"The Open Collecting Movement and its Impact on Collecting Institutions"

Bernard L. Peuto
Chairman SCC

Our Key Messages

- Multiple collecting models are coexisting
  - Some with historical roots
  - Some with non-institutional roots
- Is there an emerging or dominant model?
  - Numerically it's the non-institutional model
- An Open Software Collecting revolution is taking place which is redefining the role of Collecting Institutions
- We aim to foster and join this Open Collecting Movement

It's a Matter of Degree

- But software collecting is off the scale … compared to hardware etc…
- The numbers are massive
  - 10 to 1,100 to 1,1000 to 1 for each hardware environment
- Software is basically a content experience
  - Source code & interactive experience
    - And many more
  - Viewing the content is large percentage of the exhibit experience
  - Physical artifacts are less important
- Software collecting is contextual
- The web is eminently suited for software collecting
  - Little resources needed to store and gather
  - It's all in the interpretation / curation

Two Established Models

- Reactive Collecting
  - Staff and volunteers
  - Dropped donations
  - High level cataloging
  - Later curation
  - Later exhibits
  - Research collection cyber access is unclear
- Volunteer pro-active collecting
  - Passionate individuals
  - Deliberate information and context search
  - Some cataloguing done on the flight
  - Interpretation done on the flight
  - Combination of interpretation & access as exhibit substitute
  - Cyber access done on the flight

Which Model to Choose

- Why choose one model
  - To maximize the collection and preservation of software
  - Not to exclude other models or views
- Reactive collecting
  - More institutional museum oriented
  - More reflected in attic vs. parlor imagery
  - Closer to the current state at CHM
  - Few are using it
- Non-institutional model
  - Non-institutional passionate driven web sites
  - 10s to 100s of quality sites are available
  - No permissions needed
  - Better context coverage

A Complex Set of Factors

	Attic	Parlor	Representative volunteer driven sites
Focus	Broad based	Narrow focus	Sites are typically focused, multiple sites create a very broad base
Interpretation/ Curation	None	Critical	Significant amount
Presentation	None	Critical	From small to an integral part
Community	Large Volunteer Community	Smaller professional staff	Few passionate technical experts within a larger volunteer community
Institution, Resources	Private volunteer & Museum initiatives	Institution like museum	Non-institutional
Quality or expertise		Supposed to be of quality	Technical expertise

SCC's Feedback

- No reason to favor either approach
- The style of the work is of less interest than its quality
- Refocused questions
  - Who does the work?
    - Who do we want to participate? Who are our collaborators?
  - How are rules established…quality maintained?
  - What should the software archive…look like?
  - What collections…can be rescued?
  - How do we manage [this]?
  - How do we encourage…amateur collecting activity?
  - Collaboration among multiple institutions…?
  - International participation?
- Randy Neff's Attic Parlor Feedback (attached)

Software Collection Committee

- The work of the SCC includes:
  - Preserving and collecting software
  - Identifying best practices
  - Fostering an open alliance of the many broadly-based collecting efforts taking place today.
- SCC in numbers:
  - Established March 2003 (Started in October 2003)
  - 25 meetings
  - Approximately 30 active members
  - 90 members on the distribution list

Pro-Active Pilots Projects

- Stable collecting projects
  - Fortran – Paul McJones – Essentially done
  - Lisp – Paul McJones – Essentially done
  - NLS – Phil Gust – Completing technical, legal details
- “Far along” collecting projects
  - Multics - Olin Sibert
  - PDP-1 - Al Kossow, Judith Tauber-Lovik
  - 1401 – Bill Selmeier, Ron Mak
- New collecting projects
  - APL – Christian Langreiter and Lee Courtney
  - ACM Computer Science Book Project – Paula Newman
  - Resource Directory – Dick Blaine

Paul McJones: Collecting early FORTRAN

- Goal
  - Collect FORTRAN/FORTRAN II compiler source code, documents
- Methods
  - Relentless emails and phone calls to people, institutions
  - Blog to record progress, attract comments
- Status
  - Located documents, films, *machine-readable source code*
  - Found many helpful people, resources via the web
  - Created web site at CHM
  - Working with others to get FORTRAN II running on simulator

Paul McJones: Collecting early LISP

- Goal: collect historic source code
- Located source code for many versions from Lisp 1.5 to modern era
- Created web site at CHM
  - Includes classic books by permission of MIT Press
- Hoping to acquire Stoyan archives for CHM
- Working with others to run Lisp 1.5 on simulator

Phil Gust: NLS (oNLineSystem) Project

Goals

- Ensure that NLS system is captured and preserved
- Determine and document the chain of ownership for NLS and its present legal status
- Preserve and study related hardware and designs that were created for NLS
- Make NLS system and related documents widely available for study and use, including earlier versions where possible
- Prepare summaries that enable the historians and the public to understand NLS and its contributions
- Study NLS in detail to understand and help communicate its architecture and functionality

Accomplishments

- Captured and preserved last NLS system, running on PDP-20/TOPS-20 emulation
- Determined chain of ownership, and secured releases from most past owners (Boeing, BT: *done*; SRI: *soon*; MCI: *working*)
- Identified related hardware; in process of acquiring several artifacts and related documentation (keysets, line processor).
- Creating new, supporting software and hardware (Java Augterm, NLS XML exporter, keyset reproduction).
- Creating public distribution of NLS and related documents
- Planning for cyber exhibit of NLS, including software, documents, movies, oral histories
- Partnering with Engelbart's NSF-funded Hyperscope project to study NLS in detail and create modern version

Team

- *Active*: Jonathan Cheyer, Kathe Gust, Philip Gust, Ken Harrenstein, Raylene Pak
- *Advisory*: Doug Engelbart, Jake Feinler Henry Lowood, Shinya Yamada, Peter Yim

SCC Lessons Learned

- Passionate volunteers
  - Field knowledge
  - 12 to 18 months project duration
  - Only contacts & time are needed (very precious resources indeed)
  - No permission needed
  - Many supporting opportunities
- Collecting
  - Code, executable, manuals, emulators, oral history, business history, etc …
  - The context is critical
- Institutional support
  - Museum imprimatur
  - Community training & support
  - IP rights
- Time is running out
  - The authors are still available
  - But they may not want to collect
- Contrast with a Picasso museum

Our Opinion

- The web has fostered another killer application:
  - Open content driven collecting
- For software it is:
  - Distributed
  - Non institutional
  - Passionate volunteer driven
  - Meritocracy driven
  - Delivers quality based on field knowledge
  - Integrates collection and exhibition
  - Widely accessible

A Role for Institutions

- Geographic role
  - Physical exhibiting
  - Events and conferences
- Access Role
  - Cyber Museum
  - Research collections
- Community service role
  - IP Umbrella
  - Imprimatur
  - Community training & support
  - “Burial” services