Agenda for SCC Meeting 12-17-03 9:00 am to 11:00 am

Agenda

Proposed Charter for the Committee

Proposed Task List for the Committee

Achieving Results through Sub-committees and Moderated Discussions

Software Metadata Sub-committee

Software Taxonomy Sub-committee

Action Items Review

Museum's Software Collection Criteria

Deliberate Collecting and Collaborative Collecting

Selecting 10 Candidates for Preservation Tests

Proposed Agenda for Next Meeting

Cyber and Web Issues

Current List of Issues

Next Meetings

Proposed Charter for the Committee

For reference: 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".

We decided as other committees have done that the SCC does not have a mission only a charter, as its mission is implicitly to act in support of the museum's mission. The SCC charter may change over time as tasks are completed and the environment changes.

The charter of the SCC is:

"To develop the policies and procedures necessary to collect, preserve and archive software.

To test those standards and processes experimentally on a number of test cases.

To create an initial list of software that should be preserved in the archive.

To get interested groups and individuals collaborating with us on this preservation initiative.

To augment the resources needed to fulfill the Museum's mission of preserving software."

Proposed Task List for the Committee

This is a very preliminary list of tasks which will be improved, in the next few weeks, by the feedback from the staff and the volunteers. We will also determine the proper and most efficient way to handle them as well as the resources necessary (dollars and people) to achieve them.

Preliminary task list (no order implied):

Develop software taxonomy;
Establish the metadata structures necessary for archiving our software;
Develop the other internal standards and processes necessary for archiving software;
Determine a list of representative software to test our procedures;
Collect these representative software to test and improve our procedure;
Index our software collection;
Create a proactive list of software to be collected;
Define procedures to transcode software and media;
Create software acceptance procedures;
Prototype the software digital archive;
Find resources to hire people and cover the expenses for these tasks;
Determine the success factors for the SCC and the museum;
Help other interested groups to collaborate with us;
Establish significant collaborations with a selected group of interested parties;
Establish a process to preserve the software 10,000;
Estimate the cost and priorities for the above tasks;
Focus on an achievable list of tasks for the next 12 months.

Achieving Results through Sub-committees and Moderated Discussions

Due to the size of the SCC and the inherent difficulties in achieving results in a committee environment we plan to use sub committees and moderated discussions to complete most of the tasks required to fulfill our charter.

Sub-committees will have few active volunteers or staff members and a specific task to achieve.

They will be responsible for determining the scope of the task, its deliverable and schedule.

They will present status, proposed directions and results to the full SCC committee.

We are starting with two subcommittees: one on metadata and another one on software taxonomy.

Software Metadata Sub-committee

Mary Cicalese and Sharon Brunzel have accepted to work on the Metadata sub-committee. Mary will be the lead person on that sub-committee. They will try to recruit outside specialists and other interested parties to join them in that effort and will give us a status report at the next SCC meeting.

Software Taxonomy Sub Committee

Paul McJones, Michael Blasgen and Dag Spicer have accepted to work on the Software Taxonomy sub-committee. They will try to recruit other interested parties to join them. They will also decide who will be the lead person on that sub-committee. They will give us a status report at the next SCC meeting.

Commentary: Coming up with software taxonomy could easily become a bottomless pit, so we should try to be realistic, use the work done by other (like IEEE or ACM) and expect that the taxonomy will change and will be improved over time. Grady made some related suggestions for characterizing the software in the PCS 100 list. It was based on the market view of the software. The ACM and IEEE tend to espouse a technical view of the software classification.

Grady's suggestions:

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

Action Items Review

Please refer to the improved spreadsheet used to tack action items.

Museum's Software Collection Criteria

An initial presentation by Kirsten Tashev and the staff of the Museum to get the SCC members up to speed on what is being proposed. We will discuss the subject in greater depth at the next meeting of the SCC.

Deliberate Collecting and Collaborative Collecting

An initial presentation by Richard Gabriel. We will discuss the subject in greater depth at the next meeting of the SCC.

Selecting 10 Candidates for Preservation Tests

I was very pleased by the significant feedback we received concerning this action item. A separate document to be emailed soon will summarize the responses.

You should note that there is a significant distinction between the programs that ought to be preserved and those that we will use for this initial "experimentation". Despite my efforts to explain the distinction (see below) I got the impression that most of the responses where focused on some kind of a list of the top 10 most important software items to preserve. At the same time the energy displayed by all these responses is great and I am glad that we gave it a forum. I also noted that some good names came up that did not exist Grady's list and we got a list of software matched with volunteers that seem to have energy to invest in them. I am starting to see that this passion may be one of the most important factor for the test selection. We are clearly learning from this dialogue and after all the experimentation will need manpower and we are finding who is interested.

As you may remember we are trying to select 10 preservation candidates to evaluate and test our preservation processes. As we do not yet understand all the efforts required for such preservation we should expect that up to half of theses candidates will be dropped as test cases and that we may pursue a much smaller number to the end of the preservation tests.

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

As I tried to decide how to pick the test candidates I focused on four important factors for choosing them.

1) Their overall importance to the history of computing.

Would the software to be tested belong to what I call the Software 100, the Software 1,000, or the software 10,000. Belonging to each of the category is clearly a complex judgment call but I believe many software can be easily pegged FOR THE PURPOSE OF THIS TEST ONLY, in each of the category.

2) Their functional category

We can look at functional categories as the higher level of a software taxonomy. Is the software a programming language, an operating system, a word processor, a spreadsheet, a graphic software, UI software, a networking software, a game, a complex computational problem, etc? Again at this stage of the process we may not need to fully agree on a software taxonomy to create this short list of categories.

3) Their age

It is easy to believe that software from the 1950, 1960, 1970, etc. represent different level of preservation difficulties as well as urgency. So it makes sense to factor that in.

4) The availability of knowledgeable volunteers and original designers

We should have easy access to the required resources otherwise the test will not happen!

Proposed Agenda for Next Meeting

Revised Task List

Sub-committee Reports

Other Action Items

Software Collection Criteria

Selecting 10 Candidates for Preservation Tests

Cyber and Web Issues

Cyber and Web 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.

Current List of Issues

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.

Scanning standards and other media standards;

Process for submission;

Auditing process;

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 different time]

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]

Bernard L. PEUTO, Ph.D.
Chairman Software Collection Committee
Concord Consulting
1333 Jones Street #1106
San Francisco, Ca 94109-4118
Phone: (415) 775-7801
Fax: (415) 775-7265
blpeuto@peuto.com