STACKTRACE // News from Software Preservation Network

View this email in your browser

STACKTRACE

NEWS FROM SOFTWARE PRESERVATION NETWORK

July/August 2018
Volume 1, Issue 3

FACES OF SPN

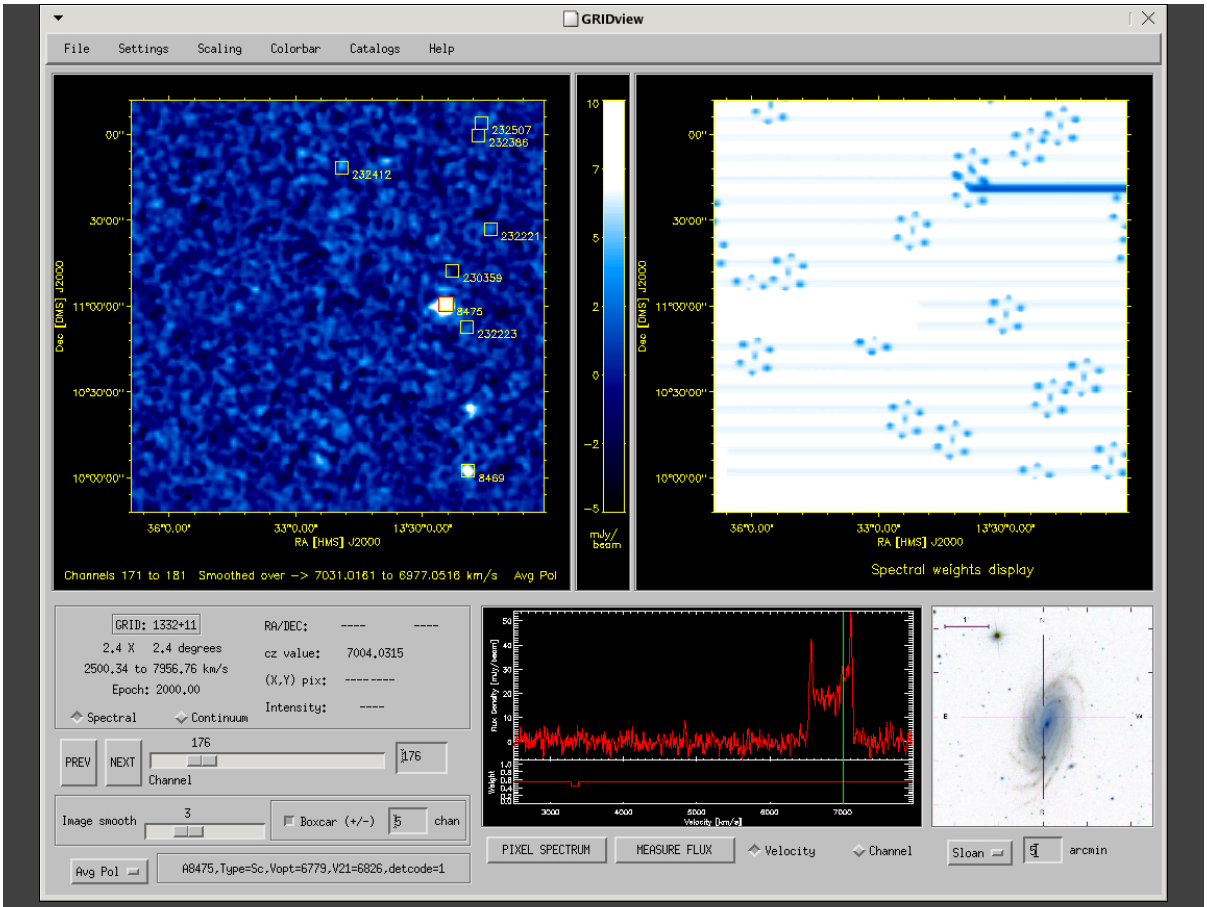
Learn about current SPN members - their current organizational roles, project activities and SPN working group membership.

MEET Daina Bouquin

Head Librarian, Harvard-Smithsonian Center for Astrophysics (CfA)

MS Data Science, MS Library and Information Science

CfA scientists pursue research about the physical processes that determine the nature and evolution of the universe. As Head Librarian I hold appointments with both the Harvard College Observatory and the Smithsonian Astrophysical Observatory and am a member of the CfA's Scientific Computation Advisory Committee. One of my current priorities is finding ways for the library to support research software preservation, access, and citation.



Example of the IDL environment - GRIDView by Dr. Brian Kent, NRAO

Software and Data from Multiple Sources

In astronomy, research software is foundational to both the future and heritage of the science. Deep intellectual contributions are made by people building software that cannot be decoupled from data and often relies on distributed computing resources and data sources. For example, new telescopes like LSST will generate 15 terabytes of data a day and the research that will be done with that staggering dataset will be driven by distributed computing and code. Even in the past, modern astronomy generally relies on data from large surveys and astronomers have had to write their own tools to do their work. It is essential then that practices are adopted to enable people to share complex, distributed, changing tools and receive credit for their work.

Example of image reliant on multiple sources of software and data: X-ray: NASA/CXC/SAO/J.DePasquale; IR: NASA/JPL-Caltech; Optical: NASA/STScI

MEET Neil Chue Hong

Director of the Software Sustainability Institute

Twitter: @softwaresaved, @npch

I am the Director of the Software Sustainability Institute (https://www.software.ac.uk/), a collaboration of UK universities which develops and promotes best practice around research software. Originally trained as a computational physicist at the University of Edinburgh, I joined a technology transfer centre before moving on to

the development, management and support of large research software projects. In my current role, I am active in software sustainability – the ability of researchers to use software for its intended function over time – and I am a part of the Software Preservation Network, co-Chair of the FORCE11 Software Citation Implementation Working Group and RDA Software Source Code Interest Group, and Editor-in-Chief of the Journal of Open Research Software.

Published Draft: Guidelines on Research Software Deposit

Last year, the SSI was approached by Jisc to develop further guidance for researchers on software deposit and preservation, as part of their Research Data Shared Service initiative. The SSI team, led by Mike Jackson, ran workshops encompassing researchers, research software engineers, librarians, research data managers, infrastructure providers, and publishers to define the key topics and gaps where guidance on software deposit was needed and collect pointers to existing materials.

REQUEST FOR FEEDBACK: WE WANT TO HEAR FROM YOU

A final workshop provided feedback on the draft guidance, which has now been published in a first public version **(**GitHub: https://github.com/softwaresaved/software-deposit-guidance**)**. We're interested in getting wider feedback to inform future versions of the guidance, as well as encouraging people to use the guidance (licensed under a Creative Commons Attribution CC BY license) and letting us know which parts have been most useful. There are several topics, in particular, the SSI team would appreciate your feedback on - please review the guidelines and complete the (BRIEF) Feedback Form for the Guidelines on Research Software Deposit.

More Information About JISC and SSI Contributions to Software Preservation

The SSI previously worked on a Jisc-funded project in collaboration with Curtis and Cartwright to understand the

benefits and challenges of software preservation (https://www.software.ac.uk/resources/guides/sustainability-and-preservation-framework).

AFFILIATED PROJECTS

Learn about SPN affiliated project activities and milestones. SPN affiliated projects focus on some aspect of software preservation/curation that supports the strategic goals of SPN.

Timeline: January 2017 – June 2020

Funder(s): Alfred P. Sloan Foundation

Awardee: Association of Research Libraries

The Code is currently in the process of external review by fair use experts. Any feedback will be addressed in a final round of research team edits and submitted for print. The Code is still scheduled for Fall 2018 publication, accompanied by supplementary resources that address licensing and digital rights management.

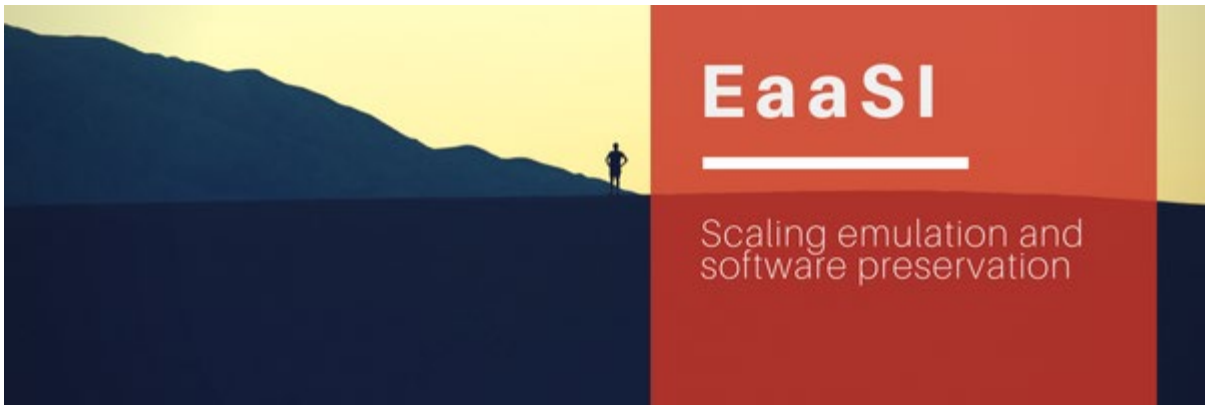
To learn more about the Best Practices methodology and the perspectives of the research team, watch these interviews with Peter Jaszi (American University), Patricia Aufderheide (American University) and Brandon Butler (University of Virginia). They were recorded in 2014 during the first ARL Code for Best Practices in Fair Use.

"Code of Best Practices Interview: Peter Jaszi and Patricia Aufderheide"

American University professors Peter Jaszi and Patricia Aufderheide discuss the development of the Code of Best Practices in Fair Use for Academic and Research Libraries. (11 minutes)

"Code of Best Practices Interview: Peter Jaszi and Brandon Butler"

American University law professors Peter Jaszi and Brandon Butler discuss the development and roll-out of the Code of Best Practices in Fair Use for Academic and Research Libraries. (16 minutes)



Timeline: January 2018 – June 2020

Funder(s): Andrew W. Mellon Foundation and Alfred P. Sloan Foundation

Awardee: Yale University

Guest Post by: Seth Anderson

July marks the official kick-off of the EaaSI Network! Please welcome the Founding EaaSI Node Hosts

- Carnegie Mellon University (Eric Kaltman, Node Lead)
- Notre Dame University (Don Brower, Node Lead)
- Stanford University (Michael Olson, Node Lead)
- University of California - San Diego (Sibyl Schafer, Node Lead)
- University of Virginia (Robert German, Node Lead)

During the Pre-Implementation & Testing Phase (July 2018 - December 2018), Node Hosts will brainstorm their organizational use cases, complete a pilot software & collections inventory, explore scenarios for use & feedback in more depth, rate their software curation readiness and articulate changes to internal policies that address the unique characteristics of software. In the September/October 2018 issue of STACKTRACE - all the members from each Node Team will be introduced and you will be invited to read their *Statements of Importance for Software Preservation and Emulation*.

OpenSLX began development of authentication and security features for the EaaSI service. Developers implemented OAuth2 procedures for login to the demo UI via public authentication services from Twitter and Google. Developers have also taken steps to improve the security of the system's REST API interfaces. Future work will focus on the security of the system's SOAP and HTTP protocols and integration of Shibboleth authentication procedures. The project team is also working on a metadata profile for the EaaSI system and its contents. EaaSI's metadata profile will identify and define properties for software, digital collection materials, and emulated computer environments. The team is reviewing and prioritizing potential properties to capture, based on their value for discovery, system functionality and automation, and provenance tracking. A draft of the profile will be distributed next month.

Governance and sustainability planning for EaaSI has begun! Working closely with Euan Cochrane (Principal

Investigator), Seth Anderson (Program Manager) and Jessica Meyerson (Community & Outreach Lead) over the next year, Katherine Skinner (Educopia Executive Director) will guide EaaSI staff through a process that includes landscape scanning, semi-structured interviews, and surveys. Outcomes include a robust market analysis and recommendations for EaaSI hosting and governance structures.

Timeline: June 2017 – May 2020

Funder(s): Institute for Museum and Library Services #RE-95-17-0058-17

Awardee: CalPoly State University

Guest Post by: Zach Vowell

Over two days in early August, the members of the Fostering a Community of Practice (FCoP) cohort assembled at the Computer History Museum in Mountain View, CA. The 2-day In-Person Kick-Off Meeting was an opportunity for cohort members to examine their software preservation projects through the lens of metadata, existing digital curation workflows, internal/external advocacy and project management.

Day 1 began with Research Lightning Talks in which each cohort members shared findings from preliminary research to the group followed by Q&A. Presentations were followed by collaborative work. Cohort members interviewed one another with questions addressing the impact of programmatic software preservation and emulation on their current digital curation tasks and workflows, as well as existing descriptive practices. Throughout Day 1, the cohort was prompted to think about how their work might relate to external organizations and projects, and how different kinds of outreach might be leveraged to connect to external stakeholders (and internal stakeholders).

I personally had the honor of pulling some themes out of Day 1 to share at the start of Day Two including:

- Impact of “upstream” work (typically carried out by people other than the cohort members) on software preservation efforts accompanied by a goal of testing internal advocacy strategies to effect their “upstream” stakeholders;
- Consensus that bit-level preservation is not enough - what does quality control for software-dependent

files look like? Is there a checklist we can develop from our collective experience?

- Software preservation and emulation still exist on the fringes of descriptive schemas even though a growing number of organizations want information about software dependencies and emulation environments to be absorbed into their workflows, repositories/catalogs, and preservation systems.

Day 2 ended on a light note: a full backstage tour of the Computer History Museum's research (and collections storage) facility in Fremont. There we saw Whirlwind computer tape, Xerox Alto disk cartridges (or "packs"), a handy card catalog from the Digital Equipment Corporation (DEC) collection, and of course boxes and boxes of software floppy disks, and the museum's digital forensics lab where the disks are imaged. It was a great reminder of all the work that goes into preserving software, and how we all need each other to preserve it.

The work completed by all of the FCoP Project Leads, including Wendy Hagenmaier (Georgia Tech) that joined us via Zoom, resulted in clear next steps for individual software preservation projects as well as the cohorts' collaborative work. Currently, FCoP cohort members are refining their detailed project plans and completing a software & collections inventory. Huge thanks to our CHM Hosts: Paula Jabloner, Andrew Berger, Elena Colon-Marrero, Hansen Hsu, and Caroline Evans.

NETWORK MAINTENANCE

Learn about changes and improvements to functional aspects of the Network. This section will feature updates on governance, funding, engagement strategies and other activities that comprise backbone support for working groups, affiliated projects and strategic partnerships.

This September, we'll be sharing a series of software preservation stories that highlight the growth and evolution of the Software Preservation Network. Ranging from the personal to the institutional, from our accomplishments of the past three years, to our goals for the future, these stories speak to the collective action that drives our organization and our efforts - to save software, together.

To follow along with the series, find us on Twitter and keep an eye out for a special edition STACKTRACE in October!

FIND OUT MORE

The Software Preservation Network (SPN) facilitates and supports software preservation efforts. SPN preserves software through community engagement, infrastructure support and knowledge generation in five core activity areas including Legal & Policy Advocacy, Metadata & Standards Development, Training & Education, Research-in-Practice and Technological Infrastructure.

Software Preservation Network

Twitter: @softpresnetwork
Web: www.softwarepreservationnetwork.org
Email: contact@softwarepreservationnetwork.org
Subscribe: https://groups.google.com/d/forum/software-preservation-network