre-press industry. Products developed for the desktop were also eliminated. Museum collection management systems, though they increasingly offer web delivery options, were judged to be out of scope, because their primary focus is on management, control and documentation of objects, rather than digital surrogates. The final list of products in the pool to be evaluated numbered nineteen (See Appendix 1).

The investigators began by gathering self-descriptive information from each company's web site in an attempt to address the various evaluative criteria outlined in the scope of work. It quickly became clear that this method would result in extremely uneven treatment, so a list of questions was compiled to ensure that each vendor would respond to each of the issues under investigation, providing some measure of comparability across product descriptions. After testing the questions in a phone interview to one vendor, the questions (see Appendix 2) were sent to each of the companies via email. The response rate was very good; follow-up messages were sent to the vendors that did not respond. However, several products were eliminated from the final review because of vendor non-responsiveness.³ Reviews of 15 products are included in this survey.

III. Characterizing the products

The products identified for review by the investigators can be broadly characterized in a number of ways.

- Of the fourteen, twelve are commercial products, and two come from the higher education sector: CONTENT from the University of Washington's School of Engineering and DLXS from the University of Michigan Library.

³ Luna Imaging did not respond to the survey; they have, however, been included in the product review because of their particular focus on delivering the kind of content ARTSTOR will include to a similar user base.

- Of the commercial products, some come from well-established companies and some from very young ones. A number of the companies are privately-held and backed by venture capital.
- Two of the commercial products, iBase's inVisage and Luna's Insight, have worked closely with the cultural heritage, museum, and library communities. The same is true of the two university-based products.
- One product, Canto Cumulus, has evolved from a desktop-based, personal image cataloging tool, to an enterprise-wide digital asset management system
- A number of the products, including Software Construction's Media Grid/Media Server and Imation's Media Manager, have evolved from a products geared to a particular specialized market (e.g. newspaper publishing), and retain biases towards meeting the needs of those vertical markets
- Several of the products are based in companies that closely integrate their own hardware and/or software products (IBM Content Manager, Informix Media 360).
- Luna's Insight product is unique among those surveyed in terms of its stated focus on supporting the needs of the scholarly user. This is not surprising given Michael Ester's research in this area dating back to the late 1980s.⁴
- A number of the products put much of their functionality into a closely integrated "fat client" and less emphasis on web browser delivery; others rely totally on a web interface for both internal, management functions and public access.
- Most of the "industrial-strength" asset management systems are understandably focused on providing functionality and tools for internal users of the system, or supporting business-to-business applications. Several of them have formed alliances with what they call "web experience" developers to assist customers in developing public interfaces to their underlying management systems.

⁴ During Ester's tenure at the Getty Art History Information Project, AHIP and Brown University's Institute for Research and Information collaborated on a study investigating how art historians work and organize their teaching and research. See Bakewell, et al. *Object, image, inquiry: The art historian at work*. Santa Monica, Getty AHIP, 1988.

These high level characterizations may assist in identifying products that are more stable and mature as well as companies more readily able to understand and respond to the requirements for ARTSTOR.

IV. Towards an Evaluation Framework

For the preliminary survey, the investigators have not attempted to evaluate the products, but simply to describe them, largely relying on vendor-supplied information reported in a standard template. At this stage, with the vision for ARTSTOR evolving and therefore its specific functional requirements unclear, we felt an evaluation to be premature. However, in the course of the investigation, it became clear that certain high-level system characteristics should be considered highly desirable in any product or combination of products that might be given serious consideration in the future.

Among those we judged most important are:

1. **Scalability:** Assuming rapid growth of numbers of users and numbers of assets, not only the underlying database but all components of the system must be able to handle the increased volume over time.
2. **Extensibility:** It is highly unlikely that any "out of the box" system will meet all the needs of any given customer. Rather than relying heavily on vendor customization, products that provide clearly defined APIs and SDKs are preferable, giving the customer the ability to make modifications as required.
3. **Distributed and open architecture:** The component architecture should be modular, distributed and open to avoid the limitations of strict client/server architecture; the business logic should be separated from the database(s) to allow for component independence to ensure a robust, scalable solution. Communication between components should be standards-based.
4. **Ease of Integration:** In addition to providing an open architecture framework, a system should facilitate integration of new hardware, applications, and functionality as requirements change and mature over time. It should be feasible to integrate new

media storage systems, such as a hierarchical storage solution, mechanisms for intelligently integrating new content types, as well as mechanisms for delivering those new content types. A track record of incorporating third-party applications into a product can be seen as an indicator of this ease of integration.

It is difficult to predict in any detail the functional requirements of the centralized content management system for ARTSTOR. However in the broad outline of assumptions about the ways in which content might come into the system, features such as data import/export, cataloging tools and metadata handling emerge as particularly critical. Again, the flexibility and extensibility of any existing product become important distinguishing criteria. Solutions that provide the ability to develop customized ingest tools and provide support for flexible data transport mechanisms like XML seem likely to be most suitable.

In the area of external user functionality, we can expect that users will want powerful searching capabilities, the ability to arrange and manipulate images in a variety of ways, to mark, save and share search sets, and annotate images. We can also predict that they will want to reuse content in a variety ways, some of which we have yet to imagine. It will be important to articulate which of these functional requirements may be more appropriately met by an application on the user's own workstation and which must be provided within any given collection or aggregation of collections, tightly integrated with the delivery system. In any case, the assertion can be made that systems that provide clear paths for customization are preferable. In addition, we know from experience that librarians are loath to support the licensing of resources that require users to download and configure custom clients; if at all possible, the required functionality should be provided through a standard web interface.

These evaluative criteria evolved over the course of our investigation and out of the process itself. It is hoped that they will provide a lens for viewing the results of our "environmental scan", for stimulating thinking as the vision for ARTSTOR continues to

mature, and for articulating next steps in investigating appropriate solutions for grounding that vision in a technical solution.