

Minutes of SCC Meeting of July 19, 2006 Hopper Conference Room, 5:30–7:30

Submitted by Paul McJones

Attending

Jonathan Cheyer	Al Kossow	Michael L. Powell
Gardner Hendrie	Paul McJones	Bill Selmeier
Nadia Ilyin	Randall Neff	Dick Toepfer
Henry Gladney	Paula Newman	Jim Wheeler
Rebekeh Kim	Bernard L. Peuto, Chair	

Short Reports

Al Kossow:

Al has been working on infrastructure. He's now copied about 150 tapes (11 gigabytes) into the CHM RAID server; this material is formally accessioned (if not fully cataloged?). He's also copied another 14 gigabytes of materials that haven't been accessioned, including many websites.

Al had hoped to be able to read a batch of 100 or so PDP-7 and PDP-10 DEC tapes, but Paul Allen acquired them before Al could copy them. Rich Alderson (Allen's archivist) may eventually provide copies.

Al is setting up a capability to do media recovery of 7- and 9-track magnetic tapes – he will be building the necessary electronics. There is a room at the Museum dedicated to this, in addition to the two existing rooms for media storage, which Al has been reorganizing. He's bought a "dry box" and will be doing experiments in order to develop a reliable protocol for handling and reading fragile tapes.

Rebekah Kim has recently joined the staff as Software Archivist; she'll be working three days a week and will work with Al on many projects.

An interesting recent acquisition was a tape with OS/360 Revision 1, circa 1966.

Paul McJones:

Paul's recent activities include:

- writing a 5-page paper for the proceedings of the Attic & Parlor workshop
- conversing with these HOPL III authors re preserving their software:
 - Bjarne Stroustrup / C++
 - See community.computerhistory.org/scc/projects/c_plus_plus/
 - Brad Cox / Objective C
 - Simon Peyton-Jones / Haskell

- Luiz Henrique de Figueiredo / Lua
- cataloging some 15 boxes of old manuals, tech reports, etc. that Mark Halpern is donating to CHM
- tracking down early Photoshop source code (starting with version 0.54) and organizing it in an internal archive at Adobe

Paul's to-do list includes:

- Figuring out how to formally transfer digital assets to the Museum
 - Example: for Fortran there are various scanned documents, tape images, etc.
- Understanding how the top-level Fortran web page should be handled: it is a "finding aid" that adds important context, but where does the Museum keep things like this?

Bernard Peuto:

Bernard previewed Mike Powell's work preparing for upgrading our Plone installation and augmenting it with a wiki and blogs. Bernard also mentioned his belief in the importance of setting up a repository server that will be robust enough to support a true open community-based software collection activity.

Jonathan Cheyer:

The NLS project now believes they have written confirmation from all the necessary institutions (SRI, McDonald Douglas/Boeing, British Telecom, MCI/Verizon) to allow the Museum to release executable images, source code, and documentation for NLS/Augment.

Jonathan has designed XML schemas for archiving the NLS files. One is "lossy" and the other is lossless (i.e., it supports "round-tripping" back to NLS). The reason for the lossy version is that it is quite a bit easier to manipulate for subsequent transformations (such as to HTML, ASCII, etc).

The original Augment "file" format is essentially a memory map of the data, which gets written to disk. It is quite efficient on the limited processing power of a SDS 940 or PDP 10, which needed to be able to read/write sections of the document in a fast, random access way. However, the tradeoff is that it isn't the simplest way to store what is essentially just hierarchical data. The lossless version requires a file reader to navigate the memory map, using pointer arithmetic, for example. The lossy version just stores nodes sequentially, and parent/child nodes are stored using standard XML child elements.

The current "cloned" and sanitized NLS system has about 5000 files, but Doug Englebart's personal system (which was the starting point for the cloned system) includes about 15000 files. The extra files are typically in user directories, and Jonathan believes many of them are work-related and potentially of historical interest. He would like to secure the appropriate permissions to gain access to as many of these files as possible.

They are working on a “kittable” version of the system.

They are hoping to interest someone in creating a Windows driver for the chord keyset.

An anniversary is coming up in October.

Nadia Ilyin:

Nadia has recruited several Convergent Technologies veterans who will be able to devote more time to collecting hardware, software, documentation, etc. Jim Wheeler and Dick Toepfer are retired now (and thus have some time to devote to this project); they had extensive experience in hardware, manufacturing, and operations at Convergent.

Dick was a graduate student at the University of Illinois; in 1958 he got a summer job with the IBM Harvest project. Later he worked on IBM’s “Cypress” photo-storage system; later still he was the engineering manager for the original HP 3000 project.

Paula Newman:

Paula discussed her progress with the book collection project. Her goal for the project as of April included selecting an appropriate list of publications, obtaining a “context entry” for each, helping CHM obtain copies (including the appropriate edition or editions of each), and, where appropriate, making available online versions.

The “context entry” would include information such as:

Content	Sources
Author bio	Authors, author home pages, colleagues, oral histories,
Authoring background & context	IEEE Annals
Content summary	Authors, and (probably older) experts in field
Significance in software history	
Discussion of similar publications	
Reference list	anyone
Links to material on other pubs by same author	??

After spending some time studying the lists of books, Paula decided that an important first step is “clustering” the publication list by subject and date. She believes this will facilitate study of the books, and will also provide opportunities to annotate groups of related books rather than just isolated books. Her current categorization is visible in a set of prototype web pages that are currently in her personal area of the SCC Plone web site:

<http://community.computerhistory.org/scc/Members/paulan/CHBooks/>

Mike Powell will move it to the top, public level soon.

She's reasonably happy with most of her categories and subcategories, except for the "software engineering" subcategories.

She's hoping to get one of the authors to do some annotations for Marvin Minsky's *Semantic Information Processing*.

The meeting adjourned a little early.