Minutes for SCC meeting of April 27, 2005

by Dave Redell

Attending

Rudy Batties	Sellam Ismail	Michael Powell
Henry Gladney	Jim McClure	Dave Redell
Kathe Gust	Paul McJones	Bill Selmeier
Phil Gust	Randy Neff	Len Shustek
Dave Gustavson	Paula Newman	Kirsten Tashev
Gardner Hendrie	Bernard Peuto	John Toole

SCC Website

Last month, we heard that the SCC website was coming to life. This month, Mike Powell gave us a tour of the site, which he, Mike Walton and others have been working on. The site is based on the Plone document management server running on the Zope server platform. It has three main goals:

1) serve to coordinate SCC activities (meetings, projects, etc);
2) act as a repository – both a resource for now and a prototype for a permanent facility; and
3) present a public face of the SCC and its activities to the outside world.

As an example of the repository function, we viewed the section of the site housing Paul's Fortran materials. This is already a sufficiently impressive collection that its visible presence on the SCC website should serve as an inspiration for other SCC projects. The site is located at: http://community.computerhistory.org/scc. SCC members are encouraged to register new personal accounts on the site. (There will be no automatic migration of existing CHM accounts onto the server.) Since Plone's handling of passwords is not exactly mil-spec, users are encouraged to choose a different password for this account. To register, send email to scc-admin@computerhistory.org, as described on the site. There is a public area with "published" content, which is visible to everyone, while other non-public areas are available to members with accounts. Each registered user has a home folder, and various navigation, search and content-creation tools are available. The site is still under construction, and some major features remain to be provided, such as integration of the Mailman email facility into the site. It is generally felt that the site should have a more functional software repository (the current facility is really just a folder) but there are open questions about what this should be and how it relates to the larger scale repository issue for the CHM as a whole. Finally, there are also some performance issues with Plone/Zope that need to be investigated.

User Requirements for a Software Archive

Lee and Mary were to report on their study of this topic, but were both out of town, so their report will have to wait until the May meeting.

System requirements for a Software Repository

Henry reported on his study of available software packages that provide software repository functionality – i.e. the underlying layer that would serve as the foundation of a

full archive (with file formats, meta-data, end-user access tools, etc.). Henry thought that it was important to take a broad view of the offerings, and be sure to avoid stopping at the first choice that seemed viable. Having found about 66 possible packages in his search, he began to focus on an orderly process for picking among them. He found that while there are a lot of opinions available, including serious studies such as that by the National Science Board, he could not find a statement of requirements so compelling that he would regard it as it definitive. He estimates that he might have a final requirements analysis in about one year, and hopes to engage two initial reviewers for early drafts soon. As a concrete example, he is looking at Greenstone, which has a combination of desirable properties, including software maturity, usage at multiple sites, a focus on small institutions and solid documentation.

A key question, that came up in connection with Henry's work on this project, is its proper relationship to the Museum and the Software Collection Committee. The scale and scope of his study suggests that it corresponds to the long-term goal of selecting a CHM-wide facility for housing all of its digital artifacts. As this is beyond the scope of the SCC's charter, it may be that Henry's work is best viewed as a component of a longer-term project of the CHM staff. The current plan is for the staff to start looking at this in about three months, and Henry reiterated his goal of minimizing staff load – at least until the staff is ready to ramp up its efforts. At the same time, he emphasized the need to make progress on the issue, since he feels it will take a long, sustained effort to get to closure.

Recovery and preservation of historic data

Randy Neff talked about the issue of what to do once bits have been recovered from old media. A general proposal that we have discussed before is to create XML-based format definitions for archival storage that capture the structure of such data. Later processing based on these definitions could include both emulations running the original software and new software written to analyze the historic data in new ways. Len pointed out that these two approaches correspond to quite different levels of structural representation – namely the low-level physical structure (blocks, physical records) vs the high-level logical structure (e.g. indexed-sequential datasets, etc.). He suggested caution in trying to tackle the high-level structure. Phil noted that in any case, the high level structure can typically be best handled by the original software, as long as the software has been preserved and a suitable emulated execution environment is available. It was also noted that as long as the bits and their physical structure are captured, the higher-level structure implicit in the data can be extracted at a later date as needed.

Questionnaire

Jim McClure, a Masters student in Museum Studies from JFK University, requested that SCC members fill out his questionnaire on exhibiting software. He is trying to gather data on successful attempts to show "non-physical artifacts". He can be reached at either <jamesmcclure@comcast.net> or mcclure@computerhistory.org.

Project Reports:

Fortran As noted above, we saw a nice presentation of Paul's collected Fortran materials on the SCC website. He is still working on obtaining all of the appropriate permissions to use the materials, and is eagerly awaiting the availability of a "real" CHM software repository.

Lisp: This is something of a background activity for Paul and no further progress was reported at this meeting.

CDC: Dave R expects that Mike Cochran from the controlfreaks and cyber1 groups will attend the May SCC meeting.

NLS: Phil reports that good progress continues on this project, with a focus on getting a "clean clone" of Doug Englebart's running system – i.e. the system but not all of the specific content files. A project website is under construction. Phil is considering how the project can/should use the SCC website as part of its activities. In terms of obtaining the necessary permissions, discussions are underway with Boeing. Apparently BT may be involved in this as well, as it is with the old CDC materials.

1401 & PDP-1 Restoration Projects: Bernard is looking for ways to tighten the ties of the SCC with the software aspects of these hardware restoration efforts.

New/Prospective projects

Multics: An effort is going on at MIT on this (which seems like a good place for it!). CHM/SCC will stay in touch with Olin Sibert at MIT to track the project and probably host a mirror of the result.

Chess Software: Nothing concrete is going on in this area, but it was raised as an obvious potential SCC project area if a quorum of volunteers materializes.

Algol-W: CHM got a CD of this and LEN had it sent to Sellam. We hope to add it to the SCC website.

Workshop

A committee now exists to push this forward, including Sellam, Bernard, and others.

Upcoming Meetings

Day	Date	Time	Conf Room
Wednesday	May 25	1:00 pm – 3:00 pm	Hopper
Wednesday	June 22	1:00 pm – 3:00 pm	Hopper
No July Meeting			
Wednesday	August 24	1:00 pm – 3:00 pm	Hopper
Wednesday	September 21	1:00 pm – 3:00 pm	Hopper
Wednesday	October 19	1:00 pm – 3:00 pm	Hopper
Wednesday	November 16	1:00 pm – 3:00 pm	Hopper
No December Meeting			