# Planning to Make Research Software More Sustainable via a US Research Software Sustainability Institute (URSSI)

Berkeley — UNIVERSITY OF CALIFORNIA

UNIVERSITY OF NOTRE DAME

## http://urssi.us

**Daniel S. Katz**, Jeffrey C. Carver, Sandra Gesing, Karthik Ram, Nicholas Weber
Assistant Director for Scientific Software & Applications, NCSA
Research Associate Professor, CS, ECE, iSchool

dskatz@illinois.edu, d.katz@ieee.org, @danielskatz

ILLINOIS
NCSA | National Center for
Supercomputing Applications

# Research Software

Software developed and used for the purpose of research: to generate, process, analyze results within the scholarly process

ILLINOIS NCSA

Research Software

|  | UK | US |
|---|---|---|
| Use | 90% | 95% |
| Can't continue without | 70% | 63% |

UK: https://doi.org/10.5281/zenodo.14809
US: https://doi.org/10.6084/m9.figshare.5328442.v3

# Software Sustainability

- The capacity of the software to endure

- The software will continue to be available in the future, on new platforms, meeting new needs

- The practices, both technical and non-technical, that allow software to continue to operate as expected in the future

I ILLINOIS NCSA

# NSF software funding

- NSF made 18,592 awards totaling $9.6 billion (1995 – 2016) that topically reference "software" in their abstracts



Funding Profile For Software
(Rectangular Area: Number of Awards , Color: Award Amount)

ILLINOIS NCSA
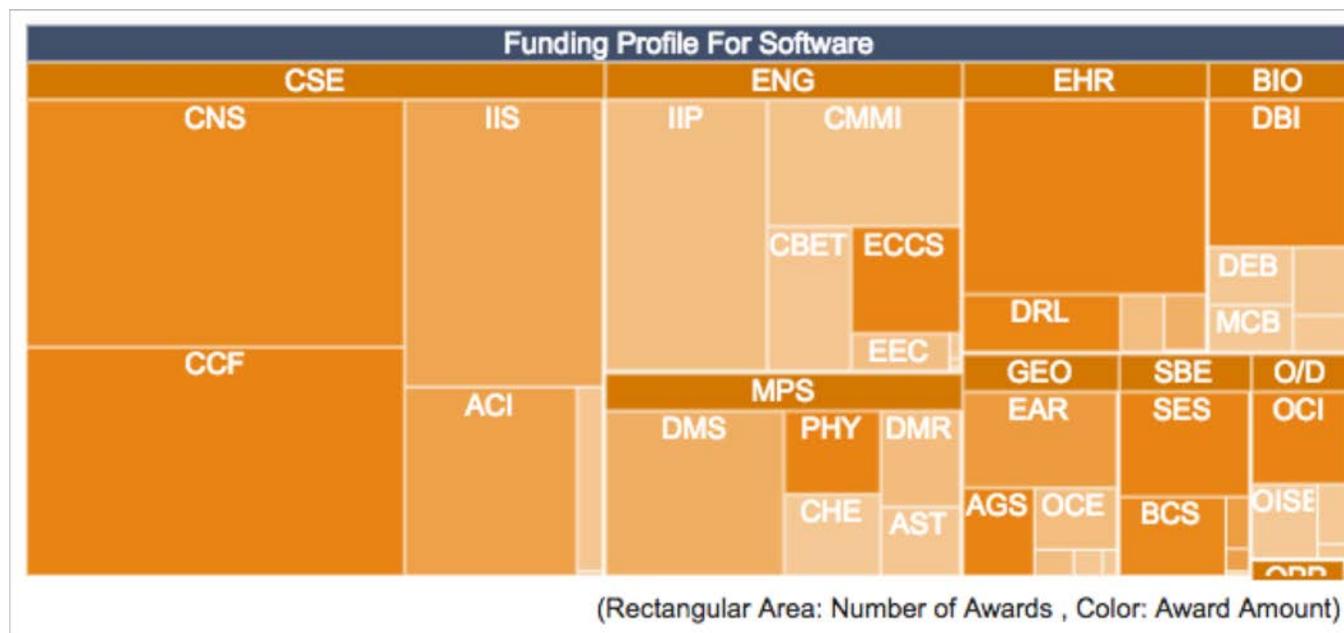
# URSSI context

- Modern research is digital, products created, analyzed, and stored electronically, using software

- Much software developed specifically for research, by researchers

- Research software essential to progress in almost all research fields
  - But often not developed in an efficient or sustainable way

- Researchers know their disciplines
  - Often don't have training and understanding of best practices to ease development & maintainability and to encourage sustainability & reproducibility

- Developers don't match the diversity of overall society or of user communities

# URSSI conceptualization goals

- Conceptualize (plan) a US Research Software Sustainability Institute

- Cut across existing activities funded by NSF and beyond

- Directly and indirectly positively impact all software development and maintenance projects

- Focus on the entire research software ecosystem, including the people who create, maintain, and use research software

- Outputs:
  - Eager supportive & inclusive community
  - Concrete institute plan configured to offer valued services
  - Published survey and data that demonstrates community need

# URSSI team

## PIs

- Karthik Ram (PI, UC Berkeley)
- Jeffrey Carver (Alabama)
- Sandra Gesing (Notre Dame)
- Daniel S. Katz (Illinois)
- Nicholas Weber (Washington)

## Advisory Committee

- Richard Arthur (GE Global)
- Michelle Barker (ARDC)
- Phil Bourne (Virginia)
- Daniel Crawford (MolSSI & Va Tech)
- Neil Chue Hong (SSI & Edinburgh)
- James Howison (Texas)
- Kurt Schwehr (Google)
- Jeff Spies (SHARE)
- Nancy Wilkins-Diehr (SGCI & UCSD)

## Senior Personnel

- Wolfgang Bangerth (Colorado State)
- Anshu Dubey (Argonne)
- Melissa Haendel (Oregon State)
- Mike Heroux (Sandia)
- Katy Huff (Illinois)
- Suresh Marru (Indiana)
- Kate Mueller (Notre Dame)
- Jarek Nabrzyski (Notre Dame)
- Kyle Niemeyer (Oregon State)
- Marlon Pierce (Indiana)
- Ariel Rokem (Washington)
- Arfon Smith (STScI)
- Tracy Teal (Carpentries)
- Matt Turk (Illinois)
- Rick Wagner (Argonne)
- Mike Zentner (Purdue)

# NSF Institutes Context
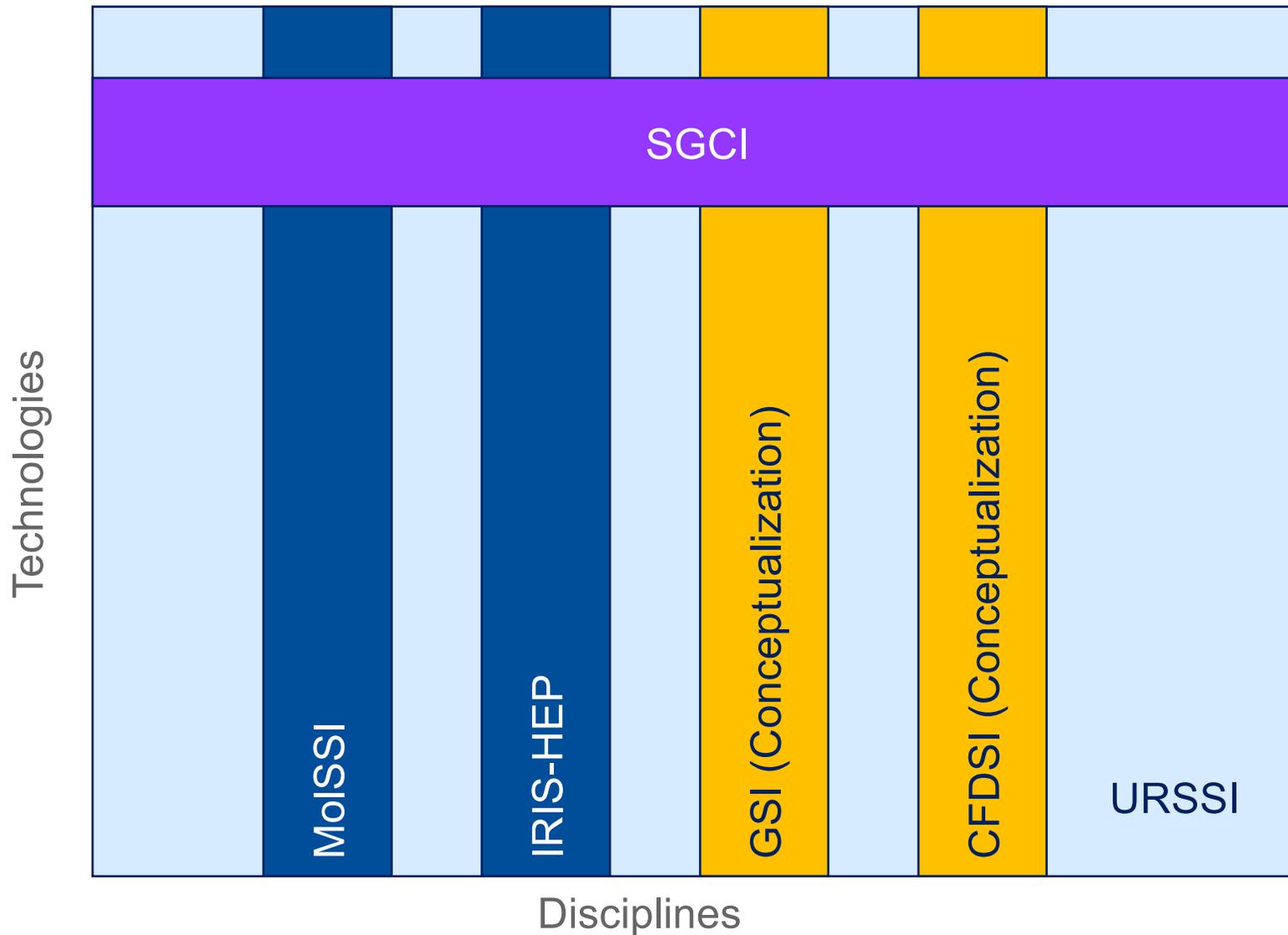
# URSSI activities

- Workshops

- Survey

- Ethnographic studies

- Communication, including blogs (urssi.us/blog)

- Through all activities, iteratively build on existing, extensive understanding of the challenges for sustainable software and its developers

# Workshop 1 – Berkeley

- General workshop, lightning talks & unconference

- Four key themes emerged from unconference discussions
  - Issues unique to software development, usability, and credit
  - Training and workforce development, including potential for building upon existing models, and designing new methods of instruction
  - The sustainability of research software engineers (RSEs) as the human infrastructure of scientific software
  - The sustainability of software projects through organizational and institutional best practices

- Initial strawhorse presented & discussion
  - How might URSSI be set up, and what will it do (see next slide)
  - Budget exercise (divide $5m/year …)

- Blog post: http://urssi.us/blog/2018/08/23/report-from-the-first-urssi-workshop/

# URSSI Strawhorse from Workshop 1

| | Supporting Software | Supporting People | Supporting the community | Science & research impact |
|---|---|---|---|---|
| Development Support (consulting & short term small project support) | X | | | X |
| Incubator (technology advice, business planning, usability advice, etc.) | X | | | X |
| Training (courses & guides) | X | X | | X |
| Policy (research & campaigns) | | X | X | X |
| Community (fellowships, workshops, blogs, website) | X | X | X | X |

ILLINOIS NCSA

# Workshop 2 – Chicago

- General workshop, lightning talks & unconference
- Four key themes emerged from unconference discussions
  - Community building
  - Career path and institutional support for RSEs
  - Training and workforce development
  - Sustainability in relation to reproducibility, usability, discoverability
- Discussion around URSSI summer school
  - Spun out of training discussion
  - Aimed at grad students, who would come back as TAs/mentors later
- Discussion around URSSI mission and vision
  - Ideas fed into ongoing process of small team drafting M&V
- Discussion around initial strawhorse
  - Realized Management was missing, including sustainability & governance
  - Development still unclear, scaling is problematic
  - Good ideas for other elements
- Report being finalized

# Workshop 3 – Santa Barbara

- Focused workshop on Software Credit, Metrics
- Lots of short talks about projects in this space
- Lots of discussion
- Then used Rumelt's Strategy Kernel concept
  - Diagnosis (problem)
  - Guiding principles
  - Coherent actions
- to plan potential URSSI work in this area
- Mostly in policy area
- With some work done by fellows (community area)
- Report being finalized

ILLINOIS NCSA

# Workshop 4 – College Park

- Focused workshop on Software Incubation
- Participants from existing software and leadership incubation organizations
  - Apache Incubator, Code for Science & Society, eScience Institute Winter Incubators, ESIP Incubator, Mozilla Open Leaders, NumFOCUS
- along with researchers who study software ecosystems
- Goals
  - Envision how URSSI could support new and existing software projects that face challenges in attracting contributors and additional maintainers, setting up a governance model, seeking out or managing funding sources, and generally development practices in producing reusable research software.
  - Gather information and recommendations from experts
  - Compare existing models of support and guidance for research software projects (or at least projects that depend on developing new software)
  - Provide inputs for final project meeting
- Model of project phases determined, and existing efforts mapped
- URSSI could coordinate and fill gaps – focus on a specific transition
- Report being developed

**I ILLINOIS** NCSA

# Stages of open community for research software



**Stage 0.** Some code and a user of it. No sustained team.

**Stage 1.** Software development team, internal use.

**Stage 2.** Multiple software teams (different institutions) on same code (team is *community*), for internal use.

**Stage 3.** Self-governing developer community deliberately supporting broad user community.

**Stage 4.** Self-sustaining organization dedicated to supporting user and dev community (e.g. through commercial support, events, software foundation, etc.).

- Direct commercialization, and open source tool development from the start are additional paths with overlapping stages

**ILLINOIS** NCSA

# Survey

- Aims – to obtain general feedback from a large community, and to let them know about URSSI

- Feedback on issues of developers and users

- ~1200 total responses, ~1000 completed most of the survey

- Results being analyzed; some demographics in the meantime

# Survey

- Aims – to obtain general feedback from a large community, and to let them know about URSSI

- Feedback on issues of developers and users
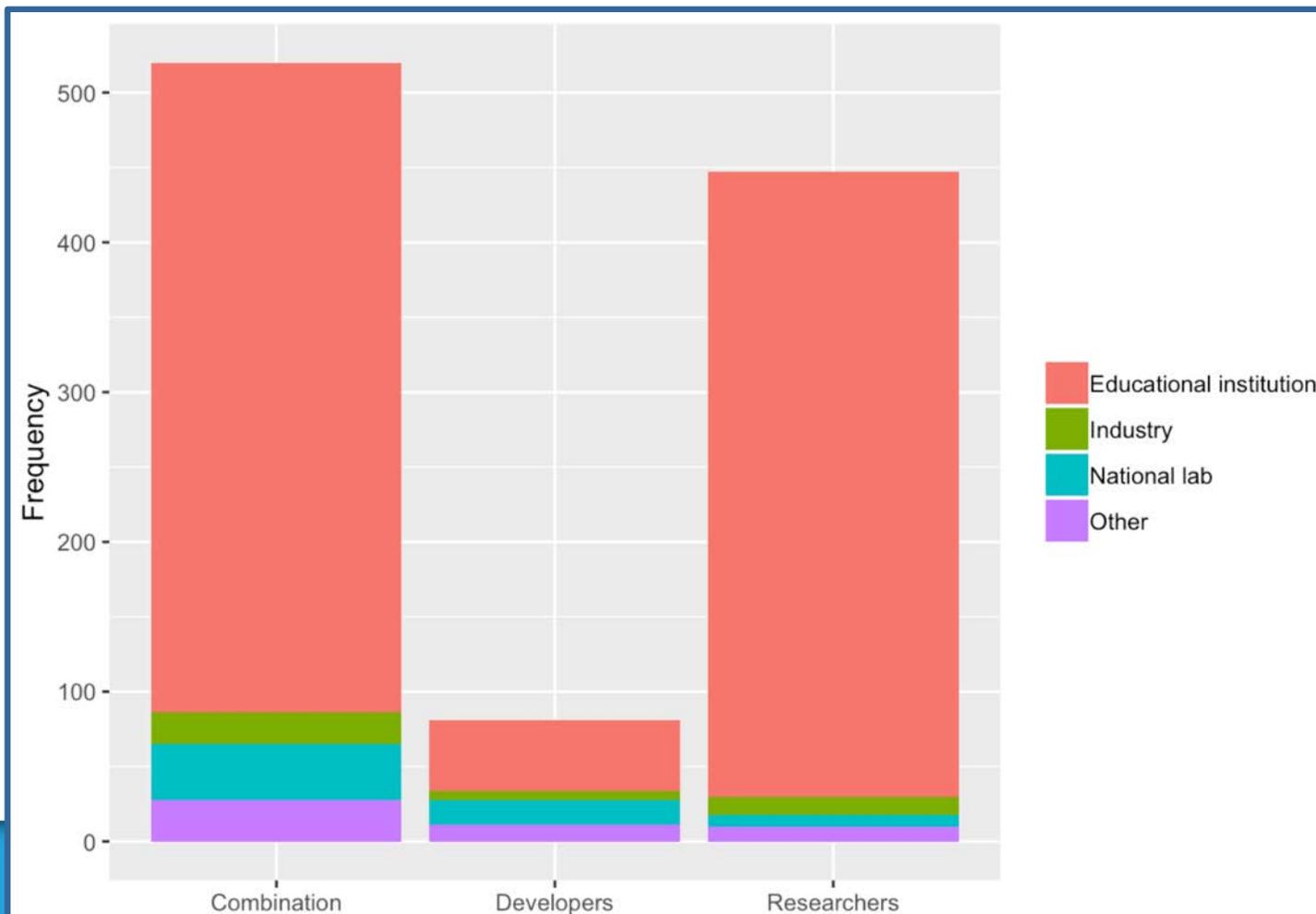
he survey

meantime

# Survey

- Aims – to obtain general feedback from a large community, and to let them know about URSSI

- Feedback on issues of developers and users
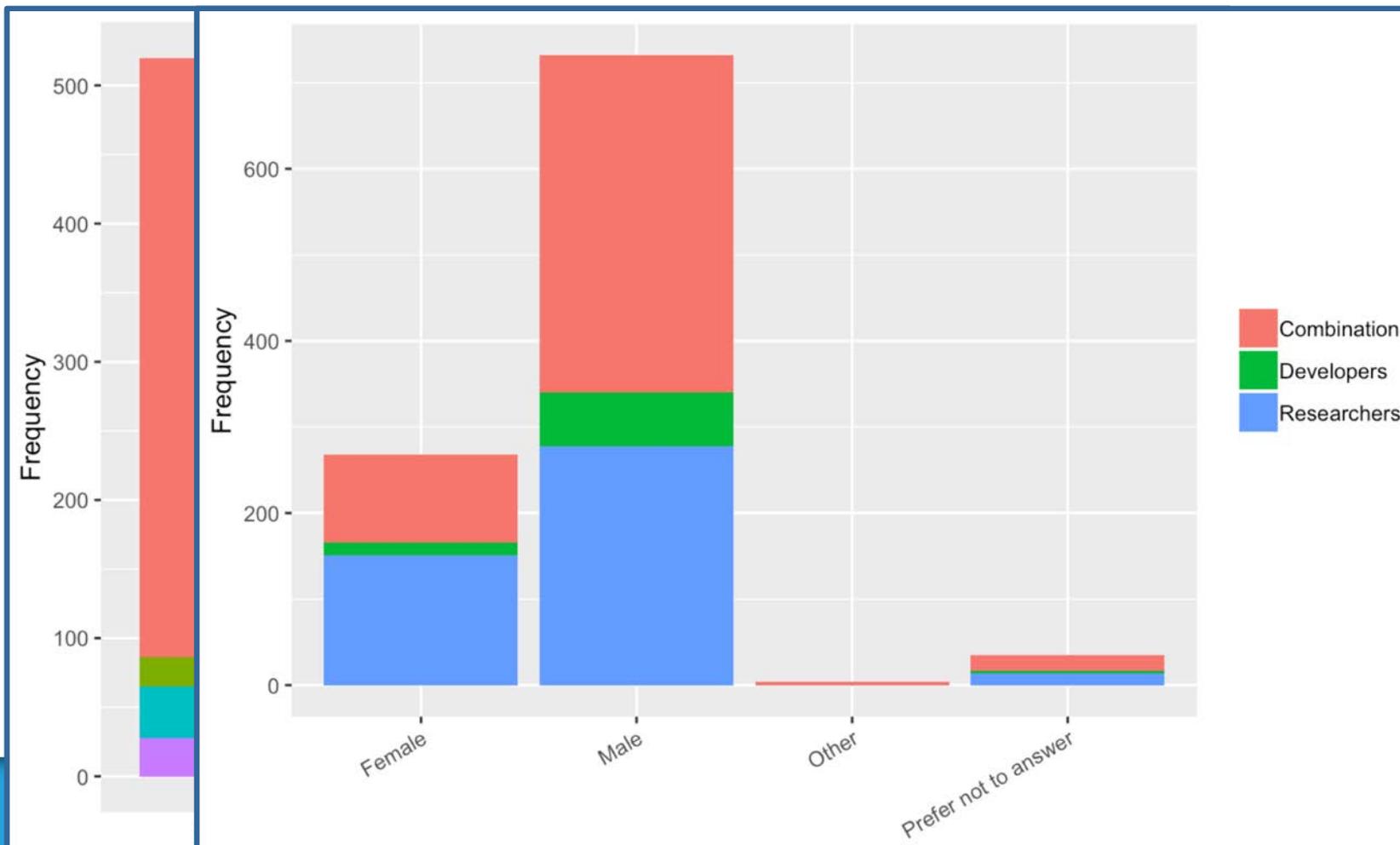
# Survey

- Aims – to obtain general feedback from a large community, and to let them know about URSSI
- Feedback on issues of developers and users
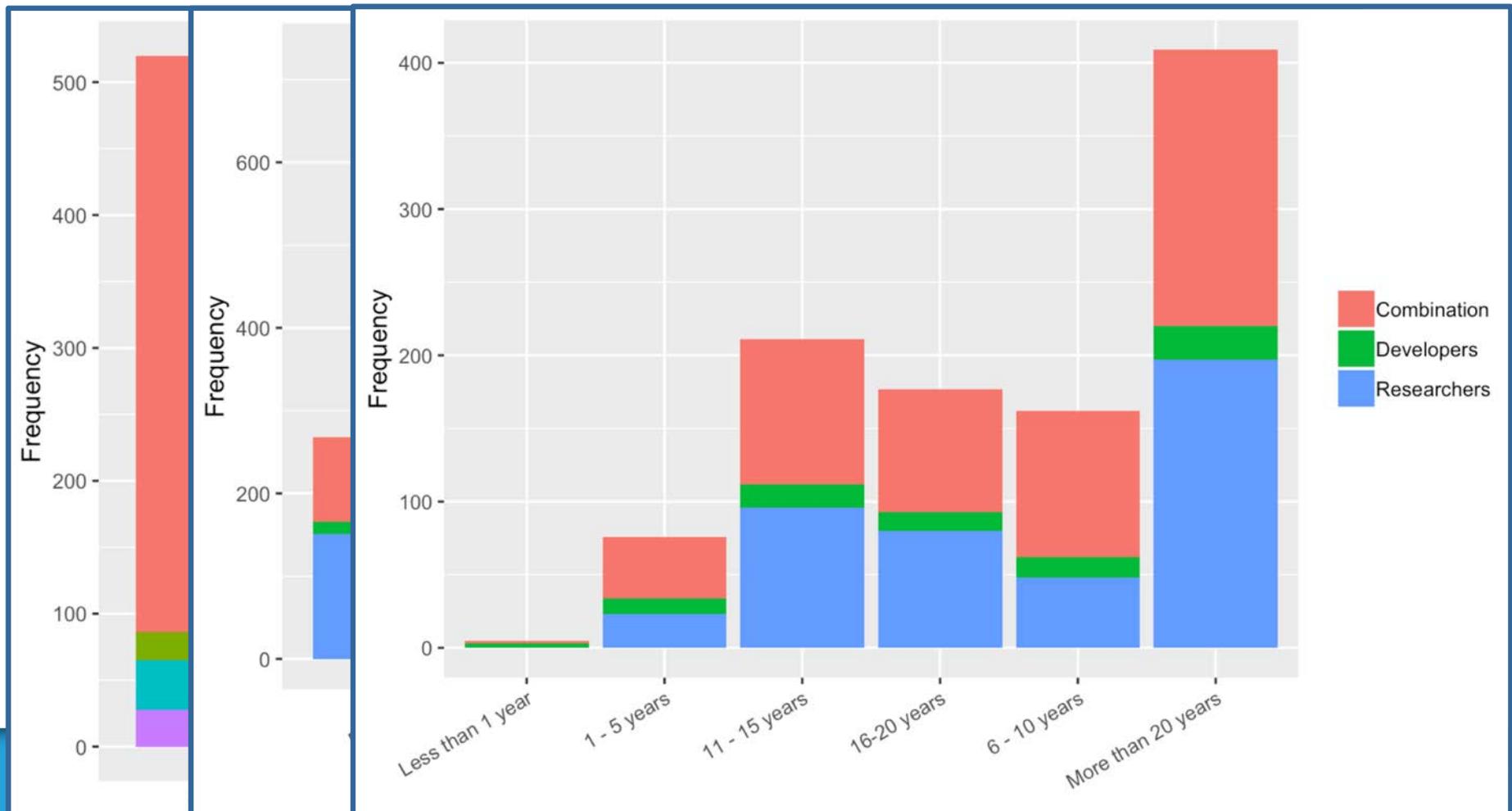
# Ethnographic studies

- Led by Nic Weber
- Over the last 11 months
- Conducted observations & semi-structured interviews
- With three projects of different types
- Gathered a series of archival documents to produce case studies of how, over time, research software projects overcome challenges in
  - Recruiting contributors
  - Building a governance model
  - Seeking funding
  - Sharing credit
- Analysis and reporting is ongoing
- Will be completed for use in final workshop

# Final meeting

- April in Chicago

- PIs, Senior Personnel, Advisory committee members, a few selected experts

- Goals
  - Outline URSSI implementation plan
  - Assign authors to sections
  - Start writing

- Will lead to final implementation in plan in May
  - To be made public
  - And submitted to NSF as part of project final report

- Then we wait to see what NSF wants to do
  - Contact your NSF program officer to provide advice
    - Such as: this is important, needed ☺

# How to get involved in URSSI

- Watch the web page: http://urssi.us

- Join the mailing list – via form on the URSSI web page

- Suggest and discuss topics: http://discuss.urssi.us

- Follow on twitter: https://twitter.com/si2urssi

- Write a blog, or suggest one we should cross-post

- Use GitHub: http://github.com/si2-urssi/
  - Repos for the web site (PRs to add) and workshops

- Talk to others about URSSI (including NSF)

- If you have questions, want to suggest something, want to volunteer, email us: contact@urssi.us

ILLINOIS
NCSA | National Center for
Supercomputing Applications