Editor's Note

In the summer of 2024, Clifford Lynch announced his retirement as executive director of the Coalition for Networked Information (CNI) after 28 years at its helm. CNI quietly launched a project to create this Festschrift to document and honor his legacy. Authors began contributing articles in early 2025, with a planned publication date of July 2025. Since the final membership meeting of Cliff's tenure was April 7–8 in Milwaukee, the plan was to surprise him, surrounded by colleagues and friends, with a presentation of the table of contents of this special issue. However, just two weeks prior to the meeting, Cliff's health worsened; he was told about the Festschrift and received project details and articles. Though unable to attend in person, he participated in the CNI membership meeting via Zoom and also virtually joined his retirement reception, which included readings of excerpts from each article in this volume. Sadly, on April 10, 2025, Clifford Lynch passed away. Festschrift contributors wrote their articles prior to his passing, and we have chosen not to alter their original language.

Stewardship of Digital Images

Howard Besser

abstract: This article presents a short list of key projects in the history of stewardship of digital images. Clifford Lynch played a role in all of these. He is responsible for popularizing the word *stewardship* to reflect management across the life cycle of a digital work.

The article begins with the mid-1980s project ImageQuery, the first to offer a variety of image database features. Next, it looks at the Dublin Core, a mid-1990s project to standardize image metadata for web discovery. It then turns to the mid-1990s Museum Educational Site Licensing Project, a 14-institution collaborative undertaking to test digital image distribution. Finally, it looks at other issues that impact image stewardship, such as copyright for digital works and the handling of vast quantities of digital works.

. . . in each of these projects in the history of stewardship of digital images, Clifford has teased out important issues and made the project grapple with them.

Across 40 years, Clifford Lynch has played a significant role in each of these. As he has done in so many other domains, in each of these projects in the history of stewardship of digital images, Clifford has teased out important issues and made the project grapple with them. And through his widespread knowledge of different projects around the world, he has encouraged collaboration.

Introduction

For over 40 years, Clifford Lynch has had a significant impact on key digital image stewardship projects. This article describes some of those projects, discusses important issues that Cliff helped to identify, and illustrates how he encouraged collaboration. It

portal: Libraries and the Academy, Vol. 25, No. 3 Supplement (2025), pp. 99–105.

is worth noting that Cliff is responsible for popularizing the term *stewardship* to reflect management across the life cycle of a digital work.

ImageQuery

In 1986, the Office of Information Systems and Technology at the University of California, Berkeley began work on a project to deliver high-quality digital images from its art museum, architecture slide library, and geography department. The developers believe that this software (eventually called ImageQuery) was the first deployed multiuser networked digital image database system. The software was first shown publicly in June 1987 at the conferences of the American Association of Museums (now the American Alliance of Museums) and the American Library Association. For most attendees, this was their first time viewing a high-resolution image on a computer screen.

ImageQuery was an X-Windows-based system with several features that were relatively new for the time: a graphical user interface, point-and-click searching, thumbnail images to permit browsing and sorting, tools for annotation of images, and the linking of images to locations on maps. In addition, ImageQuery was designed for networked accessibility, had client-server features, and permitted Boolean searches.[1]

ImageQuery featured thumbnail images linked to a list of brief records for each image (see Figure 1). Clicking on an image highlighted that image as well as the related text record. Clicking on a text record highlighted the related image. This proved to be a powerful method both for finding the correct image in a list of hits, and for quickly identifying an image displayed on the screen.

Each displayed thumbnail image was linked to both a full-text record and a larger version of that image. A pull-down menu (triggered by pointing to a thumbnail image and holding down a mouse button) would give the user the choice of displaying the full image or text. This proved to be a powerful tool to link browsing to more complete information, though in today's environment small buttons appear to be more effective than pulldown menus.

ImageQuery's architecture was modular. The user interface sent queries to a database that resided separately, so different databases and structures could serve as the "back-end." For a number of years, ImageQuery could only support back-end structures that had been collapsed into flat files, but eventually capabilities were added to support SQL-type queries. Another limitation of ImageQuery was that the text database structure had to be pre-identified and coded into a short preferences file, rather than dynamically discovered.

ImageQuery's design incorporating a separate module for text storage and retrieval is still a powerful idea. It allows image database developers to leverage off technical developments in the much larger text-database market, enabling great efficiencies in indexing and retrieval. The modularity also permits external applications to easily access the text portion of the database. The ImageQuery design is part of a movement away from closed, non-modular systems toward the modularization of user interface, query structure, search and retrieval, and storage. The system links these modules through a set of standards and protocols.

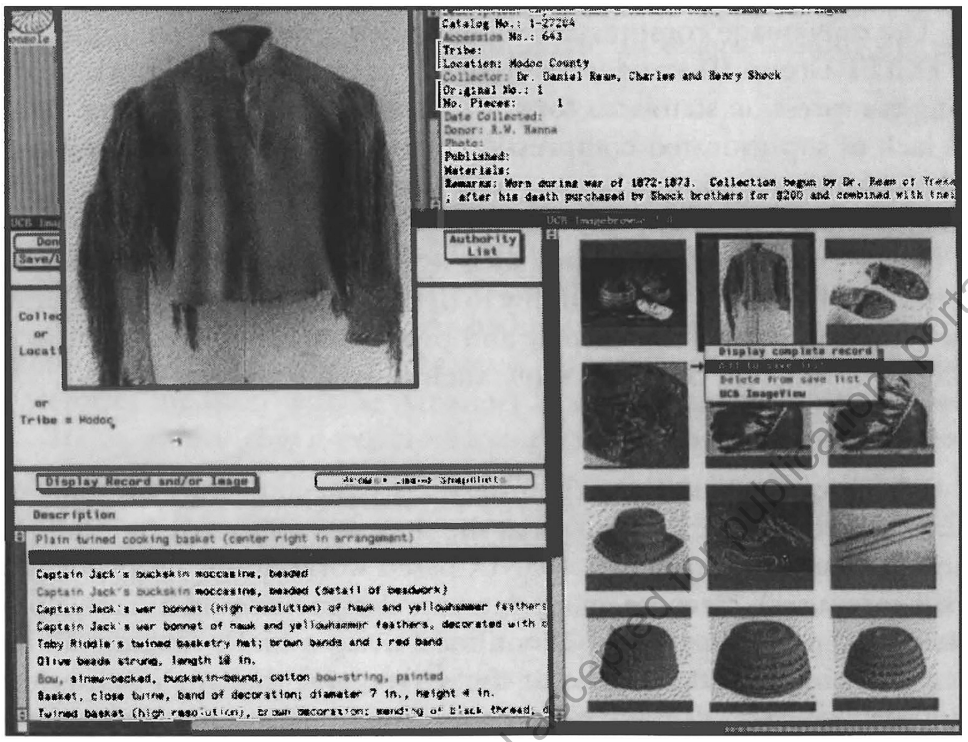

Figure 1. An ImageQuery screen grab from 1987 (image courtesy of the Phoebe A. Hearst Museum of Anthropology, University of California, Berkeley).

The system also employed modularization to link tools for users to view and process images. By pointing to an onscreen image, a user could pull down a menu and choose a variety of image-processing tools that could be applied to that image. ImageQuery would then invoke software (such as paint programs for annotation, color-map programs for balancing and altering colors, or processing programs for zooming) that would allow the user to analyze or alter the current image.

The project team's idea of links to external tools was part of a broader view of what an image database should be. The team's philosophy was that (particularly in an academic environment) simply providing access to a database was not enough; developers had the responsibility to provide the user with tools to integrate the results of database retrieval into their normal work processes.

The ImageQuery developers recognized the importance of a client-server architecture, both to assure that the image database could be accessed from a wide variety of platforms and to put less strain on the server and network by off-loading some of the functionality onto client workstations. But the ImageQuery team expected that environment would be X-Windows-based. They waited years for a variety of developments over which they had no control—the porting of X-Windows onto Intel and Macintosh platforms, an increase in the installed base of X-Windows machines, and the development of a set of extensions to X Windows called the X Imaging Extensions. No one on the ImageQuery

development team anticipated the post-1993 phenomenal growth in World Wide Web browsers that would make this the delivery platform of choice. Web browsers not only solved the multi-platform and central database load problems, but also implemented client functionality in a more sophisticated way than ImageQuery. Web browser helper applications recognize a variety of image file formats, handle decompression, and can spawn external viewing software. All these functions combine to lessen the load on the network and the server, and to increase the number of file storage options.

Another key philosophy behind ImageQuery was the implementation of a user interface that would provide a common "look and feel" across all image collections. Prior to ImageQuery, each campus object collection had its own idiosyncratic retrieval system and user interface. Users had to make a substantial investment of time to learn to use one of these retrieval systems, and many people appeared reluctant to invest the time to learn a second. The ImageQuery team believed that a common user interface would encourage cross-disciplinary use of these collections, so they designed a system that on the surface always appeared the same to the user. Only the names and contents of fields differed from database to database. An "authority preview" function was developed to permit users (particularly those unfamiliar with the valid terms associated with a field name) to view a list of terms that had been assigned within a given field. Much of the appeal of World Wide Web browsers likely lies in the fact that they act as a universal interface, providing a common "look and feel" to anything they access. Though a function to preview the actual contents of a field within a database still appears powerful, this has not yet been widely implemented.

In a number of areas, the designs for ImageQuery look naive in retrospect. Though the notion of interoperability remains important, the functionality to allow searching across image databases of different objects (each having different field names and contents) is vastly more complex than the ImageQuery team anticipated. The team also failed to appreciate the challenge of scaling up. Though they considered methods for decreasing storage cost and topologies which would limit the impact on a particular server or a particular segment of a network, little thought was put into how to handle queries that might retrieve thousands of initial hits. ImageQuery did provide for important functionality such as visual browsing to narrow down query sets (by clicking on the thumbnail images that the user wanted to save). But this process by itself would not help the user whose initial query retrieved more than 100 hits. In retrospect, relevance feedback and similar functions look critical to dealing with large image databases.

Cliff visited the project when it was in the prototype stage. He encouraged us to continue developing its modular nature and to show it more widely to the library community. He was particularly interested in the feature that tied an image and text record to a location on a map.

The Dublin Core

In 1995, Cliff asked me to attend a meeting at the Online Computer Library Center (now just OCLC) to discuss making works on the World Wide Web more accessible. This gathering took place less than two years after the release of the first web browser. Most people were not yet thinking about discovery issues, as not many resources were yet web-accessible.

The meeting brought together about two dozen stakeholders, including individuals from library information technology (IT), SGML (standard generalized markup language) publishing, and representatives from the National Center for Supercomputing Applications (NCSA). The group defined a baker's dozen set of metadata elements that became known as the Dublin Core.

The meeting envisioned a future when a vast quantity of scholarly materials would be available on the World Wide Web, and there would be a need for discovery services to locate relevant material. The idea was to give creators or intermediaries (such as libraries) tools to identify each web resource with metadata elements (such as author, date, publisher, and format).

The Dublin Core element set was designed to be simple enough that scholars or publishers could assign the metadata themselves, yet rich enough to enhance discovery. It was expected that libraries and museums could write scripts to "dumb down" their complex cataloging records into the Dublin Core metadata format.

The Dublin Core group decided to meet annually, taking up a specific topic at each yearly meeting. Subsequent meetings were in Warwick (framework) and Canberra (qualifiers). In 1998, the topic was "Extending Dublin Core to Images," and I was the moderator. The original Dublin Core had been designed for text documents. In the 1998 meeting, we tried to make it work better for other document types, particularly images. This involved both changing element names (*Author* became *Creator*), changing field definitions (*Format*), and adding fields (for a total of 15).

Cliff was a vocal contributor to all these Dublin Core meetings. He also worked hard to find important stakeholders and invite them to the meetings.

Cliff was a vocal contributor to all these Dublin Core meetings. He also worked hard to find important stakeholders and invite them to the meetings.

The Museum Educational Site Licensing Project

The Museum Educational Site Licensing Project (MESL) provided the first serious testbed for image databases in a multisite academic environment. Beginning in 1995, approximately 10,000 images and accompanying rich metadata from six museums and the Library of Congress were distributed and deployed on seven university campuses.[2]

Each university mounted the images locally and provided their own user interface. This allowed them to examine the infrastructure and tools needed to deploy an image database in an environment with many users. It also helped them understand what was needed to incorporate the use of image databases into the instructional environment. Each deployment had its own approach to discovery (metadata mapping), display (image browse and zoom), and instructional use (templates for teachers).

Over the three-year project, representatives of all 14 institutions met regularly. They were guided by a management team consisting of Max Anderson, David Bearman, Howard Besser, Clifford Lynch, Christie Stephenson, and Jennifer Trant. Participants worked together to define the terms and conditions governing the educational use of digitized museum images and related information. They also shared deployment issues.

The MESL Project was important for a variety of reasons. It helped universities test user interfaces for special collections, and it highlighted how metadata mapping could significantly affect query results.[3] It also helped collecting institutions understand the various ways their images could be used (or misused). A study of MESL commissioned by the Andrew W. Mellon Foundation led to the creation of Artstor, the nonprofit organization that builds and distributes a digital library of art images.[4]

Cliff was an important contributor to the Museum Educational Site Licensing Project.

Cliff was an important contributor to the Museum Educational Site Licensing Project. As a member of the management team, he helped identify areas of exploration and guided the participants through a variety of challenges.

Beyond the Images Themselves

As we moved into the new millennium, it became increasingly clear that the challenges to digital image stewardship were linked to broader and more systemic issues. The wider issue of copyright for all types of digital material was tackled in a National Research Council (NRC) panel on which both Cliff and I served. In that panel and the resulting book, we gave a detailed analysis of the impediments to the stewardship of all types of digital works.[5] The panel spent two years hashing over policy and technology issues before reaching the trade-offs and conclusions outlined in the book. The book was the first to tease out copyright issues in the wake of the Digital Millennium Copyright Act.

The broader issue of how to handle immense volumes of digital works was tackled in another NRC panel on which Cliff and I served. Our panel spent two years developing recommendations to the National Archives and Records Administration on how to handle their vast number of digital records.[6]

Both NRC panels brought together multiple stakeholders, with Cliff and I representing educational and cultural heritage institutions, as well as the wider public. Each panel had less than two dozen members, and Cliff was a vocal participant in both.

Some of Cliff's recent thinking has tried to differentiate between software preservation and systems stewardship. He has looked at this issue in systems ranging from games to artificial intelligence. And that has led to questions like: Should an avatar or chatbot of an individual become part of the cultural record? Should library special collections collect these?

Looking back, Cliff has observed that we have been successful at stewardship of journals because all the stakeholders (authors, publishers, libraries, and scholars) want it to happen. But we have been less successful stewarding other types of material when some stakeholders are not onboard. He offered up the example of digital books, where publishers want to license and librarians/stewards want to invoke the doctrine of "first sale," which allows legitimate purchasers of the books to do whatever they want with the books—use them, dispose of them, or sell them.

Summary

This has been a brief summary of major milestones related to the stewardship of digital images. Across 40 years, Clifford Lynch has played a significant role in each of these. As he has done in so many other domains, he has teased out important issues and encouraged project participants to grapple with them. Additionally, through his widespread knowledge of different projects around the world, he has been an important proponent of collaboration.

This has been a brief summary of major milestones related to the stewardship of digital images. Across 40 years, Clifford Lynch has played a significant role in each of these.

Howard Besser is Professor Emeritus of Cinema Studies at New York University and the founding director of the master's degree program in Moving Image Archiving and Preservation; he may be reached by email at howard@nyu.edu.

Notes

1. Howard Besser, "Image Databases: The First Decade, the Present, and the Future," in *Digital Image Access & Retrieval*, ed. P. Bryan Heidorn and Beth Sandore (Urbana-Champaign, IL: Graduate School of Library and Information Science, University of Illinois, 1997), https://hdl.handle.net/2142/25941.
2. Christie Stephenson and Patricia McClung, eds., *Delivering Digital Images: Cultural Heritage Resources for Education* (Los Angeles: Getty Publications, 1998), https://www.getty.edu/publications/virtuallibrary/0892365099.html.
3. Howard Besser, "Comparing Five Implementations of the Museum Educational Site Licensing Project: If the Museum Data's the Same, Why's It Look So Different?" in *ICHIM 97: Fourth International Conference on Hypermedia and Interactivity in Museums* (Pittsburgh: Archives and Museum Informatics, 1997), 317–25; republished in David Bearman and Jennifer Trant, eds., *Museum Interactive Multimedia 1997: Cultural Heritage Systems Design and Interfaces, Selected Papers from ICHIM 97* (Pittsburgh: Archives and Museum Informatics, 1997), 50–58.
4. Howard Besser, Rosalie Lack, and Robert C. Yamashita, eds. "Cost and Use of Digital Images on University Campuses: Lessons from the Museum Educational Site Licensing Project," special issue, *Visual Resources: An International Journal of Documentation* 14, 4 (1999), https://howardbesser.name/howard/imaging/1998mellon/.
5. National Research Council, *The Digital Dilemma: Intellectual Property in the Information Age* (Washington, DC: National Academies Press, 2000), https://doi.org/10.17226/9601.
6. National Research Council, *Building an Electronic Records Archive at the National Archives and Records Administration: Recommendations for Initial Development* (Washington, DC: National Academies Press, 2003), https://doi.org/10.17226/10707.