

Minutes
CHM Software Preservation Group
May 16, 2007
by Paul McJones

Attending

Jonathan Cheyer
Lee Courtney
Henry Gladney
Bob Goldberg

Al Kossow
Paul McJones
Randy Neff
Bernard Peuto

Michael Powell
Bob Sanguedolce
Len Shustek
Dick Toepfer

Organizational change

Al Kossow is assuming chairmanship of SPG; Bernard will continue on as a member. At this point in the group's existence, there are now a number of fairly independent activities going on such as the NLS project, Henry Gladney's pilot Greenstone project, etc. Also, Al is at a point where he has some very specific tactically-oriented activities that he believes he can enlist volunteers to participate in. Al counts approximately 2.5 million files he's brought over in the last year that need to be organized. There is also a large media recovery effort (floppies: easy; 7-track tape: difficult). The archive itself needs to be put together. And the media that are being read at the block level need to be analyzed to verify that the correct bits were read. The software needs to be cataloged and, where appropriate, curated.

Len Shustek: How about other people out in the world who have acquired other software – can we attract them to donate it to us, so it will have a permanent home beyond their lifetime? This is intended to supplement what we already have in the Museum's media collection.

Report from Bob Sanguedolce

CHM web site redesign

CHM is currently on an annualized run rate to reach 1 million total online visitors in 2007 – this is the lowest cost way to reach visitors, and provides a national and international reach. CHM is the top search engine hit for “computer history”. 75% of the monthly visitors are new (at least within six month window).

Types of visitors: technically-curious (searched on computer history term); already-aware (bookmarked, etc.); involved (multi-repeat visitor).

Web strategy for FY07→FY08:

- Build with knowledge

- Work on today's site to simplify tomorrow's redesign
- Begin with the end in mind: identify target audience; make home page welcoming and exciting
- Execute our design plan: redesign CHM home page; announce Babbage Engine; Core Magazine; 2007 Fellows Awards; Semiconductor exhibit

Bob showed a mock-up of the new home page design, which provides much more information and has flexibility to support frequent updates.

IS strategy update: museum group requirements & environment

1. Working with groups to meet immediate short-term needs
 - a. Software Industry SIG (Information Technologies Corporate Histories Collection): updated for long-term maintainability
 - b. SPG: working to improve Plone website performance and future maintainability
 - c. Semiconductor SIG: developing an online Semiconductor Timeline exhibit
2. Developing overall long-term SIG strategy
3. Begin implementing in FY'08 (starts July)

Report from Bernard Peuto: Documenting software history – “Softline” project

The motivation was a request for feedback and help from the CHM Timeline Committee regarding its coverage of software. Also, we felt it would be useful to have an overview structure for software history in the large, expressing our vision and priorities for future collection/preservation work.

The museum landscape

- Many are oriented around physical artifacts with little or no attempt to explain and provide context (examples: Railroad Museum in Sacramento; Museum of Air & Space in Washington; Musee des Art Premiers in Paris; Caravaggio & Rembrandt exhibit in Amsterdam)
- Some are “institutionally oriented” with large research collections (Louvre, Guggenheim, Getty)
- Some museums have dual physical/web presence

Software is unique

- It's information: relatively intangible – physical media not really important
- Believe requires much more context than a mainframe or other hardware artifact
- Number of items: huge number of programs written for each hardware type
- Web should be very appropriate way to exhibit software

Status of Time Line project

- There's a draft document now
- It evolved from a list of artifacts
- It's stronger on structural exhibit issues
- It's weaker on message issues
- Historians who reviewed it warned of overemphasis on timeline and artifacts

Guiding principles of Softline project

- Goal is to map software history, understanding challenges and successes
- Concentrate on software that created world as we know it (need to focus)
- Provide an extensible framework
- Recruit industry and academic experts to flesh out the framework
- Focus on a knowledge framework, not on how to design exhibit

Fighting the artifact demons

- We want insight into what happened and why
- But it's easy to fall into trap of simply enumerating "software artifacts"
- Xxx

What is the framework?

- We have categories/topics/entries, with a fairly small branching factor at each level
- There can be many views of these entries, for example:
 - Single timeline
 - "Ages": mainframes, minicomputers, microcomputers
 - Enterprise, small-to-medium business, department, individual/consumer
- So far, we've chosen entries within a topic using a "5/5/5" strategy, with Tier I of the most important, Tier II of runners-up, and Tier III of "interesting but".
- We're expecting to come up with some 100-200 entries in an initial version.

Statements (we think will be demonstrated as this project is carried forward)

- Hardware's tremendous growth in features and speed has enabled corresponding growth in software features and scope
- Hardware's tremendous decrease in cost and size have opened larger and larger markets for software solutions
- Cost effectiveness ...
- Xxx

Search for an application view

- As we worked on the project, we came to see the fundamental importance of application areas, and our own relative lack of knowledge on applications (in contrast with systems software)
- We are starting to build and prioritize lists of applications, by consulting outside experts

Take Away Messages

- The forward march of computing is fed by:
 - New better, cheaper, faster and smaller hardware
 - Which opens up existing and new markets for new users and new software applications
 - Which in turn feed demands for new hardware and software
- The fundamental principle is the interaction between technologies and user needs

Next steps

- Solicit reviews and comments
- Fill in all known holes
- Post on SPG web site
- ...

Questions and comments

Henry Gladney: What you've discussed so far sounds like it's addressed to insiders. Could you come up with a theme like "the information revolution" and address it to the general public? Up until 1970, almost all computing focused on differential equations for physics/engineering, or inventory control; after that was a vast expansion in range of applications.

Jonathan Cheyer: Does your metadata format have a field for "was influenced by"?

Henry Gladney: What were the technical innovations along the way to hide complexity and make the applications manageable?

Lee Courtney: One category that you didn't list: games/entertainment. Most of your list is applications used by institutions rather than individuals. To communicate to the public, you should take something concrete and understandable to a lay person, and then drill down to show what's involved "under the covers".

Dick Toepfer: Visicalc is a good piece of software to motivate adoption of personal computers.

Henry Gladney: You might want to think of three parallel timelines: technology, application, and ???

Jonathan Cheyer: Young people care about gaming and media (e.g., music, video).

Bob Sanguedolce: Why does this list split things finely in some areas (avionics/weapon systems, various IT subdivisions) but then omit so many others?

Al Kossow: What is your schedule?

Bernard: We are talking about in the next 2-4 weeks writing up the ideas of our framework and populating the framework far enough that people can visualize our methodology and feel comfortable that it can be extended enough to be useful.

Al Kossow: If CHM does end up building a software theme room, this would be the core of it. But for the main Time Line, there can probably only be dozen topics.

Bob Sanguedolce: And we could have an online exhibit far sooner than a physical theme room.

Bob Sanguedolce: On your "Take Away Messages" slide, shouldn't the third sub-bullet of the first bullet say competition?

Al Kossow: In general, collecting actual historic application software is going to be much more difficult than collecting operating systems and programming languages.

Randy Neff: Will you put this out on a wiki?

Upcoming Meetings (2007)

Day	Date	Time	Conf Room
Wednesday	June 27	5:30 pm – 7:30 pm	Hopper
Wednesday	July 18	5:30 pm – 7:30 pm	Hopper
Wednesday	August 15	5:30 pm – 7:30 pm	Hopper
Wednesday	September 26	5:30 pm – 7:30 pm	Hopper
Wednesday	October 17	5:30 pm – 7:30 pm,	Hopper
Wednesday	November 14	5:30 pm – 7:30 pm	Hopper.
(no December Meeting)			

Next meeting

Wednesday, June 27, 2007

Proposed topic for next month

Lee Courtney suggests discussing our agreement with HP for acquiring their old software.