Working towards Sustainable Software for Science: Practice and Experience (WSSSPE)

Daniel S. Katz
Associate Director for Scientific Software & Applications, NCSA
Research Associate Professor, ECE
Research Associate Professor, iSchool
dskatz@illinois.edu, d.katz@ieee.org, @danielskatz
WSSSPE introduction

- Progress in scientific research depends on the quality and accessibility of software at all levels
- Must address challenges
  - Related to development, deployment, maintenance, and overall sustainability of reusable software
  - And education around software practices
- Challenges
  - Can be technological policy based, organizational, and educational
  - Of interest to developers (the software community), users (science disciplines), software-engineering researchers, and researchers studying the conduct of science (science of team science, science of organizations, science of science and innovation policy, and social science communities)
WSSSPE history

• WSSSPE1 (at SC13): broad scientific community identified challenges and best practices in areas of interest to creating sustainable scientific software
  • Topics picked by organizer team:
    • Developing and Supporting Software, Policy, Communities
• WSSSPE2 (at SC14): community proposed and discussed specific mechanisms to move towards an imagined future for software development and usage in science and engineering
  • But no good way to enact those mechanisms, or to encourage the attendees to follow through on their intentions
  • Topics based on collaborative grouping of submitted papers:
• Also, WSSSPE1.1 & WSSSPE2.1 at SciPy13&14
  • Aimed at discussing WSSSPE issues in SciPy context
WSSSPE3

- Boulder, Colorado, September 2015
- Included multiple mechanisms for participation
  - Keynote & lightning talks but mostly breakout discussions
- Encouraged team building around solutions via working groups
- Included professional event organizers/facilitators
  - Helped design workshop, kept things on time, encouraged discussion, formation of working groups, outputs from working groups
- Topics: began with topics from WSSSPE2
  - CFP requested additions
  - Attendees chose which topics they wanted to work on
WSSSPE3 potential activities (1/2)

- Development and Community
  - Writing a white paper/review paper about best practices in developing sustainable software
  - Documenting successful models for funding specialist expertise in software collaborations
  - Creating and curating catalogs for software tools that aid sustainability (perhaps categorized by domain, programming languages, architectures, and/or functions, e.g., for code testing, documentation)
  - Documenting case studies for academia/industry interaction
  - Determining effective strategies for refactoring/improving legacy scientific software
  - Determining principles for engineering design for sustainable software
  - Create a set of guidance giving examples of specific metrics for the success of scientific software in use, why they are chosen, what they are useful to measure, and any challenges/pitfalls; then publish this as a white paper

- Training
  - Writing a white paper on training for developing sustainable software, and coordinating multiple ongoing training-oriented projects
  - Developing curriculum for software sustainability, and ideas about where such curriculum would be presented, such as a summer training institute

*Bold = pursued during the workshop*
WSSSPE3 potential activities (2/2)

- **Credit**
  - Hacking the credit and citation ecosystem (making it work, or work better, for software)
  - Developing a taxonomy of contributorship / guidelines for including software contributions in tenure review
  - Documenting case studies of receiving credit for software contributions
  - Developing a system of awards and recognitions to encourage sustainable software

- **Publishing**
  - Developing a categorization of journals that publish software papers (building on existing work), and case studies of alternative publishing mechanisms that have been shown to improve software discoverability/reuse, e.g., popular blogs/websites
  - Determining what journals that publish software paper should provide to their reviewers (e.g., guidelines, mechanisms, metadata standards)

- **Reproducibility and Testing**
  - Building a toolkit that could allow conference organizers to easily add a reproducibility track
  - Documenting best practices for code testing and code review

- **Documentation**
  - Develop landing pages on the WSSSPE website (or elsewhere) that enable the community to easily find up-to-date information on a WSSSPE topic (e.g., software credit, scientific software metrics, testing scientific software)

**Bold = pursued during the workshop**
WSSSPE3 working groups

• Active in the meeting
  • White paper/journal paper about best practices in developing sustainable software
  • Funding Research Programmer Expertise
  • Transition Pathways to Sustainable Software: Industry & Academic Collaboration
  • Principles for Software Engineering Design for Sustainable Software
  • Useful Metrics for Scientific Software
  • Training
  • Software Credit Working Group
  • Publishing Software Working Group Discussion
  • Building Sustainable User Communities for Scientific Software

• Started then dissolved
  • Legacy Software

• Each was asked to discuss the topic, create plans to move forward, assign a POC, create a landing page
WSSSPE3 working groups that overlap other groups

- **Funding Research Programmer Expertise**
  - Overlaps UK Research Software Engineer (RSE) group, but is intended to be international
  - Lead: James Hetherington
  - Not much progress outside of UK, but UK activity is strong
  - RSE Conference coming up in September in Manchester
  - [http://www.rse.ac.uk/index.html](http://www.rse.ac.uk/index.html)

- **Software Credit Working Group**
  - Lead: Kyle Niemeyer
  - Merged with Force11 Software Citation Working Group
  - Delivered Software Citation Principles
    - [https://www.force11.org/software-citation-principles](https://www.force11.org/software-citation-principles)
    - To be published in ~ next month
WSSSPE3 working groups with partial success (or work in progress)

• White paper/journal article about best practices in developing sustainable software
  • Leads: Sandra Gesing, Abani Patra
  • Long delay after WSSSPE3, but now moving again
  • In progress in https://github.com/WSSSPE/WG-Best-Practices

• Useful Metrics for Scientific Software
  • Lead: Gabrielle Allen
  • Survey of NSF SI2 projects & their metrics in progress
  • To be used in white paper

• Publishing Software Working Group
  • Lead: Steve Brandt
  • Created web page of venues that publish executable papers
  • To be posted on SSI site
  • SSI already has list of journals that publish software papers
WSSSPE3 working groups with no progress

- Transition Pathways to Sustainable Software: Industry & Academic Collaboration
  - Led by Nic Weber
  - No specific plans made
- Principles for Software Engineering Design for Sustainable Software
  - Led by Colin Venters and Birgit Penzenstadler
  - Some ideas proposed as plans, but no progress reported
- Training
  - Led by Nick Jones and Erin Robinson
  - Initial plans made, but no progress reported
- User Community Working Group
  - Led by Dan Gunter and Ethan Davis
  - Some plans created, but group members didn’t have time (or funds) available to pursue them
WSSSPE3 lessons

• What worked:
  • Highly motivated leads, where the working group strongly matched their interests
  • Overlap with existing effort that could be merged
• What could have worked better
  • More prodding from a coordinator to encourage leads to follow their plans
  • More prodding from leads to group members
• What didn’t
  • Groups that didn’t come up with specific plans
  • Groups that came up with specific plans, but didn’t carry them out (lack of time/funds)
• In all cases, where interest wasn’t strong enough, funding would have helped, but maybe wouldn’t have been sufficient
WSSSPE4 plans

- 2 ½ day event, Sept 12-14, Manchester, UK
- Preceding RSE Conference (also in Manchester)
- Much will be split:
  - **Track 1 – Building a sustainable future for open-use research software**
    - Goals: define vision of future of open-use research software, and in workshop, initiate activities needed to get there
    - Think about where we want to be 5 to 10 years from now, without being too concerned with where we are today, and then to determine how we can move to this future.
  - **Track 2 – Practices & experiences in sustainable scientific software**
    - Goal: improve the quality of today’s research software and the experiences of its developers by sharing practices and experiences
    - Current state of scientific software & what we can do to improve it in the short term, starting with where we are today
Links and references

- WSSSPE: http://wssspe.researchcomputing.org.uk
  - With semi-active mailing list
  - And inactive wiki
- WSSSPE1 report: http://dx.doi.org/10.5334/jors.an
- WSSSPE2 report: http://dx.doi.org/10.5334/jors.85
- @wssspe on twitter
- WSSSPE on Facebook