

## **Minutes for SCC meeting of March 24, 2004**

**by Dave Redell**

### **Attending**

Mike Blasgen  
Stewart Cheifet  
Lee Courtney  
Dick Gabriel  
Philip Gust  
Dave Gustavson  
Brewster Kahle

John Mashey  
Paul McJones  
Chacko Neroth  
Bernard Peuto  
Mike Powell  
Dave Redell  
Dave Rossetti

Len Shustek  
Dag Spicer  
John Toole  
Ed Taft  
Kirsten Tashev  
Mike Walton

### **Status: Taxonomy Subcommittee (Paul Mcjones)**

They will write a taxonomy “guide” and circulate it for review. Target a first draft suitable for review by the April meeting.

### **Status: Metadata Subcommittee**

Neither Mary nor Sharon was in attendance. Last meeting, there was a proposal to clarify the goals of the metadata effort. Target a more detailed presentation at the April meeting.

### **Status: Software Room**

At the February meeting, we visited the software room, and several members expressed concern about the disorganized state of the collection. There are now drafts of a task list and a job description for the new position of Software Curator.

### **Presentation: Proposed Curation Roadmap (Lee Courtney)**

Lee presented a proposal for how the software curation process should go forward. He distributed the related documents and solicited feedback and discussion at the April meeting. (See those documents for details, which will not be repeated here.)

Lee noted that “curation” refers to the process of “understanding what is in the collection”, through a set of processes and procedures collectively termed the “Artifact Lifecycle”. He cited the Taxonomy and Metadata efforts as providing the concrete foundation for the Artifact Lifecycle.

Len made a suggestion for jump-starting the process, based on the division of the cataloging activity into four steps: 1) Access to bits on physical media, 2) Interpretation of data formats to extract the content, 3) Definition of the repository structure, and 4) Population of the Repository with the bits from step 2. He pointed out that steps 3 and 4 can be started in parallel with 1 and 2 by initially populating the Repository with “easy stuff” that does not require handling of delicate/obsolete media and obscure data formats.

There was also some discussion of the similarities and differences between our efforts and the Internet Archive. A critical distinction is our emphasis on older (pre-Internet) software and the extra challenges that this introduces.

Ed raised the question of integrating documentation and other related materials into the software collection. Kirsten explained that the Museum’s collections and their associated curation processes are defined by *artifact type*, so documents and other materials that are logically related to items in the software collection will not be lumped into the software

collection. Instead, appropriate database linkages will be made to capture the rich set of cross-collection relationships.

### **Web Tools (Mike Walton)**

Mike described the upcoming use of a *Sharepoint meeting area* as the online collaboration vehicle for the committee. This facility is not quite ready for use, Mike and Bernard will discuss status, with the goal of having this available by the April meeting. Chacko volunteered to act as editor for materials archived on the committee site.

### **Software Curator Position (John Toole)**

John discussed the goal of having this position included in the next CHM budget. A proposal was submitted to the NSF in January which could provide at least partial funding for the position; no word on this proposal had been received from NSF as of the March meeting.

### **Community of Practice (Dick Gabriel)**

Dick presented his view on the likely structure for a *Community of Practice* – a community larger than the Computer History Museum. He prefaced his remarks with a statement that he was describing his view of the “end state” of an evolutionary process, and not a necessary or likely initial step toward that eventual situation.

Dick discussed some of the complexities and sensitive issues surrounding the relationship of the CHM effort to the many existing independent efforts at software preservation, including many small, special-interest “grassroots” projects. The whole issue of large-scale structure and governance motivated some members to question whether the approach described by Dick was too heavyweight. He responded by reminding us that he was sketching his view of a likely end state, not trying to dictate how things must be. There was general consensus that whether one agrees with all of the details of Dick’s vision or not, it is important to figure out initial steps to begin the evolutionary process. One proposal that showed significant support was the idea of a Software Preservation Workshop that would be sponsored by the Museum and to take place in perhaps six months. The idea would be to bring together the most active participants in the existing efforts and explore the advantages of a “federation” approach, which would help coordinate and rationalize the somewhat scattered nature of current efforts by encouraging sharing of methodology, lessons learned, handling of IP issues, and so on. In addition, the Museum can offer these smaller groups a permanent repository into which their work could eventually migrate, insuring that it will live on indefinitely.

Dick’s presentation went into considerable detail, which will not be repeated here. A careful reading of Dick’s and Lee’s slides side-by-side is a recommended exercise for the reader.

### **Action Items**

**Note:** Bernard announced that we will stop maintaining a separate Action Item list, and will instead fold the new/open action items into each meeting’s minutes.

*Taxonomy subcommittee:* Prepare first draft of the “taxonomy guide” for review by April meeting.

*Metadata subcommittee:* Present status of metadata project at April meeting.

*Mike Walton:* Get SharePoint meeting area ready for use by April meeting.

*Chacko Neroth:* Assume duties as editor of materials on SCC site.

*Dick Gabriel:* Draft proposal for a Software Preservation Workshop in approximately six months to jump-start the Community of Practice vision.