The ARTstor Digital Library: A Case Study in Digital Curation[1]

Max Marmor
Director of Collection Development
ARTstor
151 East 61st Street
New York City, NY 10021
max.marmor@artstor.org

Abstract

ARTstor is, like our traditional libraries, at once an *institution* – a "place" - and a *collection* of resources. As an *institution*, ARTstor seeks to work closely with both the community of content owners (archives, libraries, museums, photographers, publishers) and the the community of end users comprised of scholars, teachers, and students at the 700 institutions here and abroad currently participating in ARTstor. Above all, ARTstor strives to help bring both communities together around what should be a shared effort to create an enduring digital resource that will support the work of the scholarly, educational and cultural communities. As a *collection*, the ARTstor Digital Library currently offers more than 500,000 digital images and associated descriptive cataloging information, delivered within a networked space and a software environment that supports active use of these images in teaching, learning and scholarship while also addressing the significant concerns of content owners about appropriate – and inappropriate – uses of their "content." ARTstor evolved at a significant moment in the evolution of digital libraries. A*RTstor was conceived to test the feasibility of building a digital library based on the needs of potential users*. ARTstor sought to pose the question: Is it possible to understand the needs of potential users of an emerging digital library *before* developing a collection development strategy for that library and before making substantial investments in the business of collection building? Is it feasible to target and pursue specific bodies of "content" – to build coherent and, to use the language of bookbinding, "bespoke" digital image collections – based above all on persuasive if not compelling evidence of their potential value to end users? In its effort to test this hypothesis, ARTstor has from the outset sought consciously to assemble and, where necessary, to help create coherent digital collections that would respond in meaningful ways to the core needs of educators and scholars who use – or wish to use – images in support of teaching, learning and scholarship.

ARTstor (http://www.artstor.org) is, like other libraries, at once an *institution* – a "place" - and a *collection* of resources.

As an *institution*, ARTstor seeks to work closely with both the community of content owners (archives, libraries, museums, photographers, publishers) and the community of end users (scholars, teachers, students, museum professionals). Above all, ARTstor strives to help bring both communities together around what a shared effort to create an enduring digital resource that will support the work of the scholarly, educational and cultural communities.

As a *collection*, the ARTstor Digital Library currently offers more than 500,000 digital images and associated descriptive cataloging information, delivered within a networked space and a software environment that supports active use of these images in teaching, learning and scholarship, while also addressing the significant concerns of content owners about appropriate – and inappropriate – uses of their "content."

While the focus of this discussion will be on the building of the early ARTstor collections, a few words about ARTstor as an institution might be helpful by way of introduction.

The Genesis of ARTstor

Originally conceived as an initiative of The Andrew W. Mellon Foundation in 2001, ARTstor became an independent not-for-profit organization in January 2004. The roots of ARTstor, as well as its name, may be traced to JSTOR (http://www.jstor.org), ARTstor's sibling, initiated more than a decade ago by the Mellon Foundation. Much as JSTOR digitizes, "stores" and distributes an encompassing digital archive of the back-files of core journals in the arts, humanities, social sciences and sciences, so ARTstor assembles, "stores" and distributes digital images of visual materials that are central to teaching, learning and scholarship. In performing these activities and services, ARTstor – again like JSTOR – seeks to advance and even, in tandem with the complementary efforts of others, to transform the ways in which scholarship, teaching and learning are conducted in the evolving networked environment.

As of April 2007 ARTstor – which is only available to non-profit institutions – has more than 740 participating institutions, including scores of community colleges, colleges and universities both public and private, museums, primary and secondary schools, and independent art schools. Outreach to public libraries and, through them, to independent scholars unaffiliated with cultural institutions, is just now being announced. ARTstor is currently available for participation in the US, Canada and, quite recently, Australia/New Zealand and the UK. We anticipate making the ARTstor Digital Library available to the larger international community in stages as we define and address the many challenges related to services and support across multiple languages and time zones that will inevitably go hand in hand with expanding international access.

Building the ARTstor Charter Collection

The ARTstor Charter Collection, the first collection of aggregated content available through the ARTstor Digital Library, currently contains more than 500,000 images. The Charter Collection documents, through an expanding array of individually curated source collections, visual culture from around the globe and across all cultures and historical time periods, richly illustrating works of architecture, painting, sculpture, photography, decorative arts, and design as well as many other forms of visual and

material culture. The Charter Collection is intended to support many – but as we shall see, by no means all – of the image needs of teachers and scholars throughout the arts, humanities and social sciences. It embodies collaborations with archives, libraries, museums, photographic archives, publishers, slide libraries and individual scholars. Although the Charter Collection is an aggregate of multiple individual collections, all users at participating institutions have integrated access to ARTstor images across all these constituent collections.[2]

ARTstor evolved at a significant moment in the evolution of digital libraries.[3] *ARTstor was conceived to test the feasibility of curating and building a digital library based on the needs of potential users*.[4] Many institutions, ranging from archives and libraries to museums, have of course elected to digitize their collections on the – frequently untested – assumption that, once available online, they would prove to be of significant interest and value to scholars and teachers. Sometimes this assumption has been amply confirmed, but other digital collections created in this way have yet to demonstrate their utility and value to their targeted audiences. Put baldly, many have yet to provide the expected "return on investment" in both programmatic and – since cost is of course always relevant – financial terms. Speculative digitization – the "field of dreams" approach that says, at least implicitly, "digitize it and they will come" – has been and remains an abiding feature of digital library development in the non-profit sector.[5]

In explicit counterpoint to this abiding "basso continuo," ARTstor sought to pose the question: Is it possible to understand the needs of potential users of an emerging digital library *before* developing a collection development strategy for that library and before making substantial investments in the business of collection curation? Is it feasible to identify, target, pursue and secure specific bodies of "content" – to build coherent, "bespoke" digital image collections – based above all on persuasive evidence of their potential value to end users? In its effort to test this hypothesis, ARTstor has from the outset sought consciously to assemble and, where necessary, to help create coherent digital collections that would respond directly and in verifiable ways to the core needs of educators and scholars who use – or wish to use – digital images in support of teaching, learning and scholarship. Of course, like all significant academic and research libraries, ARTstor has not *only* pursued a focused collection development strategy; it has also been the happy beneficiary of donations from collaborators and other well-wishers, and benign opportunism has also played a role as well – since strategy minus opportunity equals mere abstraction. Strategy and serendipity have, then, gone hand in hand at ARTstor, but ARTstor has placed a premium on strategy throughout its early evolution. We will return to the role of serendipity by way of conclusion below.

How has this strategy been implemented? As an attempt to model a strategic approach to curating a digital library based expressly on the needs of potential users and participants, one of ARTstor's first steps was to seek to provide teachers, students and scholars throughout the arts and humanities with the digital equivalent of a large interdisciplinary academic slide library. It had become increasingly evident that a primary focus of investments in digitization in the higher education sector over the past two decades has been a widespread – and largely redundant – effort to replace the 35mm teaching slide with digital images. Slide and visual resource curators from Berkeley to Berlin have been, and to some extent remain, engaged in this effort. And indeed, early efforts to create digital libraries in the US museum community faced strong – and ultimately insuperable – challenges precisely because they could not, in the nature of the case, address this very practical need for a large, core body of teaching images drawing on the entire world of art and visual culture.[6] At the same time, pedagogy emerged as one key area where the needs of the academic community were relatively well-understood thanks to decades of teaching with slides, making it feasible to define

those baseline teaching needs with some degree of precision and to actively seek out and curate digital image archives that would at least begin to address them.

The outcome of this early exploration was the ARTstor "Image Gallery," which even in its initial form offered a compendium of roughly 200,000 images. Shaped around representative curricula in the arts and humanities, the Image Gallery – which continues to grow[7] – has already demonstrated its capacity to advance the transition from slides to digital images in art history instruction and associated fields fundamentally dependent on the use of teaching images, while also relieving many academic institutions and visual resource collections of the perceived need to digitize their own slide collections in support of the core image needs of teachers in the arts and humanities. The Image Gallery has had the added virtue of offering scholars, teachers and students in fields outside the arts – individuals who have typically lacked slide and photographic archives and services, let alone digital collections and services supporting teaching and learning – with the foundation of a rich and unusually accessible "campus-wide" visual resources collection.[8]

Through a series of more recent collection building efforts, ARTstor has focused on adding breadth and depth to the ARTstor Digital Library by providing ARTstor users with deeper, more specialized collections in a range of fields, including American, African and African-American, Asian, Classical, Islamic, Medieval and Renaissance Studies, the history of architecture and the built environment, graphic design and visual communication. As a result of this effort, ARTstor now offers special strengths in the arts of Asia, thanks to partnerships with the John and Susan Huntington Archive of Buddhist and Related Art at Ohio State University, with the American Council for South Asian Art and the University of Michigan, with Northwestern University and with other institutions. ARTstor is also developing significant strengths in the art of the Middle Ages, through collaborations with the Bodleian Library, Oxford, which is contributing more than 25,000 digital images of medieval manuscript painting; with Princeton University's Firestone Library and Index of Christian Art; and with the photographic archives at the National Gallery of Art Library in Washington, D.C. Through collaborations with the Frick Art Reference Library, the National Gallery of Art Library, the Samuel H. Kress Foundation and Scala Archives in Florence, Italy, ARTstor will soon offer comparable strengths in the area of Italian Renaissance art and architecture. The Scala collaboration alone will provide 13,000 high resolution images of Italian art and culture from antiquity to the present, with especially strong coverage of Renaissance painting.

In addition to these broad collections, ARTstor also provides what might best be regarded as special collections of primary source materials, the product of a series of innovative collaborations with archives, libraries, museums and publishers. A digital version of *The Illustrated Bartsch*, a monumental 100-volume reference work that offers approximately 55,000 images derived from European prints – woodcuts, engravings, etchings – from the 15th to the early 19th century, is a prime example of such a primary source collection. And a digital version of the Gernsheim Photographic Corpus of Drawings, a monumental archive that documents more than 184,000 old master European drawings, will offer similarly comprehensive coverage of this key area of European art. Both these collections draw on the holdings of scores of museums in the US and abroad.

In some instances, ARTstor is pursuing focused collaborations with individual museums or groups of museums. For example, ARTstor offers a complete digital version of the Museum of Modern Art's rich collection of architectural and design objects. And the majority of the images (perhaps 90,000) from the former "AMICO Library"[9] are now available in ARTstor, thanks to the many distinguished art museums that originally

contributed to that collaborative effort. Similarly, numerous museums here and abroad are contributing to the Mellon International Dunhuang Archive, a treasure trove of cultural materials associated with the hundreds of Buddhist cave shrines at the Dunhuang oasis site, a key node for ten centuries on the cultural crossroads of the Silk Route, in the Gobi Desert. Participating museums include the British Museum, the Hermitage, and the Musée Guimet; participating libraries include the British Library and the Bibliothèque Nationale de France.[10]

ARTstor is, finally, actively pursuing projects in key areas of "non-western" art, including African and Pre-Columbian art and archaeology, the art and architecture of Islam, and modern Latin American art.

As this brief overview suggests, ARTstor's efforts to provide early on for many of the well-defined image needs of teachers and scholars – to provide for "the canon" – should not be seen as fully defining ARTstor's collection development trajectory. ARTstor is actively seeking to texture and to update traditional teaching canons and to respond to evolving methodologies and pedagogies – in art history and throughout the humanities. In some instances, specific collections are already being created or secured to respond to these evolving interests. A case in point is a suite of collections that might be said to focus upon "social iconography," and which document the roles and representation of diverse social groups in western cultural history. A key project in this area is a collaboration with Harvard University, the goal of which is to digitize the 30,000 image archive related to "The Image of the Black in Western Art" project, to cite the title of a renowned reference publication based on this archive.[11] A tandem collection is a digital archive of 36,000 images from the renowned Schlesinger Library photographic archives, devoted to the history of women in America, also at Harvard and now available in ARTstor.

Each of these projects[12] represents a different approach to digital curation, and together they represent collaborations with museums, slide libraries, publishers, faculty photographers, teams of scholars, and photo archives. Assessment of these models of collection building and the value of these collections to ARTstor's varied audience will shape ARTstor's ongoing development.

ARTstor and "the Library as Place"[13]

ARTstor's effort to provide valued collections that would respond directly to the needs of scholars, teachers, and students – to implement the "user-driven" collection development policy described briefly above – has had important consequences for ARTstor as an institution and for the process of building collections.[14] For one thing, it presupposes a considerable degree of agility, flexibility and editorial control in targeting and pursuing potential content – wherever that content might reside. To that extent it seemed to argue against the kind of consortial membership and governance structure that is so familiar among cultural organizations. Consortia – whether of archives, libraries or museums – tend in the nature of the case to have organizational boundaries that, at least implicitly, limit the universe of candidate collections even as they help ensure contributions of content. Arranged marriages, after all, achieve efficiency at the price of narrowing the pool of potential partners! Consortia accordingly tend to produce relatively loose assemblages of content that reflect the distinctive strengths – as well, of course, as the sometimes strikingly varied capacities and priorities – of their members. What they offer in diversity they tend to lack in strategic focus, even when they do not embody the "field of dreams" approach alluded to above.

Choosing to pursue a different approach has meant that ARTstor, far from working within the collecting context of an "arranged marriage," has had to play the role

of marriage broker. Whereas a consortium may in an important sense take the securing of "content" *per se* for granted, ARTstor has had to pursue desired content actively. This has meant articulating our understanding of the needs of teachers, students and scholars to the collecting community, and then championing those needs with potential content providers.

This has been challenging on several fronts. It has frequently made for long courtships, many of which have had to begin with introducing ARTstor and its mission to individuals and institutions, here and abroad, who were not familiar with this new initiative. It has meant seeking to persuade these proposed partners to make the priorities of scholars and teachers their own priorities as well. Above all, it has meant addressing the fact that highly prized collections tend to be highly sensitive collections as well.   Even when rights and rights management have not been a central focus of discussion – and they frequently have! – other stakeholder concerns – including appropriate and inappropriate uses of digital content and sensitive questions related to cultural patrimony – have typically been at the heart of these discussions.

ARTstor's effort to bridge the interests of content owners and potential users of digital collections has entailed compromises at both ends of the spectrum. Put metaphorically, we have concluded that this essential "bridge" must in important respects remain a "covered bridge" – at least for now. By that we mean that in order to balance the concerns, interests and needs of content owners with those of end users, we have felt obliged to create a secure network on the internet, within which digital content can be used in appropriate ways by educators and scholars, without for the most part allowing that content to be removed from the digital library for use in other environments. We have, in short, wrapped ARTstor content in the ARTstor software. And we have thereby placed real limits on our ability to interoperate with other systems and services. We have taken this approach for two reasons: First, we believe that this is the only we can build the kind of valued collections our users say they most want from a service like ARTstor; and second, we believe it is important to keep these two communities in dialogue – a mission-driven goal we would jeopardize if we fully accommodated the interoperability interest some institutions and individuals have expressed.

We have nonetheless recognized from the outset that even if ARTstor were successfully to demonstrate the feasibility of building an immense digital image library shaped around the needs of its users, the ARTstor Digital Library would never have *all* the images an individual scholar or researcher needs – though we do of course hope gradually to provide an increasingly large proportion of those images. For the foreseeable future, participating institutions will continue to build their own – hopefully distinctive – digital image archives in response to local needs, and ARTstor users will inevitably continue to develop personal digital image archives much as scholars and teachers have always assembled personal collections of articles, books, photographs and slides. And both libraries and end users will wish to integrate these "local" resources alongside the hundreds of thousands of licensed images ARTstor provides, preferably in an integrated software environment. To address this essential need to use ARTstor images in a single environment alongside other visual materials available to our users, we have begun to enhance ARTstor's capacity to function as a "place" where collections as well as users come together.

This effort has taken a few forms. One is a nascent institutional "hosting service." ARTstor is now hosting institutional collections for more than thirty ARTstor participants that have built their own digital image collections and wish to make use of them within the ARTstor software environment, alongside ARTstor's own collections. These institutions have found that providing local access to their locally-developed

collections alongside ARTstor's provides a useful solution for offering easy, integrated access to local and licensed image collections; and in some instances this hosting service saves our partner the significant investment required to manage and support use of institutional digital collections.

In a similar vein, all registered ARTstor users may readily use their own images alongside ARTstor images in the classroom by using ARTstor's Offline Image Viewer. The OIV is a tool ARTstor provides for managing images offline, for creating and delivering presentations that draw on both ARTstor images and local images. Where appropriate, individual users may also upload images into a "personal collection" space within ARTstor online, for consultation by colleagues or students.

A parallel effort to enable the integration of ARTstor content alongside other information resources has focused on federated searching or metasearching. Many of our participating institutions have, of course, implemented metasearch engines that allow users to search many databases from a single interface. Recognizing the need to provide a way for such metasearch engines to search into the ARTstor Digital Library, we have created an XML gateway that provides both a stable, standardized method for querying the ARTstor Digital Library, and also a way of retrieving search results that can easily be utilized by a metasearch program.

Finally, above and beyond providing an XML Gateway, we have also begun another initiative that will improve the ability of other resources and repositories to 'interoperate' with ARTstor. We are building a standard API (Application Programming Interface) to allow ARTstor to search and retrieve results from existing image delivery systems, institutional repositories or electronic image resources already in use at participating institutions. We expect this API to be available in the course of 2006.[15]

What all these approaches to embedding ARTstor into the larger landscape of digital libraries have in common is that ARTstor's role is not merely that of a passive provider of "content," but that of a shared space – a "place" – in which collections from a variety of sources – and users of those collections – encounter one another.

Conclusion: Toward an ARTstor "Network"

We noted above that ARTstor, like all libraries, has been the beneficiary of donations from interested individuals and institutions. We are finding increasingly that ARTstor participants wish not merely to use their own image collections alongside ARTstor's licensed collections; they want to share them with colleagues outside the walls of their own institution. For example, for a year or so ARTstor hosted an important Cuban Heritage Collection for the University of Miami Library. Recently, the Library expressed an interest in making this important collection available to all ARTstor users. The collection how now been integrated into the ARTstor Digital Library. In many other instances, the hosting "stage" has, as it were, been bypassed altogether; the Cornell Libraries for example have recently contributed three significant collections to ARTstor.[16] Archives, libraries, and museums are increasingly turning to ARTstor as a vehicle for sharing already digitized collections. Frequently, these institutions express a mission-driven wish to share high resolution images for non-commercial use in teaching and learning. They are not comfortable placing such images on the open internet, but they see ARTstor as a trusted partner to whom they can safely entrust their collections, knowing they will be used only by their intended audience and only in intended ways.

This development has led us to reflect that the nearly 750 institutions now participating in ARTstor constitute, from one perspective, a nascent network. Currently, traffic across this network moves mostly in one direction – from ARTstor outward to hundreds of colleges, universities, museums and schools. But there is no reason

content should not also flow in multiple directions – back through ARTstor and outward again to other institutions, or indeed *between* individuals and institutions. We are now turning our attention to ways of fostering this kind of bilateral communication across the ARTstor "network" and we welcome the guidance of others active in this swiftly changing arena.

---

[1] This talk is based on an article first published by the author in *Collection Building*, 25, no.3 (2006): 95-99. It is made available here in accordance with Creative Commons license 3.0 (http://creativecommons.org/licenses/by-nc-nd/3.0/) , which allows copying, distributing and transmitting this work for non-commercial uses only and precludes creating or distributing derivative versions of this work. Attribution must include reference to the author as the creator and must site the initial appearance of a version of this text in the journal *Collection Building*, cited as above.

[2] Descriptions of the main sources of the Charter Collection may be found on the ARTstor public website at http://www.artstor.org/info/collections/whats_in_artstor.jsp. Forthcoming collections are announced at http://www.artstor.org/info/collections/upcoming.jsp.

[3] For an attempt at an overview of early digital library efforts, see Daniel Greenstein and Suzanne E. Thorin, "The Digital Library: A Biography" (*CLIR Reports*, no. 109, 2002).

[4] For an early statement of this goal, see Max Marmor, "Toward User-Centered Digital Image Libraries," *CLIR Issues* no. 20, March/April 2001 (online at http://www.clir.org/pubs/issues/issues20.html).

[5] The "field of dreams" metaphor has been applied to digital libraries frequently in presentations and papers by Donald J. Waters, program officer for Scholarly Communications at the Mellon Foundation. See for example "Developing Digital Libraries: Four Principles for Higher Education" (a 2001 EDUCAUSE review article) at http://www.educause.edu/ir/library/pdf/ERM0158.pdf.

[6] Early collaborative digital efforts in the American art museum community included the Museum Educational Site-Licensing (MESL) project (1995-98), the abortive Museum Digital Licensing Collective (1998) and the recently concluded Art Museum Image Consortium (AMICO) (1997-2005). The Digital Library Federation's experimental Academic Image Cooperative initiative (1999-2000) was expressly intended to help model a sustainable but *user-centered* digital image library; work on this project informed the early development of ARTstor. See the series of DLF reports at http://www.diglib.org/collections/aic.htm.

[7] An initial source of the Image Gallery was the slide collection at the University of California, San Diego. ARTstor invited UCSD to contribute this collection in recognition of the fact that is has supported a wide range of image users throughout the arts and humanities, that its collection closely mirrors in its scale and contents representative curricula in the arts and humanities, and that the UCSD Library had created item-level MARC records – with topical subject headings largely derived from LCSH – to describe each of the roughly 250,000 slides in the collection. UCSD was also the principal investigator for the Mellon-funded Union Catalog for Art Images (UCAI) initiative; see http://gort.ucsd.edu/ucai/.

[8] For a case study see the profile of ARTstor use at the University of California, Berkeley: "The right picture: finding it, organizing it, showing it, storing it," in *The Berkeleyan*, October 26, 2005 (online at http://www.artstor.org/info/news/Berkeleyan_102605.pdf). See also Barbara Rockenbach and Max Marmor, "ARTstor's Digital Landscape," *Library journal*, July 15, 2005 (online at http://www.libraryjournal.com/article/CA623002.html).

[9] See www.amico.org for documents concerning this now- dissolved effort.

[10] See http://www.artstor.org/info/collections/mida.jsp.

[11] See http://dubois.fas.harvard.edu/index.html for the project and its publications.

[12] For the projects and collections mentioned in the text, see the collections list at http://www.artstor.org/info/collections/whats_in_artstor.jsp and the chronicle of new collection announcements at http://www.artstor.org/info/collections/upcoming.jsp.

[13] See *Library as place: rethinking roles, rethinking space* (CLIR publication no. 129, Feb. 2005); see http://www.clir.org/pubs/abstract/pub129abst.html)

[14] For a discussion of "lessons learned" from the first two years of ARTstor's experience, see Max Marmor, "Six Lessons Learned: An (Early) ARTstor Retrospective," *RLG DigiNews* 10, no.2 (April 15, 2006) (online at *http://www.rlg.org/en/page.php?Page_ID=20916&Printable=1&Article_ID=1822).*

[15] See the account in Barbara Rockenbach and William Ying, "ARTstor: Enabling Cross-Resource Communication," *Library hi-tech news*, 22, no. 9 (2005): 21-23.

[16] See http://www.artstor.org/info/news/cornell_content_announce.jsp.