

Agenda for SCC Meeting 11-19-03 1:00 pm to 3:00 pm

Due to the quantity of supporting material and the complexity of the issues I do not expect that this agenda will be completely covered in our first meeting. I expect instead that it will be spread over the next few meetings (at least for a first pass!).

Agenda

- Welcome
- Introduction
- Brief Discussion of a Proposed Mission for the Committee
- Report on the Preserving Classic Software's Workshop
- Proposed Charter for the Committee
- Action Items from the Workshop
- Collections Overview/Current Practices
- Suggested Software Classification
- Establishing Standards
- Selecting 10 Candidates for Preservation Tests
- Cyber and Web Issues
- Preliminary List of Issues
- Next Meetings

Welcome

Introduction

- Brief attendee's introduction
- Update of committee membership list (Excel spread sheet attached)

Membership

- We decided not to use a CHM alias like software_collection@computerhistory.org
 - Do we distinguish between:
 - The attendees to meetings (in person or by telephone),
 - A wider email distribution list (in particular for people outside of the area)
 - And an even wider software collection interest group.

Brief Discussion of a Proposed Mission for the Committee

The current Vision and Mission of the Computer History Museum are as follows:

CHM Vision: "To explore the computing revolution and its impact on the human experience."

CHM Mission: "To preserve and present for posterity the artifacts and stories of the information age".

Proposed Mission for SCC: "To preserve and make accessible for posterity the software artifacts and stories of the information age" [Note we might replace "information age" with "software revolution" but it would create some repetition].

Commentary: Given the narrower scope of the Software Collection Committee a vision statement seems out of place at this point. It also makes sense, at first, to closely follow the current mission of the Computer History Museum by simply adding the word "software".

The other modification made ("make accessible") recognizes the strong distinctions between preserving software for interested specialists (and to an indirect degree for the museum public) and presenting the information to interested specialists (for their research) and to the museum public (in the form of exhibits). Clearly the collected information will have to be made accessible to interested specialists, but the public exhibit aspect is outside of the scope of this committee. Thus we have replaced the word "present" with implies both access and exhibit with "make accessible" only.

As we spent some time refining the committee's goals and challenges we may want to revisit the mission for the SCC, but at this point the proposed missions is likely to be adequate for this stage of our development.

Report on the Preserving Classic Software's Workshop

Some 17 of us met for 1 1/2 days on Thursday and Friday October 16-17 to discuss the challenges to Preserving Classic Software. The workshop was initiated, with the support of the Museum, by Grady Booch who had been focusing on that problem for some time and wanted to find a way to get started.

This workshop has been invaluable to this committee as it has allowed us to start much faster than we could have expected with a better understanding of what should be done and of the issues that will be challenging us. A brief set of highlights follows. The attached minutes provide a more detailed report.

Interest in the subject is at an all time high, but despite many local efforts nothing systematic seems to have been done yet.

The scope of the task is very large and it is only by channeling many different individual and institutional initiatives that we can succeed.

We must show a long term vision but pick few manageable tasks to debug the process.

The Computer History Museum is in a great position to lead these efforts.

Standard and processes must be defined.

We must find the resources to get started on something practical that can plow the way for more efforts (and more resources).

A web effort to provide access to indexed (metadata) digital version of all relevant software information is critical.

There may be a difference of opinion between being selective and being encompassing.

Proposed Charter for the Committee

Proposed charter:

To establish the standards (categorization, preservation steps and processes) necessary to preserve and make accessible software for posterity.

To test these standards on few critical and representative software with a view to improve the methodology.

To establish and coordinate a 10 year distributed process for the Software 10,000 preservation initiative.

To understand who to engage (people, institutions, etc...) in these efforts and the mechanisms for participating in this preservation initiative.

To determine the Computer History Museum proactive list of software to collect and coordinate their preservation.

Commentary: We believe that the Computer History Museum greatest opportunity for leadership will come from establishing a set of standards (with the help of interested communities) , testing them, and coordinating their application by interested individuals and institutions so that 10,000 pieces of software will be saved in the next 10 years.

Beyond its own interest in preserving software for its collection the museum will:

Host the resulting software 10,000 digital archive.

Function as an accession board for the distributed archive that will result.

I picked 10,000 pieces of software over ten years as goal that will stretch the community but still might be feasible. This will be true if, as we expect, many of the pieces worth saving will require a small amount of time to collect and preserve in the archive, while few will require significant amount of time.

Our effort at selecting ten test cases will allow us to validate or invalidate the statements above.

Action Items from the Workshop

1. Dick Gabriel: Collection goals, and principles guiding selection and collection.
2. Kirsten Tashev, Dag Spicer and Sharon Brunzel: What standards and collecting processes must be defined initially?
3. Dick Gabriel and Lee Courtney: How to structure an initial community of Practice for self-organized collection of software artifacts.
Who is the community, what form does it take, relationship with institutions such as CHM, structure of the community, what are the features/benefits for the community?
4. Len Shustek: Artifact Preservation List Revised

5. Grady Booch: Artifact Candidate List Categories (a software classification)
6. John Toole, Chuck Bronstein: Create Funding Opportunities
7. Ed Feigenbaum: Reach out to Professional Societies (this depends on all of the above)

Collections Overview/Current Practices

Presentation by Museum staff:

- Introduction to the Five Collections
- Collections Policy/Collecting Criteria
- Collections Processing/Taxonomies/Data Standards

Commentary: The purpose of this is to look to the Museum's current practices (primarily the hardware collection) to help provide context for developing the software collection. We should make use of the procedures the museum has been following so far and see how they might apply to the effort of rounding out our software collection.

Suggested Software Classification

Grady's suggestions:

- AI
- Commercial
- Communication
- Development
- Device
- Entertainment/sports
- Financial
- Game
- Industrial
- Legal
- Media
- Military
- Operating system
- Platform
- Scientific
- Transportation
- Utility

Establishing Standards

Representative list:

- Artifact Categories list (metadata);
- Artifact preservation list (type of objects from Len);
- Scanning standards and other media standards;

Process for submission;

Auditing process;

Software taxonomy guidelines (this may be questionable);

Selecting 10 Candidates for Preservation Tests

Commentary: I believe we should pick 10 preservation candidates to evaluate and test our preservation processes. As we do not yet understand all the efforts required for such preservation I expect that half of these candidates will be dropped as test cases and that we will pursue a much smaller number to the end of the preservation test.

Candidates should be from a different era, or of different preservation complexity, or of different preservation completeness.

Initial suggestions are:

- FORTRAN;
- Multics;
- Visicalc;
- Netscape/Mosaic;
- DOS or CPM;
- CAL TSS;
- Powerpoint;
- Etc...

Cyber Issues

What does it take to manage the software archive, distribute it standards and audit the efforts of a distributed community of individuals and institutions.

What cyber tools can we use (beyond email) to help in this committee's work.

Preliminary List of Issues

Not already included above, and to be added from today's feedback.

Resources:

- Seeding with volunteers today;
- Grant in 3 months (??);
- Public project in 6 months (??)

Focusing on CHM required actions vs. actions for the community at large.

What is software? What software do we collect?

Understanding the repository issues.

Conversion and data reading issues.

Other institutions to work with

Audiences

Access is universal except for standard IP restriction

IP issues

Tax issues with software donations

Escrow for commercial software currently on sale

PC Application Software Conference May 5-7, 2004 in Boston

Report on Museum Computer Network Conference

Etc..

Next Meetings

Wednesday December 17 from 9:00 am to 11:00 am.

NOTE: We are still establishing the dates below. They have high probabilities to be correct but still may change. We should have confirmation on 11/19/03.

Wednesday January 21 from 1:00 pm to 3:00 pm.

Wednesday February 18 from 1:00 pm to 3:00 pm.

Wednesday March 24 from 2:30 pm to 4:30 pm. [Note different time]

Wednesday April 21 from 1:00 pm to 3:00 pm.

Wednesday May 19 from 1:00 pm to 3:00 pm.

Wednesday June 23 from 2:30 pm to 4:30 pm. [Note different time]