yet, there is a clear lag in the capabilities of many in the research community to exploit these new opportunities, due to lacking the relevant technical skill-sets to effectively use Big Data.

### Distributed Software Training

Community-based feedback catalyzed creation of a partnership between the West Coast-based ISEES, and East Coast-based WSSI, to jointly host an “Open Science for Synthesis” Institute, OSS—a three-week, bi-coastal training program from July 21- August 8, 2014. Through a competitive process, 45 early career researchers, representing a broad geographic and disciplinary mix, attended intensive sessions involving approximately half a day in lectures and hands-on experiences, and half a day working on Group Synthesis Projects. Web-based, videoteleconferencing and IRC facilitated real-time interactions among all participants, who were located at NCEAS in California and Renci in North Carolina.

OSS topics included:
- Linux/UNIX command line and shell concepts;
- data creation, discovery, and preservation (SQL, DataONE, Dryad);
- code creation (Python, R), management, sharing, and versioning (Git);
- scientific metadata and workflow documentation;
- statistical and machine modeling techniques;
- virtual collaboration mechanisms (IRC, VTC, Wiki);
- methods for communicating scientific results (COMPASS MessageBox)

All technologies and quantitative tools presented were suitable for advancing open, collaborative, and reproducible synthesis research.

### Learning by Doing

Participants formed small synthesis teams that utilized the software skills learned each day in the context of cross-cutting science research projects. Using an open community engagement process, participants can maximize their success in collaborative research that could potentially lead to publishable results:


### Formal Assessment

All participants filled out surveys before, mid-way, and after the OSS event, to help formally assess which aspects of the curriculum were most useful, and inform future events that might run distributed, intensive, hybrid classroom/working group events such as this one.

Analyses are currently underway, under the direction of Dr. Rebich-Hespanha. Assessment results will inform our future efforts, and be published in a relevant CSCW journal for broader dissemination.

### Sample Testimonials

- I speak for the group when I say that [the OSS] was a great opportunity; it will definitely help us advance our careers in the near-term...

- I really value the knowledge and experience I gained and can’t put into words how useful it is going to be for me.

- I just wanted to let you know that our small group project from the Summer Institute has actually turned into a funded working group...

- While I felt great about the workshop when I was there, I feel even better now, using the skills I picked up...

- Thank you again for such an amazing training. I learned a lot and it really transformed how I think of open science and working across groups...

- It is my first day back in the office and I already found myself using a number of the skills I learned...

- What we learned really ought to be taught to every incoming PhD student in an ecology program, but it just isn’t...

### Training: an integral part of the Software Life Cycle

Institute for Sustainable Earth and Environmental Software, ISEES (NSF ACI-1216894)
http://isees.nceas.ucsb.edu

Water Science Software Institute, WSSI (NSF ACI-1216817)
http://renci.org/research/water-science-software-institute/

http://www.nceas.ucsb.edu/oss
(a joint ISEES/WSSI training event)

Contact information
Mark Schildhauer, PhD
Director of Computing
National Center for Ecological Analysis and Synthesis, UCSB
schildhauer@nceas.ucsb.edu

**S2I2-- Community capacity building for sustainable software ecosystems: Experiences running a real-time, bi-coastal**

*Mark Schildhauer1, Matthew Jones1, Ray Idaszak2, Chris Lenhardt2, Stan Ahalt2, Benjamin Bolker3, Laura Christopherson2, Stephanie Hampton4, Stacy Rebich-Hespanha1*

1National Center for Ecological Analysis and Synthesis, UC Santa Barbara; 2Renaissance Computing Institute, UNC Chapel Hill; 3McMaster Univ; 4Washington State Univ