

RE-FRAMING COLLECTIONS FOR A DIGITAL AGE: A PREPARATORY STUDY FOR COLLECTING AND PRESERVING WEB-BASED ART RESEARCH MATERIALS

NEW YORK ART RESOURCES CONSORTIUM (NYARC): Brooklyn Museum Libraries and Archives, Frick Art Reference Library, Museum of Modern Art Library and Archives

SUMMARY

Increasingly the materials that the NYARC institutions collect are changing from printed to digital formats. Although there has been much activity in capturing general digital research materials, less attention has been paid to some types of art historical materials, which general research libraries ignored even in the print version. Building on existing studies and practices, this study would address this gap. It would also investigate what changes need to be made to the NYARC technical infrastructure (e.g. its Millennium system), and to underlying workflows, in order to present a unified access point to the same type of art documentation (e.g. auction catalogs) whether it is available as print, digitized or only as born-digital.

Subject to the results of the study, we envisage making a future proposal to the Andrew W. Mellon Foundation for support to enable us to acquire the requisite technical infrastructure to achieve this.

It will take one year to complete the project.

NARRATIVE

a. Background

NYARC formally began in 2006 as a consortium of the libraries of the Brooklyn Museum, The Frick Collection (Frick Art Reference Library), the Metropolitan Museum of Art, and the Museum of Modern Art. With funding from the Andrew W. Mellon Foundation, three of the institutions in 2007-10 developed and launched an innovative integrated catalog, ARCADE. In December 2010 the Metropolitan Museum of Art Library formally withdrew from the consortium.

Recent enhancements to ARCADE include links to artlibraries.net, full-text digitized content in Google Books and HATHI Trust, the migration of Photo-archive collection-level and individual works of art records into Arcade etc.

About the NYARC Libraries

The Brooklyn Museum's Libraries and Archives collections parallel the encyclopedic scope of the Museum's object collections, spanning from antiquity to contemporary art. Two research libraries – the Art Reference Library and the Wilbour Library of Egyptology – and the Museum Archives

contain approximately 300,000 volumes, 2000 linear feet of archives in addition to special collections of 50,000 documentary photographs, 42,000 fashion sketches and over 5,000 rare books. Dating to 1823 the Library collections include books, exhibition catalogs, periodicals, artist and institutional files, auction catalogs, pamphlets, trade cards and catalogs, as well as visual materials. The Archives, established in 1985, collects, organizes and makes accessible documentation about the Brooklyn Museum and its predecessors. The Brooklyn Museum Libraries and Archives support research activities of the Museum's staff as well as its exhibition program. These collections are also available to the public onsite and offsite: access is provided via digital scans, email, fax and telephone services. This department is supported by private and public funding with grants from federal, state and city programs. For example, the Museum has been a recipient of IMLS grants to provide training for interns from Pratt Institute's School of Information and Library Science.

The Frick Art Reference Library was founded in 1920 by Helen Clay Frick as a public reference library for the serious student of art history. At its heart was a photoarchive based on the model of the Witt Library, London. In 1984 it formally became part of The Frick Collection but the scope of its collecting remains wider than that of the Collection. Its mission is to give public access to the research literature and documentation of Western art history; to contribute to the public understanding of the literature and sources of art history; to support the curatorial activities and study of The Frick Collection and its context; and to promote the study of the history of collecting in America. It is the home of the Center for the History of Collecting, which maintains a database of collector, dealer and gallery archives relevant to collecting, and has a program of fellowships, symposia, and publications. It also organizes a biennial book prize. Recent library activities include the migration of photoarchive records into Arcade (the NYARC integrated catalog), the provision of web access to the digitized images of the American photographic campaign negatives (a project funded by NEH), and (with Brooklyn), a gilded age digitization project funded by METRO. The Archives receive funding from the Helen Clay Frick Foundation. The Center receives funding from the Kress Foundation, Leon Levy Foundation, Billy Rose Foundation, Sotheby's, and various individual donors. The library has two interns funded by IMLS.

The Museum of Modern Art's Library and Archives is a comprehensive collection of research resources devoted to the study of modern and contemporary art. The non-circulating collection documents painting, sculpture, drawings, prints, photography, architecture, design, performance, video, film, and emerging art forms from 1880 to the present. The Library's holdings include approximately 300,000 books and exhibition catalogues, 300 periodical subscriptions, and over 40,000 vertical files of announcements and ephemera about individual artists. Collection highlights

include works on Dada and Surrealism, conceptual art, Latin American art, [The Museum of Modern Art/Franklin Furnace Artist Book Collection](#), and the [Political Art Documentation and Distribution \(PAD/D\) Archive](#).

The Museum Archives was established in 1989 to collect, organize, preserve, and make accessible documentation concerning the Museum's art-historical and cultural role in the 20th and 21st centuries. It is also an internationally recognized center of research for primary source material concerning many aspects of modern and contemporary art, including private archives that may be the papers of artists, collectives, galleries, dealers, art historians, critics, etc. Recent acquisitions of major archives of Fluxus and conceptual art make MoMA a destination for research in art documentation of the 1960s and 1970s. The Library and Archives enjoy support not only as part of the Museum's core mission but also through the generosity of its trustees, major foundations, and donors. MoMA's Library Council is a group of approximately one hundred members whose contributions benefit the historic and contemporary research collections in MoMA's Library and Museum Archives. A central focus of the Library Council is its artist's publication program. At least every other year, the Library Council publishes a specially printed artist's book or edition that is intended to shed light on the Museum collections and to give artists an opportunity to experiment with the book as an art form.

b. Rationale

Art research materials traditionally collected by art libraries have included museum and exhibition catalogs, catalogues raisonnés, auction and dealer catalogs and related price lists, photographic images, and artist and art institutional ephemera such as private view invitations and gallery checklists. These are rapidly becoming digital-only and web-based without an analog print copy. Digital documents in theory have unlimited distribution, but they remain difficult to locate, and, given that the average life of a website is between 44-75 days (Source: Internet Archive), are more ephemeral than ever.

Without an ability to collect these materials and store them on a permanent basis, there is a potential 'digital black hole' in the documentation of the art world. Because the use of web-based materials is only going to grow, failure to act now runs the risk of art libraries losing access forever to important primary research materials, a risk that could jeopardize the Libraries' long term future.

We believe that web-harvesting will be potentially the primary mode of ingest of both free and commercial born-digital material into our institutions in the future, and we need to prepare for this eventuality.

We are aware of other initiatives in this area, such as those of the International Internet Preservation Consortium (IIPC) and the Andrew W. Mellon-funded Columbia Web Archiving Project: we would leverage these projects wherever possible. However, none of the existing web-based curated collections are capturing the sort of materials that the NYARC libraries have traditionally captured in print. The art historical field has distinctive characteristics that require a unique solution. These include highly complex, multi-stranded intellectual property issues, a literature that goes beyond the traditional monographs and serials of most humanities disciplines, and a very long tail of publishers/content creators. Moreover, many art-related sites are technically innovative and pose challenges to conventional harvesting technologies. In addition, there is great interest in our institutions beyond the libraries and archives in capturing digital content and so we will work closely with our curatorial and information technology departments in determining the parameters of the project.

We are also assuming that the various national library initiatives in national domain crawling and the harvesting of websites operate under legal deposit legislation and therefore do not allow access to the archived websites outside the physical walls of the particular national library. However, such programs have acquired considerable technical expertise into which we will try to tap.

The NYARC partners, with a track record of successful cooperation in building new technological and service models, are uniquely positioned to forge new models in building collections in the digital age. The Frick Art Reference Library recently conducted a pilot project in partnership with the Internet Archive, in order to set up a test site to capture resources associated with auctions of artwork. This pilot finished in February 2011 and a report completed (See Appendix 2). The project identified issues to be resolved before we can move from an ad hoc collecting to a sustainable program – crawl efficiency and frequency, format challenges, IP issues, the need to integrate born digital resources in a common resource discovery environment, and to develop templates with OCLC to help streamline cataloguing of this material, and not least the re-training of existing staff to be able to execute these new tasks.

The Library of Congress' [Program for Cooperative Cataloging](#) in a September 15, 2011, report addressing the implementation of Resource Description and Access (RDA), the emerging international standard for cataloging library materials, stated the need for creation of "community-specific" guidelines for cataloging integrated resources like web documents in special subject areas. The NYARC group wishes to contribute to the development of best practices for cataloging art-related web resources and will utilize its team of technical contacts to work with the consultants on this project to establish local policies and procedures for description that could be adopted by the wider community.

NYARC has been active in bringing forward the issue of born-digital resources in the US art libraries community: it has arranged for a preliminary discussion for the March 2012 ARLIS/NA conference in Toronto, and is proposing a formal presentation at the 2013 conference: we think that there is scope for co-operative activities here if only in terms of distributed selection. This would be an ideal opportunity to disseminate the results of the project we are proposing.

We are asking the Andrew W. Mellon Foundation for support in scoping the problems and possible solutions to the challenges we face in changing our collecting role in the 21st century. Such a study will allow us to determine what infrastructure would be appropriate for these new functions.

c. Project Description

Because of the lack of relevant skills in-house as well as time constraints, the project will utilize several outside consultants. Three separate consultancies are needed because the skill sets are quite different and not, after our own investigations, discoverable in one or even two persons. These skill sets are 1) collection development expertise, knowledge of art publishing and focus group facilitation 2) detailed technical knowledge of selective web harvesting 3) in-depth knowledge of library systems and workflows.

Selection Criteria		
Consultancy 1	Consultancy 2	Consultancy 3
At least 5 years experience of collection development preferably in an art environment	Practical experience of web-harvesting	Knowledge of the library automation landscape, products, companies, trends
Knowledge of current trends in art publishing	Awareness of IP, ethical and data-protection issues in web-harvesting	Ability to execute gap analysis
Awareness of digital developments in art publishing	Experience of access and preservation implications	Ability to analyze workflows
Ability to facilitate workshops	Knowledge of web-harvesting initiatives around the world	Proposal writing
At least two years experience as consultant	Report writing	At least two years experience as consultant

The consultancies are cumulative as to a great degree each one will inform the following one.

1) One consultant will be employed to help determine the 'tipping point' from analog to digital for specialist art resources. The consultant will review existing 'universe of publications' studies, and interview key market leaders e.g. Pace Galleries, Adelson Galleries for e-catalogue raisonnés; Sotheby's, Christies, Bonham's for e-sales catalogues; the Smithsonian Institution for virtual exhibitions; the Public Catalogue Foundation (UK) etc. The consultant would also run focus groups with NYARC curatorial staff to help determine researcher requirements, and with IT staff who are starting to be involved in electronic archives management or even collecting web-based art.

2) A consultant (likely to be different from the first) will review existing selective web archiving projects (Columbia University, Minerva/Library of Congress, Harvard University, Yale University, UK Web Archive) and approaches e.g. Archive-It/Internet Archive, the Web Archiving Service of the California Digital Library. He or she will also review the Frick Art Reference Library's Archive-It auction catalogs project. He or she will recommend to NYARC a) what it can and should collect both from a content and technical aspect b) the best methods of web archiving c) what partners (technological, publisher, other research libraries or preservation consortia it should be working with, and d) how to address IP, ethical and access issues.

3) A third consultant will review the existing technical infrastructure in light of the new technical requirements, review technical solutions, recommend the best technical solution, and prepare a funding bid to achieve this.

The consultants will report to Deborah Kempe, Chief of Collections Management and Access at the Frick Art Reference Library. Ms. Kempe will also serve as Principal Investigator for the project, and will liaise with the Legal Counsel to the Web Archiving Program at Columbia, and commission legal advice when necessary. Monthly project meetings and additional regular meetings will ensure that management goals are achieved and timetables met. She will oversee the project budget and issue reports for the NYARC directors' monthly conference calls. She will summarize the recommendations that are the outcomes of the grant in one report, with the three consultants' reports appearing as appendices.

In addition to the involvement of NYARC curators and IT staff, NYARC library staff will be involved: NYARC content selectors will contribute to the collection development strand; interns and others involved in the Frick Art Reference Library Archive-It project to the second strand; and technical services, acquisitions, cataloging staff will input into the third strand.

The project will take 12 months.

Time Period	Activities	Objectives	Person responsible
Month 1	Drafting of terms of consultancies 1-3;		Principal Investigator (PI)
Month 2	Selection and hiring of Consultant 1 (C1)		PI
Month 3	Consultancy 1 starts	Review existing 'universe of publications' studies	C1; PI
Month 4	Hiring of Consultant 2 (C2); Consultancy 1 proceeds	Interview key market leaders on supply-side	C1; PI
Month 5	Consultancy 1 concludes; C1 Report; Consultancy 2 starts	Focus group of NYARC curators and researchers	C1; C2; PI
Month 6	Hiring of Consultancy 3 (C3); Consultancy 2 proceeds	Desk-top review of selective web archiving projects from harvesting to preservation	C2; PI
Month 7	Consultancy 2 concludes; Consultancy 2 Report; Consultancy 3 starts	Prepare collecting guidelines. Review permissions letters and process. Commission legal advice. Identify technical requirements.	C2; C3; PI
Month 8	Consultancy 3 proceeds	Identify system reconfiguration needs.	C3; PI
Month 9	Consultancy 3 proceeds;	Review of technical solutions available	C3; PI
Month 10	Consultancy 3 proceeds	Recommend technical solution	C3; PI; NYARC Directors (ND)
Month 11	Consultancy 3 Report	Prepare Funding Bid.	C3; PI

Month 12	Report Writing. Dissemination Activities	Report & dissemination events	PI; ND
----------	---	-------------------------------------	--------

d. Expected Outcomes and Benefits

- (1) Determination of the 'tipping point' from analog to digital for specialist art resources
- (2) Recommendation to NYARC as to a) what it should collect b) the best methods of web archiving c) what partners (technological, publisher, other research libraries or preservation consortia) it should be working with, and d) how to address intellectual property, ethical and access issues
- (3) Recommendation of appropriate changes to technical infrastructure to support digital capture, resource discovery, preservation and access
- (4) Preparation of grant proposal for technical infrastructure upgrading
- (5) Dissemination of results to the US art libraries community

e. Intellectual Property Issues

Intellectual property issues are involved in the archiving of and giving access to the electronic resources harvested. These will be addressed during the project.

No intellectual property in software or technology or new content will be created.

f. Long-term Sustainability

It is anticipated that staffing requirements for collecting the born digital will be handled by adjusting existing staff responsibilities so that no new staff will be required.

The three institutions are committed to sharing equally the running costs of any new technical infrastructure.

g. Reporting

The Principal Investigator, Debbie Kempe, will submit a final report to the Andrew W. Mellon Foundation three months after the end of the grant term i.e. 30 April 2013.

The criteria to evaluate the progress of the project will be whether the deliverables are successfully achieved.

The report will include:

Description of the Project and Purpose of Grant

Summary of Progress achieved

Summary and analysis of the consultants' recommendations

Setbacks and challenges

Significant staff changes if these occur during the project

Recent related seminar papers, news articles, blogs etc.

Sustainability of results

Financial Report and commentary on the grant expenditures