Interim Performance Report
Grant Number PR-50182-13
Preservation and Access Framework for Digital Art Objects
Oya Y. Rieger and Tim Murray, co-PIs & Mickey Casad, Project Manager
Cornell University Library
April 23, 2014

The project aims to develop a technical framework and associated tools to facilitate enduring access to interactive digital media art with a focus on artworks stored on hard drive, CD-ROM, and DVD-ROM. Cornell University Library's Rose Goldsen Archive of New Media Art, recognized globally as a prominent archive for such artworks, provides the test bed for the study. Our ultimate goal is to create a preservation and access practice for complex digital assets that is based on a thorough and practical understanding of the requirements from the perspectives of collection curators and users, as well as the characteristics of digital objects. The goal of this interim report is to outline the key accomplishments of the project team during the first year of the project and report changes to our original plan based on our progress and findings. Appendix A includes a sample set of working documents to illustrate our in-depth work.

Proposed Goals and Accomplishments

- We began the project reviewing our work plan and staffing configuration. We held the first advisory group meeting, which was an influential discussion to help us decide on hardware, software, operating systems, and specifications for a digital forensics workstation. We also discussed comparable initiatives and use cases. Advisors also strongly recommended emulation as feasible and scalable access strategy. Their expert recommendations guided our initial revisions to the work plan.

- We developed a media arts survey targeting researcher, artists, and curators to allow better understanding of user profiles and use scenarios (Appendix B). The questionnaire took longer than expected to create and disseminate, in part because of the complexity of our requirements. We needed a singular questionnaire that would directly engage a wide variety of disciplinary interests without assumptions or bias. We had an excellent return of 183 responses, representing a cross-section of disciplinary perspectives and allow better understanding of user profiles. The variety of responses to our questions about access confirmed that, in order to identify artworks' most significant properties for access versions, we must solicit direct input from the artists wherever possible. We have added this as a new goal for the second project year.

- At the center of our initial efforts was analyzing CD-ROM and Internet Art to determine classes and groupings, and selecting subsets of classes of material to test based on broad impact, feasibility, and scholarly value. Complementary to this effort was the identification of appropriate data models for documentation of classes and representation information. These processes have overlapped and informed each other throughout the first project year. Thus far our analysis has focused on CD-ROM artworks already cataloged in the Goldsen core collection (100 discs representing approximately 200 artworks). We began with metadata derived from artworks' existing MARC catalog

records and compared published and observed system requirements, updating MARC records as necessary. We ran the artworks on both contemporary and legacy hardware, documenting any problems, and took notes on the artworks' rendering in legacy environments to use as a baseline for emulation tests. We then created disk images—byte-perfect preservation copies—of all artworks in the test collection, noting any problems.[1] This overview led to the identification of classes based on operating systems, common software dependencies or plug-ins, the presence of audio-only formatting on a disc (so-called "hybrid" discs), or the need for an active Internet connection.

- The key purpose behind defining content, metadata, and packaging specifications for each identified classes has been supporting the development of Submission Information Package (SIP) methodology and validation definition for each class. Working with our archival repository manager, we have made progress toward the above goals much faster than anticipated. We have created an ingest agreement and a model for required files and metadata, including technical and descriptive metadata for disk images and individual file-level metadata for discs (Appendix A2). This model incorporates the initial findings of our emulation experiments.

- We analyzed a subset of complex objects within selected classes to identify and document significant properties (required metadata for documenting descriptive and rights information) and test our data model. We are strongly inclined to adopt RDF; however, we will keep our data model open until the middle of the second project year so that we can incorporate findings from artist interviews and questionnaire. We are currently working to simplify the technical metadata of this draft model, in conversation with our consultants AVPS.

- In an effort to share the outcomes of our initiative with the technical, library, museum, art, and scholarly communities, we posted a press release about the project on several mailing lists and set up a preliminary wiki. As described in Appendix C3, our team has already presented about the project in several forums.

Changes to Original Work Plan and Methodology

- Based on our Advisory Group's strong suggestion, we began exploring emulation as an access strategy early in the first project year. We recognize that these investigations must be rigorous and systematic. We have taken emulation into account in devising our data model and repository ingest protocols. As a result, we are factoring emulation platforms in our assessment of dependencies and obsolescence risks, including metadata specific to emulators in our data model, and preserving emulators as well as artworks in our

---

[1] We have used two main utilities for the disk imaging process: ISOBuster and Guymager. Guymager provides more thorough output files for technical metadata, which can be imported directly into our metadata framework. Of the two, ISOBuster is more adaptable and therefore more appropriate for hybrid discs or discs with initial imaging problems. Also, we have identified a number of digital forensics utilities that automatically capture file-level technical metadata from disc images: fiwalk, hls, disktype, FITS, and fido. We are exploring the Digital Forensics XML (DFXML) standard, which appears able to accommodate disparate and overlapping outputs.

repository. Working from our initial collection assessment, we selected several emulators for experimentation and have tested approximately 25% of the works in our target collection.[2] This experience has driven the identification of additional salient "classes," for example, compatibility of certain artworks with certain emulators, or patterns of rendering problems. First-hand experience with the imperfect rendering provided by most emulators gives us a heightened sense of how important it is to identify aesthetic "significant properties" for individual artworks—if possible, in conversation with their creators. This will inform the curatorial matrix and "best feasible" access versions we create in the second project year.

- Our Digital Forensics Analyst has an excellent set of skills, combining metadata and computing technology (such as legacy formats, file systems, operating systems and applications). Her preparedness to assume this important position has far exceeded our initial expectations for this role. This has saved the project many consultation hours. Our consultant partners from AVPS are still invaluable contributors, but their role has focused on metadata frameworks, emulation, and complex test cases. We have revised the project work plan and budget to concentrate AVPS' consultation role on ontology and restoring problem disc images. Thus far our investigations have focused on CD-ROM artworks, though it was always our intention to examine other formats as well (DVD-ROM and hard drive) with the workflows and data model developed for CD-ROM. To accomplish these tests, we will keep the Digital Forensics Analyst at 50% FTE for the final project quarter. Originally, we were planning to reduce this effort to 25% during the final three months of the project. We have furthermore decided to maintain this position beyond the duration of the project and are in the process of exploring how to accomplish this goal. This decision will further strengthen our sustainability plans and help us mainstream the methodologies developed through the NEH-funded project.

- Digital content that is not used is prone to neglect and oversight. Reliable access mechanisms are essential to the ongoing usability of digital assets. Therefore, in addition to our NEH-funded preservation activities, we are making an extra effort within the scope of existing Library staffing to introduce and embed new media art materials from the test collection in teaching and research at Cornell and elsewhere.

- We now recognize that identifying the most significant properties of individual media artworks will require direct input from artists. This confirms our belief that we need to push the integration of archival protocols as far upstream as possible, to the point of content creation. We plan to adapt pre-existing conservation-oriented questionnaires to our emerging data model and our growing sense of media art "classes," and solicit the contributions of artists in the test collection. We will simultaneously revisit rights agreements that never anticipated access strategies based on emulation. In order to complete this work revealed by the first part of our project, we will expand the appointment of the Collections Analysis Assistant from 25% to 50% during the last nine months of the project.

---

[2] We have been using BasiliskII and SheepShaver to emulate Macintosh 68K processors and PowerPC processors, respectively. For Windows works, we have had great success with VirtualBox and QEMU.

APPENDICES

A: SAMPLE PROJECT DOCUMENTS

- A1: Salient Classes of Digital Artworks
- A2: Pre-Ingest agreement and Work Plan for Cornell University Library Archival Repository (CULAR)

B: SURVEY INSTRUMENT

- B1: Sample Message Distributed to Mailing Lists and Individuals to Elicit Responses to the Survey of Media Art Researchers
- B2: Survey Distributed to Media Art Researchers
- B3: Sample Screenshot from the Online Survey of Media Art Researchers
- B4: Notice of IRB Exemption

C: DISSEMINATION AND PUBLICITY

- C1. Initial Press Release
- C2. Screenshot of Project Website
- C3. Related Presentations and Publications
- C4. Sample Project Presentation

APPENDIX A: SAMPLE PROJECT DOCUMENTS

A1: Salient Classes of Digital Artworks

The following are some of the "classes" and categories of digital media artworks on CD-ROM that we have identified through three types of collection analysis:
- Initial system analysis and test run on legacy hardware
- Disk imaging, using utilities Guymager and ISObuster
- Emulation tests, using BasiliskII and SheepShaver (Mac), VirtualBox and qemu (Windows)

This list represents an overview based on our broad gathering of information and observations. We note some potential applications and implications for these categories in a preservation workflow.

Storage format

>CD-ROM
>>Artist-burned (CD-R)
>>"pressed" (CD)
>>*Of these two categories, CD-R discs are simultaneously less stable and more likely to be unique, and therefore imply a greater level of preservation effort*
>DVD-ROM
>Hard Drive

File systems on disc

>HFS
>HFS+
>ISO9660
>Joliet
>Joliet (short)
>UDF
>Coexistence of multiple filesystems
>>*In addition to marking the obsolescence risk of each of these file system types, archivists should note that many discs include multiple filesystems – that is to say, the disc can be used in different operating systems. The coexistence of multiple filesystems effectively provides redundant versions of an artwork and multiplies opportunities for access. Discs encoded with a single file system should be seen as a greater preservation risk, particularly in cases where that file system is no longer actively supported.*

Installation requirement

>Run from disk
>Installation required
>>*Many artworks require installation and cannot be run directly from the disk. This will entail development of a specialized workflow for access.*

Hybrid Discs

Works that include data tracks and audio tracks on the same disc are relatively rare in our test collection, but require greater efforts to image and access

Software, browsers, and plug-ins

Artworks created with Macromedia Director

Macromedia Director-based works are ubiquitous across our test collection, yet pose distinct conservation challenges. Macromedia Director-compiled works appear with no source files, which reduces the potential for restoration. This is one reason why we determined that our preservation framework must include direct input from artists: among other things, we will ask artists about their original compiling environments and the availability of source files.

Browser based artworks

VRML (Virtual Reality Markup Language)

Other HTML –related works

Browser-based artworks offer excellent restoration potential, as the basic markup languages themselves are relatively straightforward. However, dependency on a browser—at times, a browser or version that may be obsolete—introduces new risks and compromises into the preservation and access strategy for such works.

Plug-ins / Media players

Specific version of player

Included on disc

Many artworks in the test collection require specific versions of media player plug-ins. In some cases, these are packaged alongside the artwork on the disc; in other cases, the plug-ins must be located elsewhere and installed for the artwork to operate properly. All of these factors will shape the process and effort required to provide access to the artwork.

Dependence on active Internet connection

Conservation indicators

Disk imaging and rendering problems

Problems / bugs in rendering test on legacy hardware

Disk imaging errors / "bad sectors"

Emulation compatibility

Processor speed / media playback conflict

A2: Pre-Ingest agreement and Work Plan for Cornell University Library Archival Repository (CULAR)

Partners
1. Tim Murray and Oya Rieger are co-PIs of the NEH grant that created these digital objects. They are informed of this deposit, but do not otherwise need to be involved.
2. Madeleine (Mickey) Casad is the associate curator of the Rose Goldsen Archive of New Media Art and will represent as the active steward for CULAR.
3. The curator will inform the artists as to the deposit as necessary.

Fitness for CULAR - PASS
1. General nature of the collection: The digital objects for this deposit are the output of a grant from the National Endowment for the Humanities (NEH) awarded for the Cornell University Library proposal "Preservation & Access Framework for Digital Art Objects" in 2012. The proposal concerns itself with the research of the effort required to adequately preserve a selection of works from the Rose Goldsen Archive of New Media Art (http://goldsen.library.cornell.edu), an archive of complex born-digital materials that trace history of artistic experimentation in digital media. The work of the grant seeks to create contemporary emulation environments for artworks selected from the archive, to classify works according to type and document research discoveries regarding the preservation effort. All aspects of this work (disk images, compiled emulators, ROMs, Operating Systems, virtual disks, notes, reports, etc.) will need to be preserved and will be part of this CULAR project/collection.
2. Desire for deposit: CULAR is expected to assist in the following ways.
    a. The works in question are currently located on CD-ROMs and DVD ROMs. CULAR will keep disk images stable and secure, addressing the threat of media rot and media obsolescence of the selected works. CULAR will also preserve related source code of operating systems and ROMS, compiled editors, and all documentation of the project.
    b. Delivery of these new media art objects will be accomplished through the archive through onsite mechanisms. It is foreseen that all relevant images, emulators, operating systems, ROMS and virtual disks necessary for delivery will be onsite, with no need to address the objects within CULAR directly to accomplish delivery.
    c. CULAR is expected to perform the function of backup and preservation of master images and derivative use copies. Access would be limited to cases where a fresh copy of any of these is required in the delivery instance.
    d. The active steward may desire to download XML MD of parts or the whole of the deposited assets, or descriptors for creating a local list of files within an image with checksums for ascertaining and/or remediating local use copies.
3. Grant output is currently backed up in daily increments using Cornell's EZBackup Service.
4. This collection is not currently preserved elsewhere.

Technical Notes – *(Will reevaluate these assumptions closer to ingest)*
1. The overall deposit will consist of master and use copy images of approximately 80 born-digital works stored natively on CDs and DVDs, source code for compiled emulators, ROMS, relevant OS versions, virtual disks, notes and metadata. See detail at the end of this document for description of deposit structure and file types of deposit objects.
2. Initial deposit size will be approximately 60GB in aggregate.
3. File system dates will not require preservation. Imaged disks will already package the dates of the original files internal to the disk images. Dates of images themselves will not require documentation, and basic timeframe of grant will adequately date project output.
4. There is no sensitive data contained in this deposit, although some of the objects are covered by commercial copyright. ***Need to check intellectual property issues for the works themselves after contacting artists.***
5. Ingest will occur towards the end of the project, anticipated to be Summer/Fall of 2014.
6. Ingest is envisioned as one large initial deposit (perhaps chunked into multiple initial deposits at the convenience of programmers), although possibly there may be periodic ingests to CULAR for this project:
    - Subsequent deposits may be added to this project if other similar works are preserved.
    - Subsequent documentation may be added after the bulk of the disk images, emulators, etc. are ingested (a sort of phased initial ingest.)
    - If emulation strategy changes there will be need to update emulators and use copies. We may need to rely on CULAR versioning to organize/mark versions and manage subsequent ingests.
7. ***Structure of ingest is described immediately below***. ***The structure of the bullet points reflects the aggregate structure anticipated in CULAR***. All will be contained within an aggregate for the depositor (named Goldsen), and the collection/project (named *ComplexBornDigWorks*). Each aggregate will contain a text file immediately within that describes the structure and contents in full.

Key:
(A) – Aggregate object
(R) – Resource object
(M) – Metadata object (understood to be descriptive to object immediately preceding i)

Expected Structure within CULAR
1. **Curation Docs** (A)
    a. Documentation of ingest plan (2 versions of this document) (R) ODT and PDF/A
    b. Grant application (R) – PDF/A
    c. Survey results (R) – PDF/A, CSV
    d. Whitepaper (R) – PDF/A
2. **CompiledEmulators** (A)
    a. Description of plan for getting emulation to work, components required and workplan (R) – TXT
    b. Child folders (A) for each emulator, each named after emulator name (Basilisk, SheepShaver, etc.)

   i. Metadata (M) that shows relationship between ROMs and emulators. (Schema unknown – Jason will check.)

   ii. Source code of specific emulators, (binaries written in C++) (R) – TAR.GZ

   iii. Descriptions (R) of how they were built, compiled, specific utility, customizations, limitations, etc. These will be in an open format (TXT and/or PDF/A).

   iv. Virtual "drives" (R) that represent an OS running on an emulator (200-300 mb apiece; 5-6 per emulator.)

   v. List of ROMS that are compatible (R) - TXT

3. **OperatingSystems** (A) will contain as little structure as necessary to keep ingest tidy. <u>EITHER</u>

  a. Operating system code for each for each preserved operating system (R) – ISO

  b. Notes (R) regarding peculiarities of the OS – TXT or PDF/A
   <u>OR</u>

  c. Child folders (A) or named after operating system it contains, containing

    i. Operating system code (R) – ISO

  d. Notes (R) regarding peculiarities of the OS – TXT or PDF/A

4. **ROMS** (A)

  a. ROM files (one file per ROM) (R) **-** BIN

   i. ***XML (M) will describe each ROM. (Schema unknown – Jason will check.)***

5. **Notes** (A – optional; use if items will not fit under _CurationDocs) – Project should rely on separate deposit to eCommons for delivery solution if items are meant to be publicly accessible.

  a. Any documentation of the research that relates to the project as a whole. (R) - TXT, PDF/A, other open formats.

6. **Works** (A) – this aggregate will contain

  a. Various folders (A) named after the Voyager BibID of the works they contain **.**

   i. Bibliographic Record of work from Voyager (M) – MARCXML
   NOTE: UNRESOLVED issue as to whether deposit will have need of both the Master folder and the UseCopy folder. Dianne will resolve this in conversation with Alex and Mickey. The solution will take into account Alex's best advice, whether the image creation process is truly a bite perfect copy of the physical disk (and therefore representative of the artifact) and the desires of Mickey as curator. Depending on the result, the (M) items below may need to be placed in the UseCopy folder instead, and/or that folder renamed.

  b. Masters (A)

   i. Disk images; RAW data; 2352 bytes per block, a bite perfect image. One image for each physical disk. (R) – BIN

   ii. Structure maps for each image (M) – XML ***(Schema TBD –will include the output of fiwalk, hls and disktype tools.)***

   iii. File manifests for each disk image (M) including original file system names dates, sizes, PRONOM entries, etc. – XML ***(Schema TBD –will incorporate the output of FITS, FIDO and HLS tools.)***

      iv. Rendering documentation (M) image specific info regarding any problems during imaging (damaged sectors, etc.) – XML *(Schema TBD)*

c. UseCopy (A)
  - i. Disk images derived from the BINs in the Master folder (R) – ISO
  - ii. Notes regarding playability; image specific anomalies, exceptions; any detail about problems during imaging (damaged sectors, etc.) Any configs specific to that piece of artwork (WINE) (M) – XML *(Schema TBD – Jason will check.)*

d. UpVersions (optional) (A)
  - i. Derivative use copies of each image in UseCopy that incorporates upgrades operating system equivalent if possible; affects only those works that play on the HFS file system natively; will make images upgraded to HFS+ (R) – **ISO**
  - ii. Notes regarding playability; image specific anomalies, exceptions; any detail about problems during imaging (damaged sectors, etc.) (M) – XML *(Schema TBD – Jason will check.)*

e. Covers (A)
  - i. Image files; photo documentation of original media, packaging, inserts, etc. (R) – TIFF, JPEG
  - ii. Description of abovementioned files (M) – XML *(Schema TBD – Jason will check.)*

APPENDIX B: SURVEY INSTRUMENT

B1: Sample Message Distributed to Mailing Lists and Individuals to Elicit Responses to The Survey of Media Art Researchers

Dear Colleagues,

We would appreciate your assistance with an NEH-funded digital media art preservation project currently underway at Cornell University. This project aims to develop scalable preservation strategies for complex, interactive, born-digital media artworks, using the collections of Cornell's Rose Goldsen Archive of New Media Art as a test bed (http://goldsen.library.cornell.edu/).

In developing a preservation framework that will address the needs of the broadest range of archive users, we seek the input of artists, researchers, educators, curators, and others who work with interactive digital artworks and artifacts. Would you please take a few minutes to respond to this questionnaire about your practices?

https://cornell.qualtrics.com/SE/?SID=SV_6mPEBGQWr2K4nmR

Depending on your responses, we estimate that this questionnaire will take 10-25 minutes to complete.

Information about questionnaire results will be published and made available to the broader media archives community. Information about this preservation initiative is available at:

http://news.cornell.edu/stories/2013/02/humanities-grant-helps-library-preserve-digital-art

Feel free to contact Mickey Casad, project manager, at mir9@cornell.edu for more information.

Many thanks for your help with this investigation, and apologies for any cross-postings.

Yours on behalf of the project team,

Madeleine Casad
Associate Curator, The Rose Goldsen Archive of New Media Art
Digital Scholarship & Preservation Services
Cornell University Library

B2. Survey Distributed to Media Art Researchers

The following is a PDF version of the online survey we distributed to media art researchers.

Introduction

Thank you for responding to this questionnaire about your media art research and preservation practices.

Your responses will help guide an NEH-funded project to create Preservation and Access Frameworks for Digital Art Objects in the holdings of Cornell University Library's Rose Goldsen Archive of New Media Art. If you would like more information about this initiative, or about the Goldsen Archive, please share your contact information at the end of the questionnaire.

The questionnaire begins by asking you to decide whether to respond as an individual researcher or as a representative of an archiving institution. You will have an opportunity to return to this decision and take both versions of the questionnaire, if you choose.

We estimate that it will take 15-30 minutes to complete these questions, depending on your responses.

○ Respond as an individual researcher / practitioner

○ Respond on behalf of an archive / museum / cultural heritage institution

The respondent's selection determines which set of questions appear next--"Individual Researcher" (I.), or "Museums/ Archives" (II.). In either case, the respondent will have an opportunity to answer the second set of questions.

INDIVIDUAL RESEARCHER

What is your name?

[   ]

Which of the following best describe your engagement with media art? (Select any that apply)

☐ Researcher / Writer

☐ Educator

☐ Curator   The respondent's selections here trigger questions targeted to educators, curators, artists, etc., later in the questionnaire.

☐ Artist

☐ Other (please describe below)

[   ]

When you visit collections of digital media artworks and artifacts, do you have a specific research agenda or question in mind?

○ Always

○ Often   Questions marked in green are targeted to researchers, but appeared to all respondents.

○ Rarely

○ Never

Please describe some of the questions that have guided your research with digital media artworks and artifacts.

[   ]

What kinds of digital artifacts do you work with most often?

[   ]

What is your ideal scenario for interacting with archive-owned versions of such materials? (For example, your response could reference spatial surroundings, hardware and peripherals, control over settings, availability of physical media, remote or networked access, or anything that seems especially important to you.)

What are the most serious impediments you have encountered in researching such materials?

Which of the following genres or content types do you emphasize in your media art research? (Check any that apply.)

- ☐ Interactive artist portfolios or compendia of works
- ☐ Installation / performance / media sculpture
- ☐ Video / cinema
- ☐ Code
- ☐ Hypertext
- ☐ Games
- ☐ Interactive narrative
- ☐ Expanded cinema
- ☐ Virtual reality / augmented reality
- ☐ Other (please describe below)

The respondent's selections here appear again to be ranked later in the questionnaire.

Use this space to elaborate on any of your responses from the previous question, if desired.

Which of the following platforms, forms, formats, or physical media do you emphasize in your media art research? (Check any that apply.)

- ☐ CD-ROM or DVD-ROM
- ☐ Floppy discs
- ☐ Hardware / peripherals
- ☐ Personal computers or devices
- ☐ Web-based artworks
- ☐ Locative media
- ☐ Installation / sculpture / performance
- ☐ Participatory artworks
- ☐ Specific software or filetypes (please describe below)
- ☐ Other (please describe below)

The respondent's selections here appear again to be ranked later in the questionnaire.

Use this space to elaborate on any of your responses from the previous question, if desired.

[text box]

Which of the following research frameworks do you emphasize in your media art research? (Check any that apply.)

☐ Cultural or art-historical contexts (includes exhibition and reception histories, artists' writings, and other work)

☐ Social contexts and histories

☐ Technological contexts and histories

☐
Production histories

☐ Media theory

☐ Aesthetic theory

☐ Other (please elaborate below)

[text box]

The respondent's selections here appear again to be ranked later in the questionnaire.

Use this space to elaborate on any of your responses from the previous question, if desired.

[text box]

Please rank your selected genres or content types in order of importance. (1 = most important)

1 2 3 4 5 6 7 8 9 10

» Interactive artist portfolios or compendia of works

» Installation / performance / media sculpture

» Video / cinema

» Code

» Hypertext

» Games

» Interactive narrative

» Expanded cinema

» Virtual reality / augmented reality

» Other (please describe below)

[text box]

A respondent would see only those genres he or she had selected as relevant earlier in the questionnaire.

Please rank your selected platforms, forms, formats, or physical media in order of importance. (1 = most important)

1 2 3 4 5 6 7 8 9 10

» CD-ROM or DVD-ROM

» Floppy discs

» Hardware / peripherals

» Personal computers or devices

» Web-based artworks

» Locative media

» Installation / sculpture / performance

» Participatory artworks

» Specific software or filetypes (please describe below)

[text box]

» Other (please describe below)

[text box]

A respondent would see only those platforms he or she had selected as relevant earlier in the questionnaire.

15

Please rank your selected research frameworks in order of importance. (1 = most important)

1 2 3 4 5 6 7

» Cultural or art-historical contexts (includes exhibition and reception histories, artists' writings, and other work)   ○ ○ ○ ○ ○ ○ ○

» Social contexts and histories   ○ ○ ○ ○ ○ ○ ○

» Technological contexts and histories   ○ ○ ○ ○ ○ ○ ○

A respondent would see only those frameworks he or she had selected as relevant earlier in the questionnaire.

» Production histories   ○ ○ ○ ○ ○ ○ ○

» Media theory   ○ ○ ○ ○ ○ ○ ○

» Aesthetic theory   ○ ○ ○ ○ ○ ○ ○

» Other (please elaborate below)   ○ ○ ○ ○ ○ ○ ○

In cases where full interactive access to a complex digital art object is not possible, what documentation strategies work the best for your research purposes? Please cite specific examples if possible.

You indicate that you are an educator. In what contexts and subject areas does your teaching involve interactive digital media artworks and artifacts? (Select any that apply.)

☐ Post-secondary education (colleges, universities, post-secondary art or technical schools); please describe subject areas

☐ Primary or secondary educational institutions (ages 5-18); please describe subject areas

Questions marked in blue appeared only to respondents who described themselves as educators earlier in the questionnaire.

☐ Museum, gallery, or archival settings; please describe below

☐ Short-duration workshops, seminars, or lectures; please describe below

☐ Other; please describe below

What kinds of interactive digital artworks or artifacts do you emphasize in your teaching?

How do you use, assign, or reference such materials in your teaching?

What impediments have you found to teaching interactive digital media materials?

16

What aspects of these works are most important for your students to understand?

What kinds of technical and archival support would most enhance your teaching with digital media materials?

As an artist, do you create interactive digital media artworks? Please elaborate on your response as desired.

○ Yes

○ No

Questions marked in purple appeared only to respondents who identified themselves as artists earlier in the questionnaire.

What preservation measures have you taken with your own digital work?

Note that these questions target preservation concerns in general. The project team is creating a follow-up interview questionnaire for artists in the Goldsen collections, to address more specific technical and aesthetic concerns.

What are your biggest concerns about preserving your own media artwork?

How do practices of archiving and accessing complex digital media artworks affect you most, with regard to your own creative and professional work?

As a curator, are you responsible for interactive born-digital artworks and artifacts? If yes, what kinds of works are in your collections?

○ Yes

○ No

Questions marked in pink appeared to respondents who identified themselves as curators earlier in the questionnaire. This section is very short, as most of our curator-focused questions appear in the second part of the questionnaire ("Museums/ Archives/Libraries").

As a curator, what are your biggest concerns about preserving access to older digital artworks?

==Please use the space below to share any additional comments about this survey or its subject matter.==

==Would you be willing for members of this preservation and access initiative to contact you in the future?==

○ Yes--use this contact information:

○ I would rather not be contacted.

==Thank you for your responses. You may stop here, or retake the questionnaire on behalf of a library, archive, museum, or other cultural heritage institution with which you are affiliated.==

○ Finish questionnaire.

○ Retake questionnaire as a representative of a library, archive, museum, or other cultural heritage institution.

MUSEUMS / ARCHIVES / LIBRARIES

The following series of questions appeared to respondents who selected "respond on behalf of an archive…" in the introductory question.

What is your name?

What institution do you represent?

What is your role there?

==Does your institution include born-digital interactive media artworks and artifacts in its holdings?==

Respondents' selection here determined which questions they would see next.

○ Yes

○ No

==You indicate that your institution does not include born-digital interactive media artworks in its holdings. What are the main reasons why not? (Check any that apply.)==

☐ Such materials fall outside collecting scope

☐ Intellectual property questions too complex

☐ Procedures for providing access too complex or unsustainable

☐ Procedures for preservation or conservation too unclear or technologically challenging

☐ Lack of technological infrastructure

☐ Lack of specialized staff support

☐ Lack of audience interest

☐ Other (please describe below)

These selections marked in blue appeared only to respondents who answered 'no' to the previous question

18

☐ Other (please describe below)

Use this space to elaborate on your responses from the previous question, if desired.

What kinds of interactive born-digital materials do you have, and approximately how many items fall into these categories?

These questions marked in pink appeared only to respondents who answered 'yes' to the earlier question.

How do you currently provide access to these materials?

What preservation strategies do you currently employ for these materials?

What conservation measures do you take, and when?

What measures do you take to secure access, preservation, and migration rights to digital materials in your holdings? (This questionnaire aims to learn more about your interactve born-digital media art preservation practices, but other kinds of collections may be relevant here as well.)

These questions marked in green appeared to all respondents, regardless of whether their institutions hold collections of born-digital interactive media art.

What kinds of patrons or audiences seek out your media collections?

19

Do you support online access to such materials? If so, how?

Do you support on-site access to such materials? If so, how?

What are the most serious institutional challenges you face, with regard to providing continued access to born-digital interactive artifacts in your holdings?

These questions marked in pink appeared only to respondents whose institutions hold collections of born-digital interactive media art

When you have questions about preservation, conservation, or providing access to interactive born-digital media objects, where do you look for answers?

Please use the space below to share any additional comments about this survey or its subject matter.

Would you be willing for members of this preservation and access initiative to contact you in the future?

◯ Yes--use this contact information:

◯ I would rather not be contacted.

Thank you for your responses. You may stop here, or retake the questionnaire from your perspective as an individual researcher rather than a representative of an archival institution.

◯ Finish questionnaire.

◯ Retake the questionnaire, responding as an individual researcher.

Thank you for your responses. The information you have shared will be enormously helpful for our project of developing Preservation and Access Frameworks for Digital Art Objects in the collections of Cornell University's Rose Goldsen Archive of New Media Art.

B3. Sample Screenshot from the Online Survey of Media Art Researchers

Cornell University

Q58. Which of the following genres or content types do you emphasize in your media art research? (Check any that apply.)

- ☐ Interactive artist portfolios or compendia of works
- ☐ Installation / performance / media sculpture
- ☐ Video / cinema
- ☐ Code
- ☐ Hypertext
- ☐ Games
- ☐ Interactive narrative
- ☐ Expanded cinema
- ☐ Virtual reality / augmented reality
- ☐ Other (please describe below)

Q60. Use this space to elaborate on any of your responses from the previous question, if desired.

>>

B4. Notice of IRB Exemption

Cornell University
Office of
Research Integrity and Assurance

East Hill Office Building, Suite 320
395 Pine Tree Road
Ithaca, NY 14850
p. 607-254-5162
f. 607-255-0758
www.irb.cornell.edu

Institutional Review Board for Human Participants

Concurrence of Exemption

To: Madeleine Casad

From: Amita Verma, Director, ORIA

Approval Date: January 10, 2014

Protocol ID#: 1401004398

Protocol Title: Questionnaire on usage and preferences of media art archives patrons and researchers

Your above referenced request for **Exemption from IRB Review** has been approved according to Cornell IRB Policy #2 and under paragraph 2 of the Department of Health and Human Services Code of Federal Regulations 45CFR 46.101(b).

Please note the following:

- ° Investigators are responsible for ensuring that the welfare of research subjects is protected and that methods used and information provided to gain participant consent are appropriate to the activity. Please familiarize yourself with and conduct the research in accordance with the ethical standards of the Belmont Report (http://www.hhs.gov/ohrp/policy/belmont.html).
- ° Investigators are responsible for notifying the IRB office of change or amendments to the protocol and acquiring approval or concurrence **BEFORE** their implementation.
- ° Progress reports or requests for continuation of approval are not required for thie study.

For questions related to this application or for IRB review procedures, please contact the IRB office at irbexemptions@cornell.edu or 254-5162. Visit the IRB website at www.irb.cornell.edu for policies, procedures, FAQs, forms, and other helpful information about Cornell's Human Participant Research Program.

Please download the latest forms from the IRB website www.irb.cornell.edu/forms/ for each submission.

APPENDIX C: DISSEMINATION AND PUBLICITY

C1. Initial Press Release
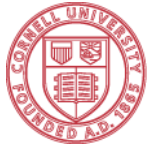
 Cornell University

Feb. 19, 2013

Humanities grant helps library preserve digital art

By
Gwen Glazer
grg59@cornell.edu

How can librarians protect the historical record, now that archives include digital images, audiovisual files, photographs and manuscripts?

A $300,000 grant from the National Endowment for the Humanities (NEH) will allow the library to begin to develop a framework to ensure continued access to complex digital media objects, using the interactive born-digital artworks in the library's Rose Goldsen Archive of New Media Art as a test site.

The NEH grant is the year's largest in New York state and one of the largest in the country -- a fact that recognizes how vulnerable and valuable the Goldsen Archive's collection is.

Timothy Murray, a principal investigator on the grant and curator of the Goldsen Archive, said: "The grant acknowledges the cultural importance of the archive's 'born digital' collections -- those online or stored on hard drives, CD-ROM and DVD-ROM. Those formats will be a particular focus for the new technical framework and associated tools this grant will help the library develop."

Despite its "new" label, new media art has a rich 40-year history, making obsolescence and loss of cultural history an imminent risk. To play a media artwork requires machines that are themselves vulnerable to technological obsolescence. This is especially true of digital art, which requires hardware and software support and is often stored in extremely fragile formats.

The Goldsen's collection includes interactive work from the two most crucial decades in the emergence of digital media art, 1991 to the present, tracing media culture's shift from disk-based to networked and Web-based applications.

"Some of the digital artworks in Goldsen are designed for ephemeral experiences," said Associate University Librarian Oya Rieger, the other principal investigator on the grant.

"Reproduction of an artwork's digital files does not always ensure preservation of its most important cultural content. It is essential that we anticipate the needs of future researchers and acknowledge the core experiences that need to be captured to preserve these artifacts."

The library will develop an archival strategy based on understanding what users need to use digital artworks. Eventually, it will create generalizable new media preservation and access practices that are applicable for different institutional types and sizes.

The preservation model to be developed will apply not only to new media artworks but to other digital media environments. Beyond the Goldsen Archive, the project will inform digital preservation services at the library and help explain how rich media objects are used in learning, teaching, research and creative expression by scholars and students.

Housed in the library's Division of Rare and Manuscript Collections, the Goldsen Archive was founded in 2002 to amass an international collection of artwork and research materials on CD-ROM, DVD-ROM, videotape and the Internet. It is one of the most prominent new media archives in the world, containing thousands of digital and video artworks and encompassing a wide variety of formats and research materials.

The Goldsen Archive is also the repository of the annual competition in new media art funded by the Rockefeller Foundation and the video art collection of the Experimental Television Center. It keeps the country's largest collections of Chinese and Taiwanese electronic art, as well as significant holdings from Australia and Europe, and serves as the repository of extensive historical collections of video and new media art created in Cyprus.

The library will collaborate with AudioVisual Preservation Solutions, and the project will have an advisory board composed of international leaders in curation, arts and preservation.
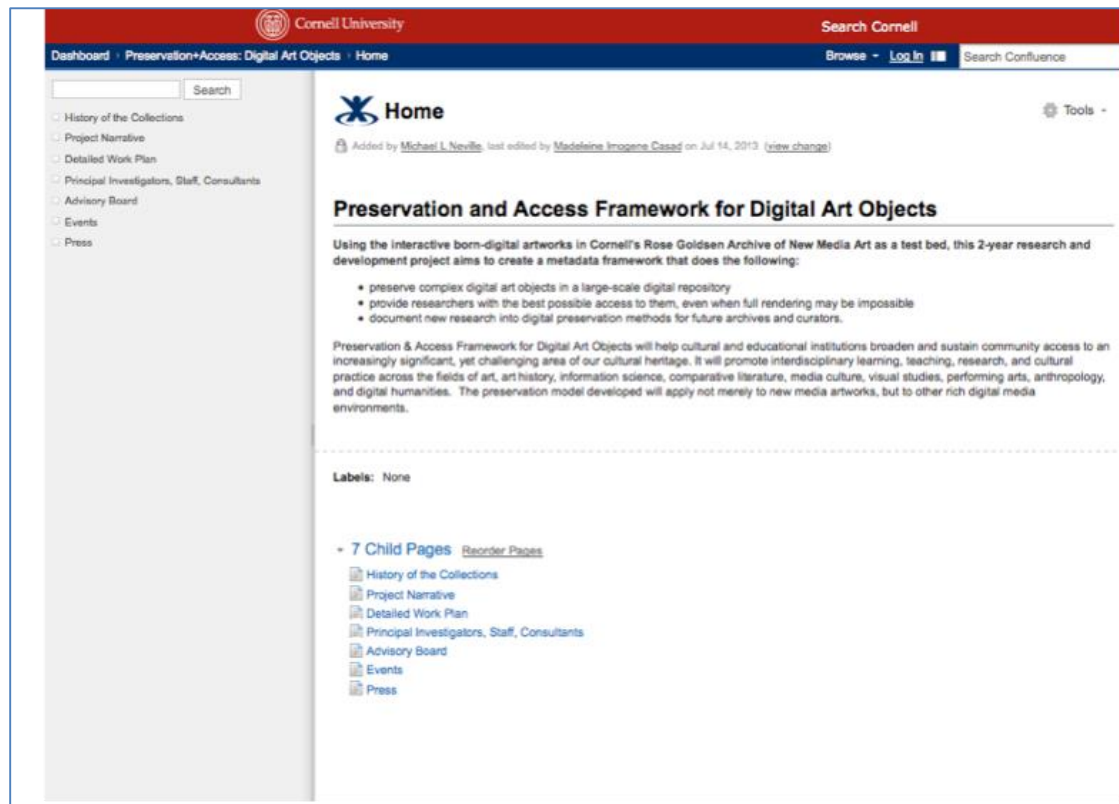
Gwen Glazer is the staff writer for Cornell University Library.

Find more Cornell news online at news.cornell.edu.

Source URL: http://news.cornell.edu/stories/2013/02/humanities-grant-helps-library-preserve-digital-art

C2. Screenshot of Project Website

C.3. Related Presentations and Publications

Fino-Radin, Ben; Lacinak, Chris; Mericle, Danielle; Minard, Johnatan. "Archiving Complex Digital Artworks." New York Archives Conference, LIU Post Campus of Long Island University, Brookville, NY, 7 June 2013. Conference Presentation.

Murray, Timothy, "Comparative (New) Medias: Screening Digital Memory Between East and West," ACL(X), Conference of the American Comparative Literature Association, Pennsylvania State, University, University Park, Pa, September, 2013. Conference Presentation.

Murray, Timothy, "Philosophy of Technology and Medial Practice," Department of Film Studies, Dongguk University, Seoul, Korea, October, 2013. Invitational lecture.

Rieger, Oya. Future Opportunities for Collaboration. *Aligning National Approaches to Digital Preservation II* (Barcelona. Spain), November 2013

Murray, Timothy, "Archival Baroque: Shifting the Tableau of Memory to the Scene of New Media," Department of Comparative Studies, Ohio State University, Columbus, Ohio,

November, 2013. Invitational lecture.

Murray, Timothy, "New Media Misfits: Artistic Anarchive," A Collection of Misfits Symposium, Carnegie Museum of Art, Pittsburgh, Pa., November, 2013. Symposium presentation.

Murray, Timothy, "Deep Time or Distributed Surface: Alternative Approaches to Archival and Media Theory," Department of English, Syracuse University, November, 2013. Invitational lecture.

Murray, Timothy, with Renate Ferro, "Archive, Memory, and Artistic Practice," College of Humanities, Arts, and Social Science, Nanyang Technological University, Singapore, January, 2014. Invitational lecture.

Alexander, Desiree and Dianne Dietrich. "The Rose Goldsen Archive of New Media Art at Cornell University." Technology Experiments in Art, Washington DC, 17 January 2014. Symposium Presentation.

C4. Sample Project Presentation

The Rose Goldsen Archive of New Media Art at Cornell University

Desiree Alexander
Dianne Dietrich

Technology Experiments in Art
17 January 2014

Cornell University

NATIONAL ENDOWMENT FOR THE
Humanities

History of the archive

Founded in 2002 to create a home and place of research for growing collections of new media art

Its conception coincided with exhibition: Contact Zones, The Art of CD-ROM, curated by Timothy Murray.

contact zones:
the art of cd-rom
@ Cornell University
March 9 - April 16, 1999
An International Exhibition Curated by Timothy Murray

Sites:
Herbert F. Johnson Museum of Art
A. D. White House
Carpenter, Mann, Olin, Uris Libraries
Cornell Information Technologies/MVR Lab

Catalogue & Schedules:
Http://contactzones.cit.cornell.edu/

Sponsored by:
The Society for the Humanities
Academic Technology Services/CIT
Herbert F. Johnson Museum of Art
Graduate Program in Film & Video Studies
Technology Connection at the Campus Store
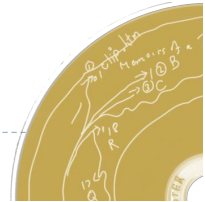Apple Computer

Cornell University

NATIONAL ENDOWMENT FOR THE
Humanities

Housed in Cornell's Rare and Manuscript Collection

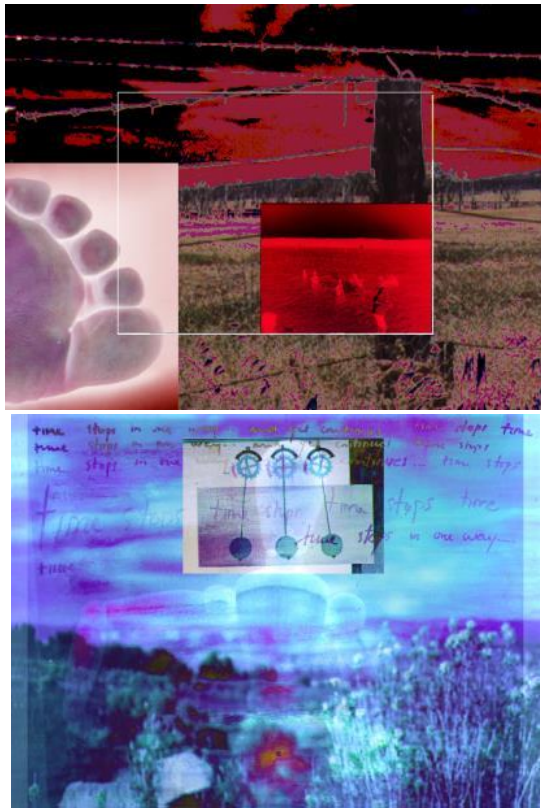**Library environment allows for discoverability & access**

Supports the variety and scope of the Goldsen collections

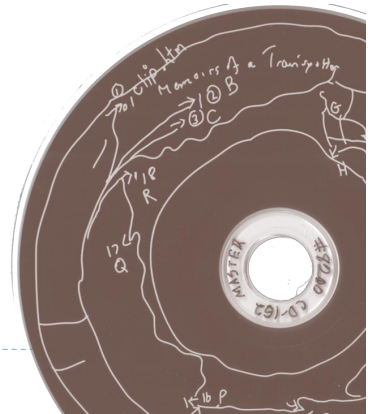Cornell University — NATIONAL ENDOWMENT FOR THE Humanities

---

Complexity of Works

Visual, Aural and Textual Interplay
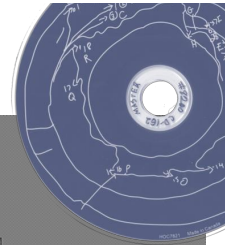
Non-Linear Form

Storage Capacity

Cornell University — NATIONAL ENDOWMENT FOR THE Humanities

Scope of collection

"Macintosh System 7.0 and up"

"Macintosh with 680x0 series processor"

"Internet connection and web browser with Flash and Shockwave"

"PowerPC processor-based Macintosh computer"

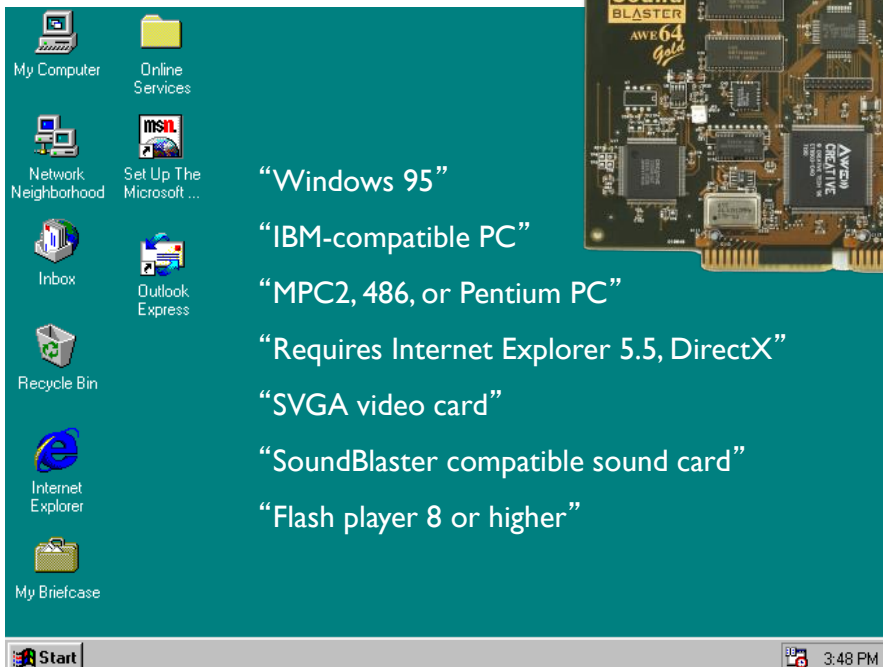"16 bit color and 13 inch monitor"

"Quicktime 2.1 or better"

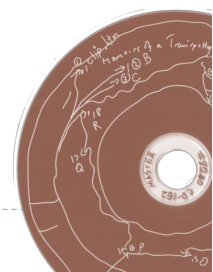Cornell University • NATIONAL ENDOWMENT FOR THE HUMANITIES

Scope of collection

"Windows 95"

"IBM-compatible PC"

"MPC2, 486, or Pentium PC"

"Requires Internet Explorer 5.5, DirectX"

"SVGA video card"

"SoundBlaster compatible sound card"

"Flash player 8 or higher"

Cornell University • NATIONAL ENDOWMENT FOR THE HUMANITIES

Preservation & Access Framework

Develop methodology for analysis of complex objects and their dependencies

Define and capture significant properties

Develop ingest protocols for technical, descriptive and structural metadata

Profile Media Art Researchers and their needs

Cornell University   NATIONAL ENDOWMENT FOR THE Humanities

Preserving the objects – Photography

Preservation of the entire object

Maintain original context and artist's intentions

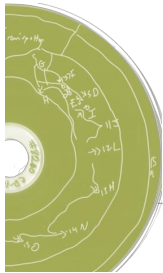**Develop naming conventions for consistency and integration**

Cornell University   NATIONAL ENDOWMENT FOR THE Humanities

Preserving the objects – Disk imaging

Format	← 2,352 byte sector structure→						
CD digital audio	2,352 (Digital audio)						
CD-ROM Mode 1	12 Sync pattern	3 Address	1 Mode, 0x01	2,048 Data	4*	8**	276***
CD-ROM Mode 2	12 Sync pattern	3 Address	1 Mode, 0x02	2,336 Data			

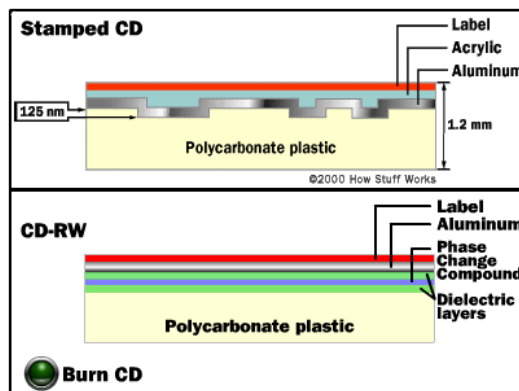* Error detection
** Reserved, zero
*** Error correction

Preserving the objects – Disk imaging

CDs and CD-Rs

File systems

- **ISO 9660**
- **Joliet**
- **HFS**
- **HFS+**
- **UDF**
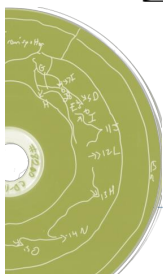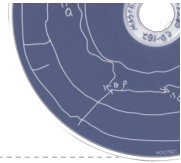
Preserving access – Emulation

Cornell University

NATIONAL ENDOWMENT FOR THE
Humanities

Preserving access – Emulation
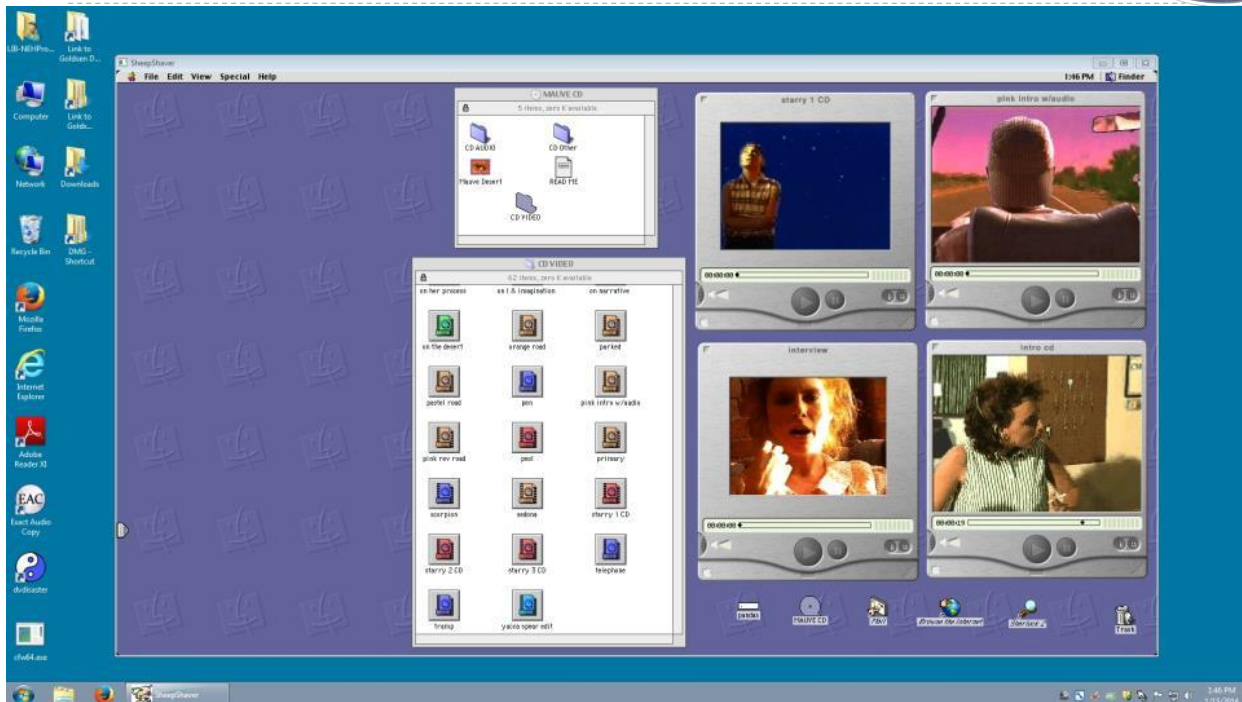
Cornell University

NATIONAL ENDOWMENT FOR THE
Humanities
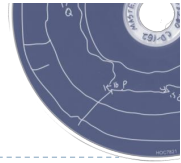
Preserving access – Emulation

It is not enough to preserve individual files, we need to preserve the experience and interactivity of these works

Emulate as many platforms as feasible to give users the opportunity to see the work in various systems

Cornell University

NATIONAL ENDOWMENT FOR THE
Humanities

Preserving access – Emulation

Emulators are often enthusiast-driven software.
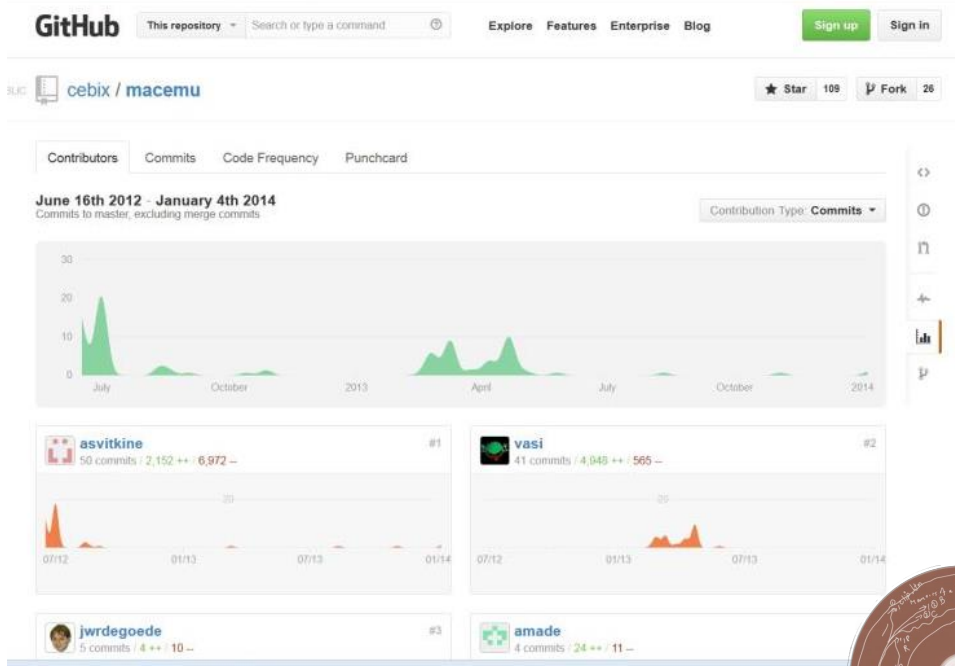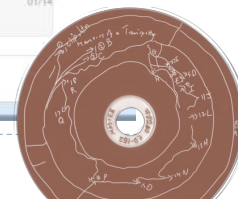
Cornell University

NATIONAL ENDOWMENT FOR THE
Humanities

Preserving access – Classifying

▸ **Does this need an emulator to run?**

▸ Does this run in a web browser?

▸ Does this work need to be "installed" to a computer (or emulated system) in order to run?

▸ What plugins or extra software is required to run the work? Are these still available?

▸ What issues have we run into running these works on modern systems? (Audio, video, overall speed.)

▸ Are there any disconnects between the official system requirements and what we observed running the work?

Cornell University   NATIONAL ENDOWMENT FOR THE Humanities

Preserving access – Classifying

▸ Does this need an emulator to run?

▸ **Does this run in a web browser?**

▸ Does this work need to be "installed" to a computer (or emulated system) in order to run?

▸ What plugins or extra software is required to run the work? Are these still available?

▸ What issues have we run into running these works on modern systems? (Audio, video, overall speed.)

▸ Are there any disconnects between the official system requirements and what we observed running the work?

Cornell University   NATIONAL ENDOWMENT FOR THE Humanities

Preserving access – Classifying

▸ Does this need an emulator to run?

▸ Does this run in a web browser?

▸ **Does this work need to be "installed" to a computer (or emulated system) in order to run?**

▸ What plugins or extra software is required to run the work? Are these still available?

▸ What issues have we run into running these works on modern systems? (Audio, video, overall speed.)

▸ Are there any disconnects between the official system requirements and what we observed running the work?

▸   Cornell University   NATIONAL ENDOWMENT FOR THE Humanities

---

Preserving access – Classifying

▸ Does this need an emulator to run?

▸ Does this run in a web browser?

▸ Does this work need to be "installed" to a computer (or emulated system) in order to run?

▸ **What plugins or extra software is required to run the work? Are these still available?**

▸ What issues have we run into running these works on modern systems? (Audio, video, overall speed.)

▸ Are there any disconnects between the official system requirements and what we observed running the work?
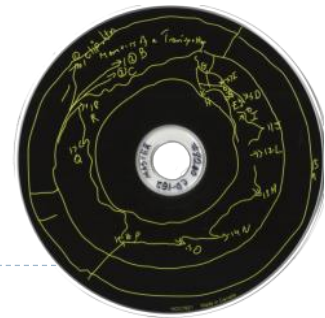
▸   Cornell University   NATIONAL ENDOWMENT FOR THE Humanities

Preserving access – Classifying

- Does this need an emulator to run?
- Does this run in a web browser?
- Does this work need to be "installed" to a computer (or emulated system) in order to run?
- What plugins or extra software is required to run the work? Are these still available?
- **What issues have we run into running these works on modern systems? (Audio, video, overall speed.)**
- Are there any disconnects between the official system requirements and what we observed running the work?
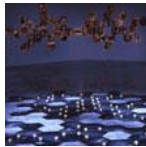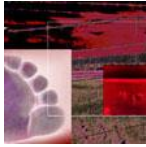
Cornell University   NATIONAL ENDOWMENT FOR THE Humanities

Preserving access – Classifying

- Does this need an emulator to run?
- Does this run in a web browser?
- Does this work need to be "installed" to a computer (or emulated system) in order to run?
- What plugins or extra software is required to run the work? Are these still available?
- What issues have we run into running these works on modern systems? (Audio, video, overall speed.)
- **Are there any disconnects between the official system requirements and what we observed running the work?**

Cornell University   NATIONAL ENDOWMENT FOR THE Humanities

Preserving the collection – Archiving

Ultimately, we want to preserve these works

- as objects (CD-ROMs)
- as software and digital files
- as experiential, interactive works

CORNELL UNIVERSITY LIBRARY ARCHIVAL REPOSITORY

Cornell University

NATIONAL ENDOWMENT FOR THE Humanities

Gathering feedback

Cornell University

NATIONAL ENDOWMENT FOR THE Humanities

Thank you

Cornell University · NATIONAL ENDOWMENT FOR THE Humanities

- Cornell PAFDAO Team:
  Desiree Alexander, Madeleine Casad, Dianne Dietrich, Jason Kovari, Danielle Mericle, Liz Muller, Michelle Paolillo
- Grant PIs: Oya Rieger, Timothy Murray
- Consultants:
  AVPreserve: Alex Duryee, Chris Lacinak, Kara VanMalssen
- Advisory Board: Ben Fino-Radin, Jean Gagnon, Rebecca Guenther, Matthew Kirschenbaum, Jon Ippolito, Norie Neumark, Christiane Paul, Richard Rinehart, Simeon Warner
- This work supported by a grant from the National Endowment for the Humanities

---

Image credits (in order of appearance)

- Physical CD – David Thomas; The Encoded Eye, The Archive, and its Engine House
- Contact Zones poster – Valerie McMillen; https://contactzones.cit.cornell.edu/index.html
- Rare and Manuscript Collections (interior) – http://rmc.library.cornell.edu/presidents/view_image.php?img=97
- Norie Neumark; Shock in the Ear – (screenshot from Contact Zones webpage: https://contactzones.cit.cornell.edu/artists/neumark.html/)
- Old Macs – Wikipedia; http://en.wikipedia.org/wiki/File:Macintosh_IIci.png (Public Domain) and http://en.wikipedia.org/wiki/File:Indigo_iMac_G3_slot_loading.jpg (CC By 2.0)
- "Welcome to Macintosh" – Dianne Dietrich; Screen capture of System 7 in BasiliskII emulator
- Windows 95 Desktop – Wikipedia; http://en.wikipedia.org/wiki/File:Windows_95_Desktop_screenshot.png (Used with permission from Microsoft via Wikipedia)
- Creative Labs Soundblaster card – Wikipedia; http://en.wikipedia.org/wiki/File:KL_Creative_Labs_Soundblaster_AWE64_Gold_CT4390.jpg (GPL)
- Shattered CDs – photo by lydyth; http://www.pxleyes.com/photography-picture/4b58e0e25d4d1/Dreaded-Broken--amp--Damaged-CD-s.html/)

Image credits (in order of appearance)

- Cover art—Zoe Beloff; Beyond;LTC Commune, Live Spring Scream '97
- CD-ROM Layout Table – Adapted from Wikipedia; http://en.wikipedia.org/wiki/CD-ROM
- CDs and CD-Rs Physical Structure – From HowStuffWorks.com; CD vs CD-R: http://computer.howstuffworks.com/cd-burner7.htm
- Mauve Desert in OS9 – Work: Adriene Jenik; Mauve Desert: a fiction in translation for CD-ROM. Dianne Dietrich, Screen capture. Tre Berney and Dianne Dietrich, Video capture of user interacting with work in emulation.
- Github Contribution page for SheepShaver/BasiliskII -- https://github.com/cebix/macemu/graphs/contributors (Retrieved January 10, 2014)
- Selected screen captures – Norie Neumark, Shock in the Ear; Renate Ferro, Panic Hits Home; Yao Jui Chung, The Soul Is But One Instant; Mark Hansen, Ben Rubin, Listening Post; Reginald Woolery, Million Man March/World Wide Web (From Rose Goldsen Archive of New Media Art homepage, http://goldsen.library.cornell.edu/)
- CULAR Logo – Courtesy of Michelle Paolillo