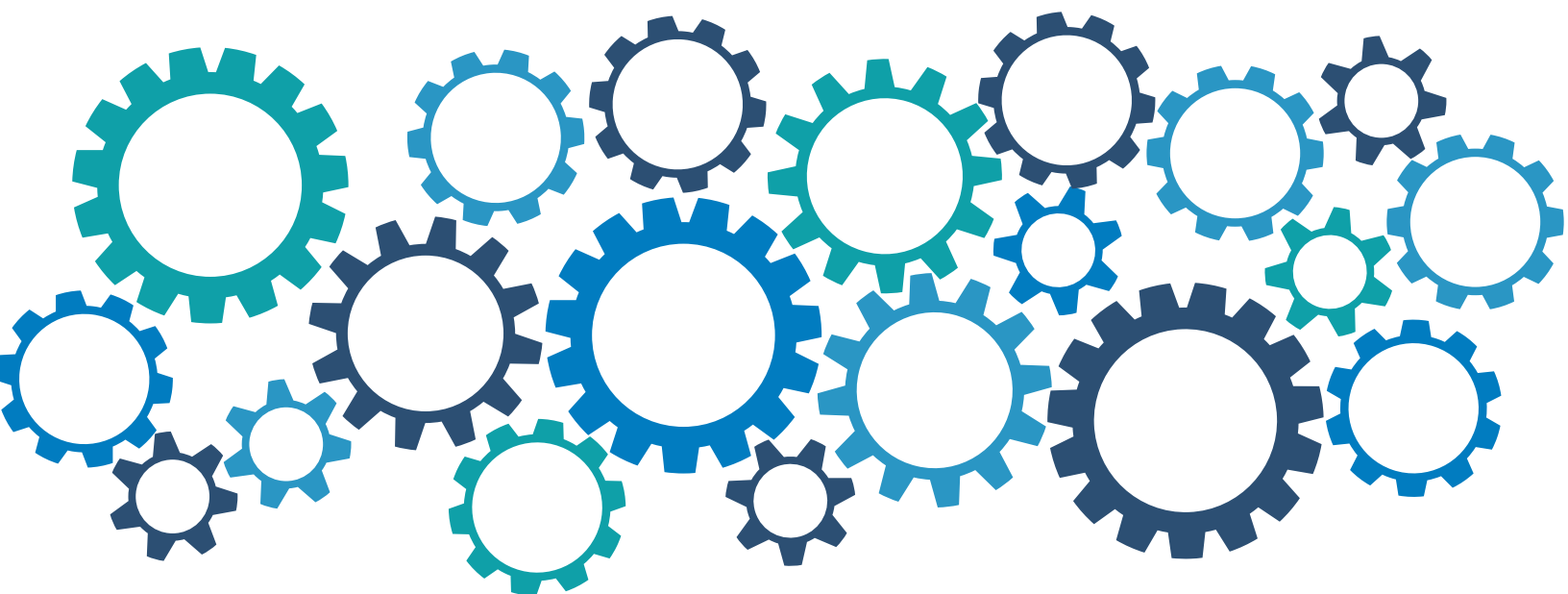




Towards a National Software Ecosystem

Report of the 2018 NSF Software Infrastructure for Sustained Innovation (SI²)
Principal Investigator (PI) Workshop

Washington D.C. Apr 30 – May 1, 2018

<https://si2-pi-community.github.io/2018-meeting/>

Frank Timmes (Chair)	Arizona State University
Sandra Gesing	University of Notre Dame
Paul Bauman	University of Buffalo
Frank Löffler	Louisiana State University, Friedrich Schiller University Jena
Kyle Niemeyer	Oregon State University
Rafael Ferreira da Silva	University of Southern California

This report is based on activities supported by the National Science Foundation under award number 1831393. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.



Contents

1	Executive Summary	2
2	Introduction	2
3	Planning and Execution	2
4	Workshop Agenda	3
5	Invited Talks	4
5.1	Conceptualizations and Implementations of SI ² Institutes	5
5.2	SSI and SSE Highlights	6
5.3	Additional Talks	7
6	Lightning Talks and Poster Sessions	7
7	Panel Discussion	9
A	Registrant List	10
B	Post-Meeting Poll Results	13

1 Executive Summary

“Towards a National Software Ecosystem”, the sixth NSF Software Infrastructure for Sustained Innovation (SI²) Principal Investigator (PI) workshop, was held at The Westin Washington, D.C. City Center on April 30–May 1, 2018. The workshop was attended by 161 PIs who collectively spanned the SI², Early-concept Grants for Exploratory Research (EAGER), Innovation Corps (I-Corps), Critical Resilient Interdependent Infrastructure Systems and Processes (CRISP), Computational and Data-Enabled Science and Engineering (CDS&E), and Designing Materials to Revolutionize and Engineer our Future (DMREF) programs. The workshop was also attended by at least 5 NSF Program Directors.

The workshop featured 146 posters and one-minute lightning talks, one for every project, distributed over four sessions. Significant blocks of time were allocated for the poster sessions to allow participants and Program Directors the opportunity to engage in deeper scientific interactions, forge new collaborations, and conduct on-the-ground discussions of how the “Towards a National Software Ecosystem” theme of the workshop relates to their software projects and the new Cyberinfrastructure for Sustained Scientific Innovation (CSSI) program. The remainder of this report presents the rationale, organization, and contents of the workshop.

2 Introduction

NSF’s vision of a Cyberinfrastructure Framework for 21st Century Science and Engineering (CIF21) identifies advancing new computational infrastructure as a priority for driving innovation in science and engineering. Innovation occurs through advances in computing facilities, software environments, advanced networks, data storage capabilities, and human capital and expertise. Software and data is thus an integral enabler of computation, experiment, theory and a central component of new computational infrastructure. Scientific discovery and innovation are advancing along new pathways opened by the development of increasingly sophisticated software. Software and data is also directly responsible for increased scientific productivity and significant enhancements of researchers’ capabilities. The sunset of the Software Infrastructure for Sustained Innovation (SI²) program, and the sunrise of the new Cyberinfrastructure for Sustained Scientific Innovation (CSSI) program, supports vibrant partnerships among academia, government laboratories, and industry for the development and stewardship of a sustainable software infrastructure that accelerates innovation in science and engineering.

As of January 2018 there are $\simeq 250$ SI² and SI²-related projects. The purpose of the “Towards a National Software Ecosystem” workshop is to reach five strategic objectives that align with the old SI² and new CSSI program goals and contribute to the broader CIF21 vision: (1) Serve as a focused forum for PIs to share technical information with each other and NSF Program Officers; (2) Explore innovative topics emerging within software communities; (3) Discuss emerging best practices across the supported software projects; (4) Stimulate thinking on new ways of achieving software sustainability; (5) Record the shared experiences in a workshop report. The URL of the workshop website, <https://si2-pi-community.github.io/2018-workshop/>, contains links to the websites of the previous 2013–2017 workshops.

3 Planning and Execution

The Organizing Committee for the 2018 workshop were:

- Frank Timmes (PI), Professor, School of Earth and Space Exploration, Arizona State University – <http://cococubed.asu.edu>
- Sandra Gesing (Co-PI), Research Assistant Professor, Department of Computer Science and Engineering, University of Notre Dame – <http://sandra-gesing.com>
- Paul Bauman (Co-PI), Assistant Professor, Department of Mechanical and Aerospace Engineering, University at Buffalo – <https://engineering.buffalo.edu/mechanical-aerospace/people/faculty/p-bauman.html>

- Kyle Niemeyer (Co-PI), Assistant Professor, School of Mechanical, Industrial, & Manufacturing Engineering, Oregon State University – <https://niemeyer-research-group.github.io>
- Rafael Ferreira da Silva (Co-PI), Research Assistant Professor, Department of Computer Science, University of Southern California – <http://rafaelsilva.com>
- Frank Löffler (Collaborator), Researcher and IT Consultant, Louisiana State University, Friedrich Schiller University Jena, Germany – <https://www.cct.lsu.edu/~knarf/>

The Organizing Committee held one-hour Google hangouts approximately once a week between October 10, 2017 and April 25, 2018. In early November 2017 the Organizing Committee, in consultation with the SI² PI Program Directors, converged on the vision, theme, and title of the meeting “Towards a National Software Ecosystem”. The meeting website went online in mid-November 2017. In January 2018 the Organizing Committee converged on featuring a Lightning Talk + Poster centric workshop to promote participant interactivity and networking. Integrated in between the four (4) Lightning Talk + Poster sessions were invited talks spanning SI² Institutes, Conceptualizations, Integrations (SSI), and Elements (SSE).

Leveraging continuity from previous SI² workshops, the 2018 workshop aimed to (1) Expand the number of non-SI² researchers who attend the workshop; (2) Invite EAGER, I-Corps, CRISP, CDS&E, and DMREF PIs to explore new or hitherto untargeted SI² areas by engaging non-SI² PIs and conversely, give the SI² community a sense of what is possible through the EAGER, I-Corps, CRISP, CDS&E and DMREF mechanisms; (3) Showcase the diversity of projects within the SI² program; (4) Feature domain scientists from SI² under-served domains; (5) Nurture a sustainable organizing committee process. Before, during, and after the workshop the Organizing Committee sought volunteers for the 2019 CSSI workshop Organizing Committee.

4 Workshop Agenda

Monday Apr 30, 2018

Time	Event	Speaker	Moderator
7:15am	Breakfast with Jim Kurose*		
7:30am	Breakfast and Registration		
8:30am	Welcome Address	Frank Timmes	Frank Timmes
8:45am	NSF CISE and OAC Directions	Jim Kurose Amy Friedlander	
9:30am	Coffee Break		
10:00am	Accessing Commercial Potential of Deep Technology Innovation	Cindy WalkerPeach	
10:05am	Lightning Talks #1		
10:50am	Posters Session #1		
noon	Lunch SI2		
1:00pm	Network for Computational Nanotechnology Cyber Platform	Gerhard Klimeck	Frank Löffler
1:30pm	SI2 Institutes: Science Gateways Software Institute	Nancy Wilkins-Diehr	
2:00pm	SI2 Institutes: Molecular Science Software Institute	Daniel Crawford	
2:30pm	Industry Highlight: Approaching Sustainability and Industry: A Longest but not Last Mile	Mike Zentner	
3:00pm	Coffee Break		
3:15pm	Lightning Talks #2		Kyle Niemeyer
4:00pm	Posters Session #2		
5:00pm	Reception		

Tuesday May 1, 2018

Time	Event	Speaker	Moderator
7:30am	Breakfast and Registration		

8:30am	Conceptualization: Geospatial Software Institute	Shaowen Wang	Rafael Ferreira da Silva
9:00am	Conceptualization: Scientific Software Innovation Institute For High Energy Physics	Peter Elmer	
9:30am	Conceptualization: Conceptualizing a US Research Software Sustainability Institute	Karthik Ram	
10:00am	Coffee Break		
10:15am	Lightning Talks #3		Paul Bauman
11:00am	Posters Session #3		
noon	Lunch		
1:00pm	SSI Highlight: STORM: A Scalable Toolkit for an Open Community Supporting Near Realtime High Resolution Coastal Modeling	Joannes Westerink	Sandra Gesing
1:15pm	SSE Highlight: Fast Dynamic Load Balancing Tools for Extreme Scale Systems	Mark Shephard	
1:30pm	REU Highlight: NCSA's INCLUSION (Incubating a New Community of Leaders Using Software, Inclusion, Inclusion, Innovation, Interdisciplinary and Open-Science)	Dan Katz	
1:45pm	SSI Highlight: Distributed Workflow Management Research and Software in Support of Science	Ewa Deelman	
2:00pm	SSE Highlight: Making Software Engineering Work for Computational Science and Engineering: An Integrated Approach	Jeff Carver	
2:15pm	Software Security: Selecting engineering and security practices to enable robust CI and trustworthy science	Von Welch	
2:30pm	Coffee Break		
2:45pm	Lightning Talks #4		Frank Timmes
3:30pm	Posters Session #4		
4:30pm	Panel Discussion: Combining Software and Data Programs	Ewa Deelman Boyce Griffith DK Panda Abani Patra Haiying Shen	Paul Bauman
5:00pm	Closing		

*: The Breakfast with Jim Kurose was limited to 12 attendees. Those spots were available by short application and were selected by the organizing committee, based on a mix of seniority levels with a focus on junior researchers, a wide field of award sciences and applicant-given reasons to attend. The following researchers were invited to attend: Christina Bandaragoda (University of Washington), Jerry Bernholc (North Carolina State University), Nathan Goldbaum (University of Illinois Urbana Champaign), Piotr Luszczek (University of Tennessee, Knoxville), Abhijit Majumder (Wayne State University), Kyle Mandli (Columbia University), Hyowon Park (University of Illinois at Chicago), Gregory Sharp (Massachusetts General Hospital), Haiying Shen (University of Virginia), Ali Shojaie (University of Washington), Frank Tip (Northeastern University), and Vincent Weaver (University of Maine).

5 Invited Talks

The workshop was kicked off by the NSF directors Jim Kurose and Amy Friedlander presenting “NSF CISE and OAC Directions”. The talks gave an overview on topics and areas of funded projects in the past years and an outlook for the next five years on key aspects and core areas. Furthermore, they went into detail for available funding per area. NSF director Cindy R. WalkerPeach presented the I-Corps program with “Accessing Commercial Potential of Deep Technology Innovation” emphasizing that the program helps to think like an entrepreneur. It supports to investigate target communities and fills the gap between research projects and setting up a start-up company. Several PIs in the audience shared their own experience within the I-Corps program, with the unanimous opinion that going through that program was, despite being intense, one of the best decisions they made, independent of the recommendation to commercialize or not.

5.1 Conceptualizations and Implementations of SI² Institutes

Invited talks included conceptualizations and implementations of SI² institutes and a long-existent institute that has been sustainable for years:

Network for Computational Nanotechnology Cyber Platform (Gerhard Klimeck). nanoHUB.org is a place for computational nanotechnology research, education, and collaboration. The site hosts a rapidly growing collection of simulation tools for nanoscale phenomena that run in the cloud and are accessible through a web browser. In addition to simulations, nanoHUB provides online presentations, cutting-edge nanoHUB-U short courses, animations, teaching materials, and more. These resources help users learn about our simulation tools and about nanotechnology in general. This presentation discussed the strategies, techniques used to reach $\simeq 500$ apps in the cloud, $\simeq 5000$ lectures and tutorials, $\simeq 100$ online courses, $\simeq 1.4$ million users annually. Particular emphasis was placed on making a software instrument or app usable by the target community.

SI² Institutes: Science Gateways Software Institute (SGCI) (Nancy Wilkins-Diehr). This presentation focused on connecting people and resources to accelerate discovery by empowering the science gateway community. The evolution of the SGCI from a 2009 EAGER award to a vision for an SI² institute emerged in 2016. SGCI Software-as-a-Service clients include COSMIC2 (Michael Cianfrocco, University of Michigan), LSU Systems Biology (Michal Brylinski, Louisiana State University), SimCCS (Kevin Ellet, Indiana Geological and Water Survey), ChemCompute (Mark Perri, Sonoma State University), nSides (Rami Vanguri, Columbia University), and Interactive Parallelization Tool (Ritu Arora, University of Texas). Data Distribution clients include Coastal Emergency Risk Assessment (Carola Kaiser, LSU) CitSci.org (Greg Newman, Colorado State University), ENIGMA (Lisa Eyler, UC San Diego), Ocean Observatories Initiative (Ivan Rodero, Rutgers University), and Aquavit (Jack Smith, Marshall University). Next steps for SGCI include interoperation amongst institutes, attendance/presentation at several institute planning workshops, engagement with S2I2 and CSSI projects, and increasing the reach of individual software projects.

SI² Institutes: Molecular Science Software Institute (Daniel Crawford). The Molecular Sciences Software Institute (MolSSI) is a nexus for science, education, and cooperation for the global computational molecular sciences community. It is a collaborative effort by Virginia Tech (TDC), Rice U. (C. Clementi), Stony Brook U. (R. Harrison), U.C. Berkeley (T. Head-Gordon), Stanford U. (V. Pande), Rutgers U. (S. Jha), U. Southern California (A. Krylov), and Iowa State U (T. Windus). The goals of MolSSI are to (1) provide software expertise and infrastructure; (2) to provide education and training; and (3) to provide community engagement and leadership. This presentation offered a broad overview on MolSSI, the communities involved, the challenges faced (e.g., millions of lines of hand-written source code across a diverse terrain of languages and algorithms), and mechanisms in place to address those challenges.

Conceptualization: Geospatial Software Institute (Shaowen Wang). The goal is to Conceptualize a Geospatial Software Institute (GSI) is a long-term hub of excellence in software infrastructure that can serve diverse research and education communities. Some of the scientific and societal challenges to be addressed include climate change, disasters, emergencies, food security, population growth, energy sustainability, water environment, and urbanization. The challenges encountered include data that is dynamic, distributed, heterogeneous, massive, and multi-scale, but also involve data quality and uncertainty. This presentation discussed the needs and challenges associated with leadership and software for transforming such geospatial data into actionable information, knowledge, and intelligence.

Conceptualization: Scientific Software Innovation Institute For High Energy Physics (Peter Elmer). The primary goal of the S2I2-HEP conceptualization project is to prepare a strategic plan for a potential NSF Scientific Software Innovation Institute (S2I2) to develop software for experiments taking data in the “High-Luminosity Large Hadron Collider” (HL-LHC) era in the 2020s. This presentation focused on ongoing community workshops and conceptual work to “take advantage of the significant data and computing requirements of the Large Hadron Collider as a science driver for next generation high-performance software and sustainability developments.” The HEP Software Foundation (HSF) was created in early 2015 as a means for organizing our community to address the software challenges of future projects such as the HL-LHC. The HSF has the following objectives: 1) catalyze new common projects, 2) promote commonality and collaboration in new developments to make the most of limited resources, 3) provide a framework for

attracting effort and support to software and computing common projects (new resources!). 4) provide a structure to set priorities and goals for the work.

Conceptualization: Conceptualizing a US Research Software Sustainability Institute (Karthik Ram). The goal of this conceptualization project is to design a US Research Software Sustainability Institute (URSSI). The motivations for a US-based institute to support the research community are legion: modern research relies on research software, lack of reproducibility is rampant even in applied computational research, available software frequently doesn't work or build, and lack of recognition of software as a scholarly product. The conceptualization project organizers have held an initial workshop with many community members, with the goal of identifying issues URSSI can address. These include: (1) training in effective computational skills; (2) policy recommendations aimed at creating better credit mechanisms and supporting career paths for developers of research software; (3) supporting software development by incubating projects, providing consulting support, and offering startup grants; (4) building community by disseminating best practices and handling governance. Additional workshops are being planned around the topics of software credit/citation metrics and a software incubator, followed by a final workshop to plan the institute.

5.2 SSI and SSE Highlights

SSI Highlight: STORM: A Scalable Toolkit for an Open Community Supporting Near Real-time High Resolution Coastal Modeling (Joannes Westerink). Coastal sustainability and risk means understanding water levels, currents, and wind waves in tides, global ocean circulation, storms, rainfall runoff, coastal flooding, wave forces, sinking deltas, wetland degradation, coastal dead zones, and marine larval transport. Physical processes and scale separation can be decomposed into shallow water equations, Boussinesq equations, spectral action balance equation, kinematic wave equation, dynamic wave equations, prognostic ocean circulation equations, Reynolds averaged Navier-Stokes equations. This presentation focused on the past, present, and future of coastal ocean hydrodynamic models. The past was represented by the ADCIRC and SWAN software instruments. The present was embodied by a dynamic interleaving of the ADCIRC, SWAN and HYCOM software instruments. The vision for the future included 1) fully mesh dynamic computations that during the simulation, 2) dynamic grid optimization for multi-physics, 3) higher order interpolation methods, 4) advance engines for load balancing, 5) developing frameworks that allow dynamic and coupled physics.

SSE Highlight: Fast Dynamic Load Balancing Tools for Extreme Scale Systems (Mark Shephard). Parallel simulation-based engineering workflows using unstructured meshes require adaptive methods to ensure reliability and efficiency. Starting with a problem specification on a geometric model, an effective workflow automatically executes parallel mesh generation, analysis, and analysis-based mesh and/or model adaptation. The analyze-adapt cycle is repeated until a desired level of solution accuracy is reached. Between each step in the cycle is an opportunity to improve scalability and efficiency through dynamic partitioning. This presentation focused on this project's tools for parallel unstructured mesh simulations, generalizations to multicriteria partition improvement procedures, and user applications currently being addressed. Efforts currently underway at Scientific Computation Research Center (SCOREC) include services for Parallel Unstructured Mesh Infrastructure (PUMI), Parallel Curved Mesh Adaptation (MeshAdapt), Partitioning using Mesh Adjacencies (ParMA), and its generalization to EnGPar for multicriteria load balance improvement. Domain specific applications include fields in particle accelerators, flow control to aircraft tails, deformation of mechanical parts, and plasma fields in tokamaks.

SSI Highlight: Distributed Workflow Management Research and Software in Support of Science (Ewa Deelman). In this talk, Dr. Deelman highlighted the sustainability plan for the development of the Pegasus Workflow Management System. Although from the past 17 years, Pegasus has been mainly funded by NSF programs, it is imperative to seek for alternative and collaborative opportunities. The talk discussed the lessons learned along these years (e.g., scientific tools development should be guided by the scientist's needs), and the challenges of working with evolving, heterogeneous, complex computational infrastructures and applications. The presentation also highlighted major achievements (e.g., supporting the discovery of gravitational waves), and the diversity of scientific application domains supported by Pegasus.

SSE Highlight: Making Software Engineering Work for Computational Science and Engineer-

ing: An Integrated Approach (Jeff Carver). This software engineering for science presentation focused on metrics and a metrics dashboard. Surveys identified architecture, code complexity, general quality, methodology, performance, process, recognition, and testing as categories having or needing unique metrics. A metrics dashboard was presented that listed tracked projects (e.g., Astropy, abjad, Golang, iPython, NumPy, Simbody, Spray, and SymPy) and several metrics for those tracked projects: thousands of lines of code (KLOC), defect density, and issue spoilage.

5.3 Additional Talks

In addition to the program-funded highlights, we had also a series of invited talks, on which the goal was to empower the SI² program with their complementary expertise:

Industry Highlight: Approaching Sustainability and Industry: A Longest but not Last Mile (Mike Zentner). This presentation focused on bridging the gap between federally funded research and a new commercialized technology, known as the “valley of death” where software projects can die. Specific examples of successful transitions included scheduling and supply chain optimization, structured and unstructured data analytics, medical device informatics, nanoHUB going to I-Corps, and HUBzero. Generally the transition can be made when offering solutions, not capabilities, to problems. One Science Place is a new address for the HUBzero Foundation, where a community of science gateways solve sustainability together.

REU Highlight: NCSA’s INCLUSION (Incubating a New Community of Leaders Using Software, Inclusion, Innovation, Interdisciplinary and Open-Science) REU Site (Dan Katz). The National Center for Supercomputing Applications (NCSA) at University of Illinois Urbana-Champaign ran the first year of the INCLUSION (Incubating a New Community of Leaders Using Software, Inclusion, Innovation, Interdisciplinary and Open-Science) REU Site. This REU Site supports undergraduate students of diverse backgrounds with a \$500/week stipend, room and board, and travel allowances to work on a summer project at NCSA to develop, study, and/or apply open-source software to socially impactful problems in all research areas. Projects are proposed by NCSA faculty and staff, and pairs of students work with pairs of mentors. The program includes pre-summer mentor mentoring of faculty, postdocs, and staff and a Software Carpentry workshop for student participants. Regular student activities during business hours include work on their projects, weekly group meetings with the INCLUSION PIs, weekly lightning talks by staff and other students, and some additional special events. In evenings, students meet with UIUC Summer Research Opportunities program staff, have weekly writing meetings, and participate in organized social events. The first year outcomes include relatively high portion of female students (but less diversity in other areas), overall student satisfaction with the program (aside from networking opportunities). Goals include targeted recruitment of minority students, increased diversity of mentors, better communication with mentors about program goals, and better continuity of projects.

Software Security: Selecting engineering and security practices to enable robust CI and trustworthy science (Von Welch). This software security presentation focused on steps for transitioning software from development to operations to science from a cybersecurity Perspective. A goal is help the NSF community achieve their science mission with actionable, reasonable guidance regarding cybersecurity, technology, operations, software development, management, and budgeting. It is stressed that “software as research” is not equal to “software for research”. This NSF Cybersecurity Center of Excellence aims to build consensus between cyberinfrastructure operators and developers on secure software. A Level 1, minimally functional software project, includes 1) use of revision control, 2) documenting dependencies and the build process, 3) providing or documenting the build infrastructure, 4) a changelog mechanism, 5) development status, and 6) a license. A level 2, basic software engineering practice includes 1) all of Level 1 requirements, 2) semantic versioning, 3) distributing software, 4) code signing, 5) basic security policy, including vulnerability management, 6) dependency selection, 7) succession, 8) issue tracker, 9) testing.

6 Lightning Talks and Poster Sessions

The poster session in past events was often problematic. Attendees to those poster sessions gave the feedback that they (1) did not have a lot of interaction, (2) could not go to other posters because they had to stay at

their own posters, (3) were not aware which projects were represented at the workshop, and (4) could not easily to search for suitable collaborators.



Figure 1: Networking during Poster Session #3.

To address these issues, we organized four poster sessions. This strategy divided the number of posters per session by four. Thus, poster presenters had an overlap with at most $\simeq 39$ other posters in a session, and were otherwise free to visit other poster in the three remaining sessions. To give everyone the opportunity to give a brief overview on their project to the community and give the audience an impression of the upcoming poster session, before each poster session we arranged a *lightning talk* session. The lightning talks were bound to *one minute* and *one slide* to give a summary about a project. We have asked for the slides one week in advance and we put them together in a slide deck per session (available at https://si2-pi-community.github.io/2018-meeting/poster_sessions.html). To keep the presenters to their time, a session chair assisted with lining up the speakers in the order published. A time taker would advance the presentation to the next slide after strictly a minute so that the each presenter and the chair had a clear sign that the presenter was exceeding the time. In addition, a seconds-timer was visible to each speaker to allow easy self-control. This worked very well and we stayed within the allocated time frame of each session. Upon being aware of the projects and topics presented, attendees enjoyed noticeably more lively and interactive poster sessions.



Figure 2: Networking during Poster Session #4.

For defining the order of the lightning talks we used the following algorithm:

1. Assignments of people to a specific session if they requested it in advance due to time constraints. We did not offer this opportunity and only reacted when people contacted us directly;
2. Balance between the number of SI2 institutes, SSIs, SSEs, SI2-Conceptualizations, and other award

types in each session; and

3. Order of registration of the attendee.

The final assignments for poster/lightning talks sessions can be found at https://si2-pi-community.github.io/2018-meeting/poster_sessions.html.

Session	SI2-Institute	SI2-SSI	SI2-SSE	SI2-Conceptualization	Other
#1	1	14	20	2	2
#2	1	13	18	1	5
#3	1	14	17	1	6
#4		6	17	1	6
Total	3	47	72	5	19

Table 1: Distribution of posters/lightning talks per session.



Figure 3: Left: Panel discussion on combining software and data programs. Right: panelists seated from left to right Ewa Deelman, Boyce Griffith, DK Panda, Abani Patra, and Haiying Shen. The panel discussion was chaired by Paul Bauman and Vipin Chaudhary.

7 Panel Discussion

In the last session of the workshop we have organized a panel with the theme “Combining Software and Data Programs”. The panel was composed Ewa Deelman (University of Southern California), Boyce Griffith (UNC-Chapel Hill), DK Panda (The Ohio State University), Abani Patra (University at Buffalo), and Haiying Shen (University of Virginia). There were six prepared questions that were sent to the panelists the day before the panel convened:

1. What does a national ecosystem mean to you - shared tools, people, projects, mother of all libraries?
2. What challenges, if any, and opportunities do you see in combining data and software infrastructures?
3. What is the difference between a software project and a data project?
4. What training is needed for the next-generation of computational (software and data) scientists?
5. How do we motivate long-term or tenured positions bridging domain science, software engineering, and data research?
6. Should all software and data projects live forever?

It was anticipated that there would not be time to address all questions in the panel, but we wished to allow the panelists to be as prepared as possible. The moderators posed each question to the panel and allowed each panelist to answer in turn. Then, the audience was allowed to ask questions to panel or comment on a panelist’s response or to the question posed to the panel.

A Registrant List

A total of **165** participants registered for the workshop, including three employees of NSF. Out of these, 140 were SI² PIs. The registered PIs represented 154 NSF awards:

Award Type	Count
SI2-SSE	75
SI2-SSI	50
Other	11
EAGER	5
SI2-Conceptualization	4
SI2-Institute	3
CDS&E	3
DMREF	2
CRISP	1

The 165 registered participants and the projects they represented were:

Name	Organization	Award Type	NSF Award Number
Abani Patra	Univ at Buffalo, SUNY	SI2-SSI	ACI1339765
Abhijit Majumder	Wayne State University	SI2-SSI	ACI-1550300
Alberto Passalacqua	Iowa State University	SI2-SSE	ACI 1440443
Ale Strachan	Purdue University	SI2-SSE	1440727
Aleksei Aksimentiev	University of Illinois at Urbana-Champaign	SI2-SSE	OAC-1740212
Alexander D. Kaiser	Stanford University	SI2-SSI	1663671
Alexander Withers	NCSA	SI2-SSE	1535070
Ali Shojaie	University of Washington	CDS&E	DMS-1722246
Alyssa Goodman	Harvard University	SI2-SSE	1739657
Amarda Shehu	George Mason University	SI2-SSE	1440581
Andreas Mueller	Columbia University	SI2-SSE	OAC-1740305
Andreas Stathopoulos	College of William and Mary	SI2-SSE	1440700
Andrew Connolly	University of Washington	SI2-SSE	1739419
Andrew Lumsdaine	University of Washington	SI2-SSE	1716828
Andrew Miner	Iowa State University	SI2-SSE	1642397
Ankur Srivastava	University of Maryland	SI2-SSE	1642424
Anton Van der Ven	University of California Santa Barbara	SI2-SSE	1642433
Azzam Haidar	University of Tennessee	SI2-SSE	OAC-1740250
B. S. Manjunath	UCSB	SI2-SSI	1664172
Barry Drake	Georgia Institute of Technology	SI2-SSE	1642410
Boyce Griffith	UNC-Chapel Hill	SI2-SSI	ACI 1450327
Brian Bockelman	University of Nebraska-Lincoln	SI2-SSI	1450323
Brian Demsky	University of California, Irvine	SI2-SSE	1740210
Bruce Berriman	Caltech/IPAC-NExSci	SI2-SSE	1642453
Bruce Childers	University of Pittsburgh	SI2-SSE	ACI-1535232
Bryan Heidorn	University of Arizona	SI2-SSE	1642446
Carl Boettiger	UC Berkeley	EAGER	1549758
Carlos Maltzahn	University of California, Santa Cruz	SI2-SSI	1450488
Chad Hanna	Penn State	SI2-SSE	ACI-1642391
Charles Torre	Utah State University	SI2-SSE	1642404
Christina Bandaragoda	University of Washington	Other	1810886
Christine Goulet	SCEC/USC	SI2-SSI	1450451
Christopher Iacovella	Vanderbilt University	SI2-SSE	1535150
Christopher Iacovella	Vanderbilt University	SI2-SSI	1047828
Christopher Paciorek	UC Berkeley	SI2-SSI	ACI-1550488
Christopher Roland	NC State University	SI2-SSE	1534941
Damian Dechev	University of Central Florida	SI2-SSE	1740095
Dane Morgan	University of Wisconsin, Madison	DMREF	1728933
Dane Morgan	University of Wisconsin, Madison	SI2-SSI	1148011

Daniel Crawford	Virginia Tech	SI2-Institute	ACI-1547580
Daniel S. Katz	University of Illinois Urbana-Champaign	Other	1659702
David Anderson	UC Berkeley	SI2-SSI	1664190
David Hudak	Ohio Supercomputer Center	SI2-SSE	1534949
David Kofke	University at Buffalo	SI2-SSE	OAC-1739145
David Schloen	University of Chicago	SI2-SSI	1450455
David Tarboton	Utah State University	SI2-SSI	1664061
Davide Curreli	University of Illinois at Urbana Champaign	SI2-SSE	1740310
Devangi Parikh	The University of Texas at Austin	SI2-SSI	ACI-1550493
Dhabaleswar K (DK) Panda	The Ohio State University	SI2-SSI	1664137
Douglas Thain	University of Notre Dame	SI2-SSE	1642409
Edward Valeev	Virginia Tech	SI2-SSI	1550456
Elbridge Gerry Puckett	University of California, Davis	SI2-SSE	1440811
Emanuel Gull	University of Michigan	Other	DMR-1606348
Emery R. Boose	Harvard University	SI2-SSI	1450277
Emre Brookes	University of Texas Health Science Center at San Antonio	SI2-SSE	1740097
Eric Polizzi	University of Massachusetts, Amherst	SI2-SSE	1739423
Ewa Deelman	University of Southern California	SI2-SSI	1664162
Francesco Paesani	University of California San Diego	SI2-SSE	ACI-1642336
Frank Löffler	Friedrich Schiller University Jena	Other	
Frank Petriello	Northwestern University	SI2-SSE	1740142
Frank Timmes	Arizona State	SI2-SSI	1663684
Frank Tip	Northeastern University	Other	1715153
Ganesh Gopalakrishnan	University of Utah	SI2-SSE	OAC 1535032
Gene Cooperman	Northeastern University	SI2-SSE	OAC-1740218
George Bosilca	University of Tennessee	SI2-SSI	1664142
Gerhard Klimeck	Purdue University	Other	EEC-1227110
Grady Wright	Boise State University	SI2-SSE	1440638
Greg Newman	Colorado State University (CitSci.org and the Natural Resource Ecology Laboratory)	SI2-SSI	1550463
Greg Tucker	University of Colorado / CSDMS	SI2-SSI	1450409
Gregory Sharp	Massachusetts General Hospital	SI2-SSE	1642380
Gregory Voth	The University of Chicago	SI2-SSE	OAC-1740211
Haiying Shen	University of Virginia	CDS&E	1404981
Haiying Shen	University of Virginia	Other	1724845
Harish S. Bhat	University of California, Merced	CDS&E	DMS-1723272
Hassen Saidi	SRI	SI2-SSE	1440800
Heike Jagode	University of Tennessee Knoxville, Innova- tive Computing Lab (ICL)	SI2-SSE	1642440
Henri Casanova	University of Hawai'i at Manoa	SI2-SSE	1642369
Hyowon Park	University of Illinois at Chicago	SI2-SSE	1740112
Icaro Alzuru	University of Florida	SI2-SSE	1535086
Ivo Jimenez	UC Santa Cruz	SI2-SSI	1450488
James Bordner	University of California, San Diego	SI2-SSE	1440709
Jan Verschelde	University of Illinois at Chicago	SI2-SSE	1440534
Jason Leigh	University of Hawaii at Manoa	SI2-SSI	1441963
Jeffrey Carver	University of Alabama	EAGER	1445344
Jeffrey Potoff	Wayne State University	SI2-SSE	ACI-1642406
Jerry Bernholc	NC State University	SI2-SSE	OAC-1740309
Jin-Jian Zhou	California Institute of Technology	SI2-SSE	1642443
Jindal K. Shah	Oklahoma State University	SI2-SSE	OAC-1339785
Joannes J Westerink	University of Notre Dame	SI2-SSI	ACI-1339738
John Mellor-Crummey	Rice University	SI2-SSI	ACI-1450273
Joseph Stubbs	University of Texas at Austin	SI2-SSE	1740288
Juan Zhai	Purdue University	EAGER	1748764-CCF
Karthik Ram	University of California, Berkeley	SI2-Concept.	1743188
Kenneth Jansen	University of Colorado Boulder	SI2-SSE	1740330
Kesong YANG	UC San Diego	SI2-SSI	ACI-1550404

Kyle Chard	University of Chicago	SI2-SSI	1550588
Kyle Mandli	Columbia University	CRISP	1735609
Kyle Niemeyer	Oregon State University	SI2-SSE	1535065
Laxmikant Kale	University of Illinois	SI2-SSI	1339715
Lincoln Carr	Colorado School of Mines	SI2-SSE	1740130
Louise Kellogg	University of California, Davis	EAGER	1448633
Luis Agapito	Caltech	SI2-SSE	1642443
Luis Oliveira	University of Pittsburgh	SI2-SSE	1535232
Madhusudhan Govindaraju	SUNY Binghamton	SI2-SSE	1740263
Mark Ghiorso	OFM Research	SI2-SSI	1550346
Mark Neubauer	University of Illinois at Urbana-Champaign	SI2-Concept.	ACI-1558233
Mark Shephard	Rensselaer Polytechnic Institute	SI2-SSE	1533581
Marlon E Pierce	Indiana University	SI2-SSI	1339774
Matteo Turilli	Rutgers	SI2-SSE	1440677
Matthew Newville	The University of Chicago	SI2-SSI	1450468
Matthew Parno	Cold Regions Research and Engineering Laboratory	SI2-SSI	1550487
Maxine Brown	University of Illinois at Chicago	SI2-SSI	ACI-1441963
Michael	Purdue University	SI2-Institute	1547611
Michael Bell	Colorado State University	SI2-SSI	1661663
Mike Sokoloff	University of Cincinnati	SI2-SSE	OAC-1740102, OAC-1450319, OAC-1558219
Mohammad Ebrahim Khalaj	Wayne State University	SI2-SSI	1550300
Muhammad Osama	University of California Davis	SI2-Institute	OAC-1740333
Nancy Wilkins-Diehr	San Diego Supercomputer Center	SI2-Institute	ACI-1547611
Nathan Goldbaum	University of Illinois Urbana Champaign	SI2-SSI	#1663914
Neil Heffernan	Worcester Polytechnic Institute	SI2-SSE	1440753
Omar Ghattas	The University of Texas at Austin	SI2-SSI	1550593
P. Bryan Heidorn	University of Arizona	SI2-SSE	1642446
Paul Bauman	University at Buffalo	SI2-SSE	1642388
Paul Butler	The University of Tennessee/NIST	SI2-SSI	CHE-1265821
Paul Navratil	Texas Advanced Computing Center	SI2-SSI	ACI-1339863
Paul Rulis	University of Missouri - Kansas City	DMREF	1729227
Paul Saxe	Molecular Sciences Software Institute	SI2-Institute	ACI-1547580
Peter Diener	Louisiana State University	SI2-SSI	1550551
Peter Elmer	Princeton University	SI2-Concept.	1558216
Peter Volgyesi	Vanderbilt University	SI2-SSE	1740151
Philip A. Wilsey	University of Cincinnati	SI2-SSE	ACI-1440420
Philip Maechling	Southern California Earthquake Center	Other	EAR-1551411
Piotr Luszczek	University of Tennessee	SI2-SSE	1642441
Rafael Ferreira da Silva	University of Southern California	SI2-SSE	1642335
Raheem Beyah	Georgia Institute Of Technology	SI2-SSE	1643032
Rajiv Ramnath	National Science Foundation	Other	
Ramakrishnan Kannan	Oak Ridge National Laboratory	SI2-SSE	1642410
Ray Zimmerman	Cornell University	SI2-SSE	1642341
Reuben D. Budiardja	University of Tennessee	SI2-SSE	1535130
Rion Dooley	University of Texas	SI2-SSI	1450459
Ritu Arora	TACC, UT Austin	SI2-SSE	1642396
Ryan May	UCAR/Unidata	SI2-SSE	1740315
Sameer Shende	University of Oregon	SI2-SSI	ACI-1450471
Sandra Gesing	University of Notre Dame	SI2-Institute	1547611
Shantenu Jha	Rutgers	SI2-SSI	1265788
Shaowen Wang	University of Illinois at Urbana-Champaign	SI2-Concept.	1743184
Sohrab Ismail-Beigi	Yale University	SI2-SSI	ACI-1339804
Stefan Robila	National Science Foundation	Other	
Suresh Marru	Indiana University	SI2-SSI	1339774
T. Daniel Crawford	Virginia Tech	SI2-Institute	ACI-1547580
Tim Menzies	NC State	EAGER	

Timo Thonhauser	Wake Forest University	Other	1712425
Upulee Kanewala	Montana State University	Other	1656877
Vincent Reverdy	University of Illinois at Urbana-Champaign	SI2-SSE	1642411
Vincent Weaver	University of Maine	SI2-SSI	1450122
Vipin Chaudhary	National Science Foundation	Other	
Volker Blum	Duke University	SI2-SSI	1450372
Von Welch	Indiana University	Other	1547272
Warren B. Mori	UCLA	SI2-SSI	ACI-1339893
Wen-Chau Lee	NCAR	SI2-SSI	1550597
Wolfgang Bangerth	Colorado State University	SI2-SSI	OCI-1148116
Xiangyu Zhang	Purdue University	EAGER	1748764
Xiaosong Li	University of Washington	SI2-SSI	OAC-1663636
Ye Zhao	Kent State University	SI2-SSE	1535031
Ye Zhao	Kent State University	SI2-SSE	1739491
Yifei Mo	University of Maryland	SI2-SSI	1550423
Yifei Mo	University of Maryland	SI2-SSI	ACI-1550404
Ying Li	University of Connecticut	Other	1755779
Yosuke Kanai	University of North Carolina at Chapel Hill	SI2-SSE	1740204
Yung-Hsiang Lu	Purdue University	SI2-SSE	1535108

B Post-Meeting Poll Results

After the meeting the Organizing Committee conducted an anonymous poll of seven questions on the value and mechanics of the meeting. A total of 47 PIs responded to at least one question, about 35% of PIs attending.

1) *Please describe your primary expertise in your own words (e.g., computer scientist, software engineer, physicist).*

Computer scientist - 15
Physicist - 9
Software Engineer - 4
Computational Scientist - 3
Mathematician - 3
Chemist - 2
Ecologist - 2
Astronomer - 2
Hydrologist - 2
Software Developer - 2
Geochemist - 1
Geoscientist - 1
Computer Engineer - 1
Molecular simulator - 1
Mechanical engineer - 1
Material Scientist - 1
Statistician - 1
Information Science - 1
Science - 1
Educator - 1
Manager - 1
Supercomputer Center Director - 1

2) *If you have attended past SI2 PI Meetings, does attending the SI2 PI Meetings continue to add value? Describe how or why.*

- Found out about what the agency thinks, and its future plans, and opportunities I didn't know about.
- Yes. Interacting with other SI2 PIs and learning about the SI2 funded projects.
- yes. networking. becoming aware of what is new.

- yes, mostly for the people, but also for the information about the projects that might collaborate meeting/interacting with other SI2 PIs is interesting
- More this year than last year. Got to interact with more people on 1-1. the lightning talks and poster sessions exposed well the breadth of work happening.
- No. Not very useful or stimulating.
- I like the SI2 PI meeting to build further collaborations and get an overview, which SI2 projects are funded. yes; it is useful to find out what people in the community are working on
- Yes
- Yes! It's a great mechanism for seeing what else is going on in the community and meet potential collaborators.
- Hardly. What these meeting are missing is a vision of what the meeting is supposed to achieve. Is it parading everyone in front of NSF program officers? That works. But I would hope that the meeting would facilitate building a community and cohort of scientific software developers, and for this cramming everyone into a basement room with a forced agenda is not working. If the goal is to achieve an *exchange* of knowledge, then the program ought to focus on bringing people together in a productive environment, rather than one where 3/4 of the participants are sitting at their tables bored looking at their email inboxes.
- Yes. PIs come and go, the SI2 program changes (merges), so there's always new things to discuss.
- This meeting was a welcome departure from previous SI2 PI meetings. The previous "unconference" format typically devolved into gripe sessions or pet issues with little actionable content, generally or for specific projects. The format for this year's meeting allowed for both general awareness of the scope of projects through the lightning talks and time for more meaningful interaction through the smaller and separate poster sessions. I obtained more ideas for future work and collaborations, and more and more meaningful interactions at my poster, than the combination of all the previous SI2 workshops I have attended.
- Yes. The most valuable aspects to me were networking and opening new potential collaborations for my software project.
- This was by far the best SI2 PI Meeting i have attended. Yes!
- I always enjoy and get something from these meetings. It's interesting to see how other domain communities deal with common software challenges. I like learning about tools and techniques that might translate between domains. It's also valuable to learn about the many challenges that communities share in common.
- Yes. Important contacts - the most important ones related to my own domain, however.
- Yes and no. The topics are generally so far flung that most talks and posters are of no use to me (and I assume vice versa): I don't really understand what they are about and how they can help me or people similar to me. 1-2 talks/posters were interesting and potentially useful, so I guess it was better than nothing. It was good of course to see where the program is going (data, data, data like so many other agencies and universities as a giant bandwagon).
- Learning from and interacting with other PI's facing the same challenges. Seeing the broader perspective and scope of this program beyond individual projects.
- I have attended last year and this year's SI2 meetings. I think the real value is in communicating with other PI's and seeing the other projects. Other projects are good places to look for collaboration.
- Yes - networking with other developers and integrating software!
- Yes. Networking with other PIs was very useful as a community of researchers working on software engineering issues.
- The main contribution is a chance to learn about other related projects. These change year by year so this continues to be useful.
- The value is to meet people from the community and to find new collaboration opportunities. There was lots of food for thought and it is very instructive to see examples of success.
- I thought it was valuable to see all the other projects and especially to connect with people who are doing work related to my project that could help.
- Yes. It helps identify people with similar interests and expertise.
- this meeting was a better format than last year

3) *Did the lightning talks provide economy and value for the ensuing poster sessions?*

- Yes
- Absolutely. I found they provided an overview of the posters in the ensuing session and helped me focus where I wanted to spend my time.
- minimal.
- yes
- no
- Yes. It helped me identify which people I wanted to talk to (several of whom I would not have thought were relevant by the title of their grant) while also giving a clear sense of the breadth of the projects funded by nsf in this area.
- Yes!!
- yes
- Yes, they were excellent
- Yes! They were great actually.
- Yes, it allowed me to quickly identify which posters I wanted to spend a lot of time looking at and discussing.
- Absolutely, extremely helpful.
- Yes, keep it in
- They were fantastic. It was a really nice way to get the “big picture” range of all the different projects represented, and a great way to make connections that I wouldn’t have made just milling around the posters.
- I should note that the poster format itself was sub-optimal, in that the tight space layout made it really hard to circulate, especially on the second day.
- No.
- I liked the lightning talks a lot. I took notes to which of the posters I would like to go in the following poster session. It worked really well!
- yes
- Not really
- Yes
- Hardly.
- most certainly
- Yes, absolutely. The lightning talks were a great format to capture the breadth and diversity of topics and to highlight “must see” posters during the session. Much better being able to know to look for a poster than to chance upon it in the morass.
- Yes, loved the format, cannot believe we actually accomplished that
- Yes! I think the lightning talks were very nicely executed (better than other places I have seen them). They provided great “guidance” for posters in the session (since it was sometimes not possible to get to each one). Looking around the room, the lightning talks often had more “eyeballs” paying attention than the regular talks.
- The lightning talks were very useful to highlight the various posters that were being shown—it made it easier to find those I really wanted to see.
- Yes. This format should be continued.
- It helped me a lot to see the breadth of the program and identify which posters I was most interested in
- Honestly, I didn’t love the lightning talks. I understand the motivation, and it was worth trying, but I think it’s simpler and a better use of time to allow participants can read poster titles and abstracts on their own (or simply browse, as they prefer).
- yes
- Yes. I was skeptical at first, but keeping them short and to a minimum was great. Would recommend doing it again.
- Yes, found the lightning talks beneficial.
- Yes they were very good.
- Yes.

- Yes - helped me focus in on posters most relevant to my domain - very helpful!
- Yes, definitely. Very well done.
- yes, it did (a bit to my surprise)
- It is a plus. It is difficult to make it through the online posters and understand the key points. The lightning talks are faster.
- no, I'd rather just have perused a list of titles/abstracts and wandered amongst the posters. The lightning talks didn't add anything for me.
- Yes.
- Yes, it gave me an idea of what posters I definitely wanted to check out.
- Yes, they were quite helpful.
- The lightning talks were great, and definitely helped me to target discussions. Also, it was good to have an overall view of what is going on on the SI2 community.
- Yes, it makes it possible to identify the posters that are of interest.
- Yes! It really helped not missing specific posters of higher interest.
- I definitely had more foot traffic than other PI meetings

4) *Did you prefer the format of this year's meeting (more presentations) compared with previous meetings (more breakout sessions)? If you did not attend last year, did you like the format?*

65% Yes, this format was great.

16% No, I prefer more breakout discussions.

6% I don't like either format.

13% Other.

5) *Did you get what you wanted out of this meeting? Why or why not?*

- Yes.
- Yes. I learned about most (all?) of the SI2 projects and got to interact with several PIs, some of whom I knew and ones that I didn't previously know.
- partially. I wanted more feedback on reproducibility and replicability, and there was very little of that at the meeting.
- mostly
- Yes. I was hoping to make connections with some other groups that would be interested in my software but also who have projects that could augment my project and my research in general.
- Yes. Met with those people that I needed to.
- I think yes. The work done in this program is all very impressive, but the group seems so diverse and broad that it is sometimes hard to find common interests or even language between PIs. Having been to these before, I had fairly low expectations for hearing new things that readily applied or informed my work. So even 1 or 2 potential new projects of direct interest seems like a positive outcome.
- It was not what I expected, but very valuable
- Yes.
- This was a great meeting for networking and making contacts. I also felt like I was able to build excitement about my project with other interested PIs.
- Yes! I learned of many new tools I can use for my open simulation platform. I also got an idea of how to build from an SSE to an SSI
- Yes, very useful broad overview with targeted one-on-one interactions
- Yes
- No. The discussions are very superficial.
- Yes; I got to see people and hear what NSF program directors are thinking about.

- Yes, I got the chance to interact with new collaborators
- Yes. I discussed my work with two high-potential collaborators, as well as a few other lower-potential collaborators.
- I satisfied my contractual obligations to come to the meeting, and I had some interesting conversations with colleagues who are doing things I find stimulating. But I think this was independent of the program, rather than facilitated by the program.
- I did. I met people I knew and expected to meet, but I also met people I didn't know before: something I was especially looking forward to.
- I actually look forward to next year's meeting (if in this format), which is a marked and welcome departure from previous experience.
- Yes, I think so.
- Yes - met a few new contacts and I got a good idea of the "breadth" of the program
- I did manage to do some networking, but the focus on formal talks, from projects themselves, seemed to hinder this objective.
- Yes. I wanted in-person meetings with Program Directors, get a larger view of the current SI2 program, and forge new collaborations.
- I wanted to relearn the scope of the SI2 community, and this meeting helped me reach that.
- Yes, basically
- Yes. New contacts / collaborators.
- I did my duty coming to it, so that was good. I saw some of my colleagues in my field, that was good. I got to see where the program is going, that was good. But frankly 90
- Was not sure what to expect, being my first meeting.
- Generally yes
- Yes, I did. I got to meet a couple of project PI's that complement what I'm doing and talk about collaborative next steps.
- I really liked this meeting and valued the networking time as well as learning about the institutes in the various longer presentations.
- Yes. Networking was great, and also it was great to be able to provide/see inputs from diverse groups of people with different disciplinary backgrounds.
- yes. I talked to the people I should have talked to.
- I didn't get much out of it. Waste of time.
- I found a new collaborator, maybe even two.
- I didn't know what to expect out of the meeting, so it's hard to say I did or did not get what I wanted. I did find the experience valuable.
- Yes, I had the opportunity to engage with several PIs and also briefly discuss with the PM.
- It was interesting, but somewhat overwhelming due to the extremely broad scope.
- Yes - my goal was to connect with other researchers and better understand the range of science covered. That was achieved.
- Yes. I grew the user base of our software.
- Yes! I got introduced to new potential collaborators and was able to improved ideas and confidence in our current project approach

6) *How many potential new collaborators did you meet?*

0 - 4 replies, 8.9%
 1 - 3 replies, 6.6%
 2 - 7 replies, 15.6%
 3 - 12 replies, 26.7%
 4 - 8 replies, 17.8%
 5 - 8 replies, 17.8%
 8 - 1 reply, 2.2%
 10 - 2 reply, 4.4%

7) *Comments or suggestions for improvements for the 2019 Cyberinfrastructure for Sustained Scientific Innovation (CSSI) PI workshop organizers?*

- Repeat this format if possible. I found no value from the breakout sessions the previous year. Of course this will be challenging given the merger of DIBBS and SI2 into CSSI, which may double (or more) the size of the meeting. Perhaps make lightning talks voluntary to keep the time down? Or only lightning talks from projects that have gone through their first year?
- better room (day 2). more time for poster sessions.
- better physical space - posters were too crammed together, and the basement room on the second day was really bad for the talks. When there is a single meat option at a meal, it should not be pork.
- The room on the second day was too small. I think it would be interesting if next year there were some slightly more technical talks about how to utilize some of the emerging technologies that could help with software development, testing and deployment. Things like why everyone should be developing using git (and services like github), communication tools that integrate with code development (slack, gitter), continuous integration, deployment schemes (anaconda, docker, singularity). I think surveying those with grants in terms of what they do in terms of development and deployment would give a sense of what areas we as a community need guidance in terms of best practices. The organizing committee did a great job overall.
- The hotel location in downtown DC was great, but the meeting rooms (especially the 2nd day) was terrible, and the poster sessions were cramped and poorly lit. Making the meeting 1.5 days would also be helpful.
- Lightning & poster were great! The longer talks were OK, but I think some could have been shortened.
- I think have posters printed at or near the venue and already put up according to the area of study that they are associated with, so there isn't a lot of searching and finding for things posters that you are interested in.
- It was perfect. Please do it like this again, and thanks for your hard work!
- Minor thing: better layout of tables so the chairs face the screens.
- The meeting room on day 2 was really sub-optimal. ample space to move around and see the presenters is so important, and all the more so for the poster sessions. Otherwise, excellent meeting and job well done.
- Remove them.
- 2 full days seemed too long; the most useful thing is the networking; I also like to hear what NSF program directors have to say — the problem with adding in disciplinary talks is that it is a very interdisciplinary audience ...
- Stick to 20 minute limits for talks; 30 is a bit too long
- Articulate a vision for what these meetings are supposed to achieve, and then ask how the program can get you there.
- Some of the program types (lightning talks) that are just barely still viable with 160 people will be very, very hard or impossible to do with double the size. If possible, avoid >300 people for this type of meeting. It encourages people too much to stay isolated and anonymous instead of a part of the discussion like at SI2 meetings.
- Good idea combining with a relevant conference in the area, and for moving it later in the year, since we always seemed to manage to catch a February snow storm in previous years. However, running until 5pm the last day makes heading home on the same day difficult for those not able to stay for the conference the remainder of the week. The two-day format seemed better paced though than the previous 2.5 day form.
- Nothing off hand, thought you all did great considering the circumstances.
- The Westin was only "ok" as a location. I'm always a big fan of having inexpensive workshops in inexpensive cities.
- I thought it was run very well this year.
- The same group should run the CSSI meeting!
- Try running 2 hour clinics, as CSDMS does at their annual meetings. These can be hands-on mini-workshops at which participants get to learn and try out a new tool or technique. It's a great way to dip your toe in the water while meeting new people. You can have as many running concurrently as there are rooms available (it helps to have people sign up in advance, so undersubscribed ones can be canceled and popular ones capped). I'd love to see a clinic on Pegasus, for example, or on how to publish in JOSS.

- Format was good. Replacing 1-2 non-domain specific speaker sessions with domain specific parallel sessions could be even better, similar to what was done at the last PI meeting of MGI. See here: <https://app.certain.com/profile/form/index>. The MGI meeting also had the very good idea of inviting representatives from key centers funded by other agencies along and encouraged developments of joint activities with those. Finally: Requiring the finished product for a conference a week ahead of time is a nice idea but I don't think this approach works in practice.
- Less talks. Most talks were not useful. But the ones from people who actually did something (like nanohub or the "business" talk from other nanohub guy) were so useful because they were about something concrete you could relate to.
- Pay MUCH more attention to facilities. The room the second day was REALLY TERRIBLE. On both days there was insufficient space around posters to have a good conversation. Apart from the 2nd day room being too small and not right for the job with columns blocking the view, and posters being way too cramped, the meeting was great.
- There is a commonality within SI2 that I think should be addressed. We are all creating software for other researchers to use. In a sense, we all have a shared customer: the academic research group. This customer base has aspects that we all individually address. Perhaps more focus on identifying and understanding these customers' needs would be helpful and broadly applicable?
- Add a discussion thread on financial sustainability of keeping software supported
- This was fantastic, and the format should be kept the same. Could to be good to have a longer "panel session" so that we have ample time for participants to provide feedbacks/inputs, etc
- Everything was great. kuddos.
- Location and room on the first day was great. Room on the second day not so much.
- Shorten the meeting to 1.5 days so those of us on the west coast can easily get flights back home in the afternoon of the last day.
- The short longer talks could be even shorter.
- Maybe reduce the time for longer talks (e.g., 30min to 20min + questions), and increase the time for the poster sessions (discussions were great).
- The meeting covers an extremely broad set of areas, and I felt that many of the presentations were far removed from my area of expertise. I wonder if it makes sense to cluster the presentations on related topics somehow.
- More time for posters
- Keep this structure.
- The room on the second day was too small...probably heard this 100 times already. Could organize poster sessions around some sort of themes - not sure what the themes should be - but randomly navigating 160 talks and posters is overwhelming - especially when everyone's work is so interesting. Perhaps organize poster session groups by age of projects?