⬅️ Connect with CASS
https://tinyurl.com/2024-CASS-BOFS

CASS Community BOF Days

The Consortium for the Advancement of Scientific Software

June 11 – 13, 2024

https://cass.community/bofs

Announcing CASS

The Consortium for the Advancement of Scientific Software

CASS Basics

- A newly-formed organization
- Sponsored by DOE Office of Advanced Scientific Computing Research (ASCR)
- Established by DOE Software Stewardship Organizations (SSOs)

CASS Goals

- Forum for SSO collaboration and coordination
- Bigger than the sum of its parts
- Vehicle for advancing the scientific software ecosystem

CASS Status

- Defining governance structure
- Establishing community awareness
- Building a team of teams
- Collaborating on outreach

Software Stewardship Organization (SSO) Basics

- Each SSO represents a specific software ecosystem concern
- **Product SSOs:** Programming systems, performance tools, math packages, data/viz packages
- **Portfolio SSO:** Curating & delivering software stack to the community
- **Community SSOs:** Workforce, partnerships

Engage with CASS

- Participate in June 11-13 CASS Community BOF Days: https://cass.community/bofs
- Visit https://cass.community

8 Software Stewardship Organizations (SSOs)

DOE Office of Advanced Scientific Computing Research (ASCR) Post-ECP Projects

COLABS

Training, workforce development, and building the RSE community

CORSA

Partnering with foundations to provide sustainable pathways for scientific software

FASTMATH

Stewardship, advancement, and integration for math and ML/AI packages

PESO

Stewarding, evolving and integrating a cohesive ecosystem for DOE software

RAPIDS

Stewardship, advancement, and integration for data and viz packages

S4PST

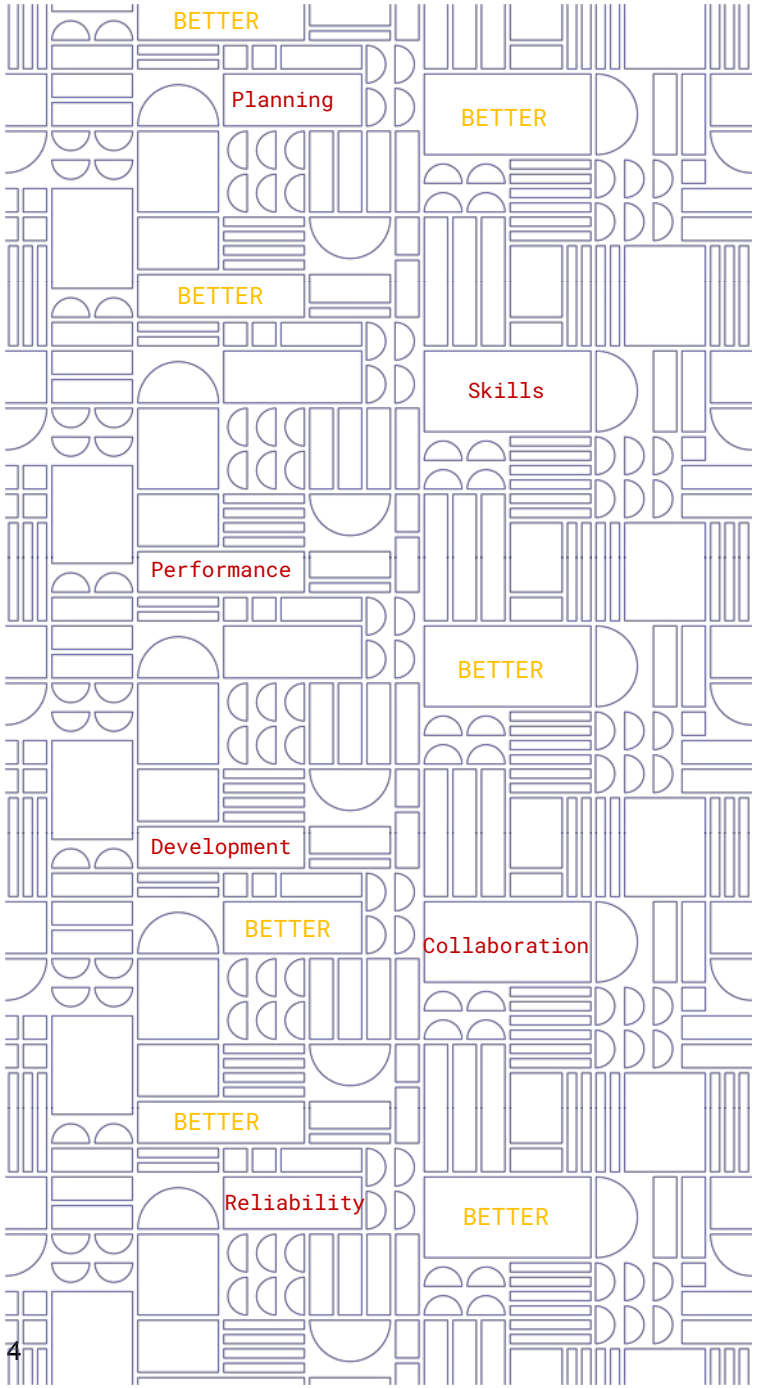
Stewardship, advancement and engagement for programming systems

STEP

Stewardship, advancement of software tools for understanding performance and behavior

SWAS

Stewardship and project support for scientific workflow software and its community



better scientific software

https://bssw.io

SO MY CODE WILL SEE THE FUTURE

CASS BOF Days

BSSw Fellowship BoF

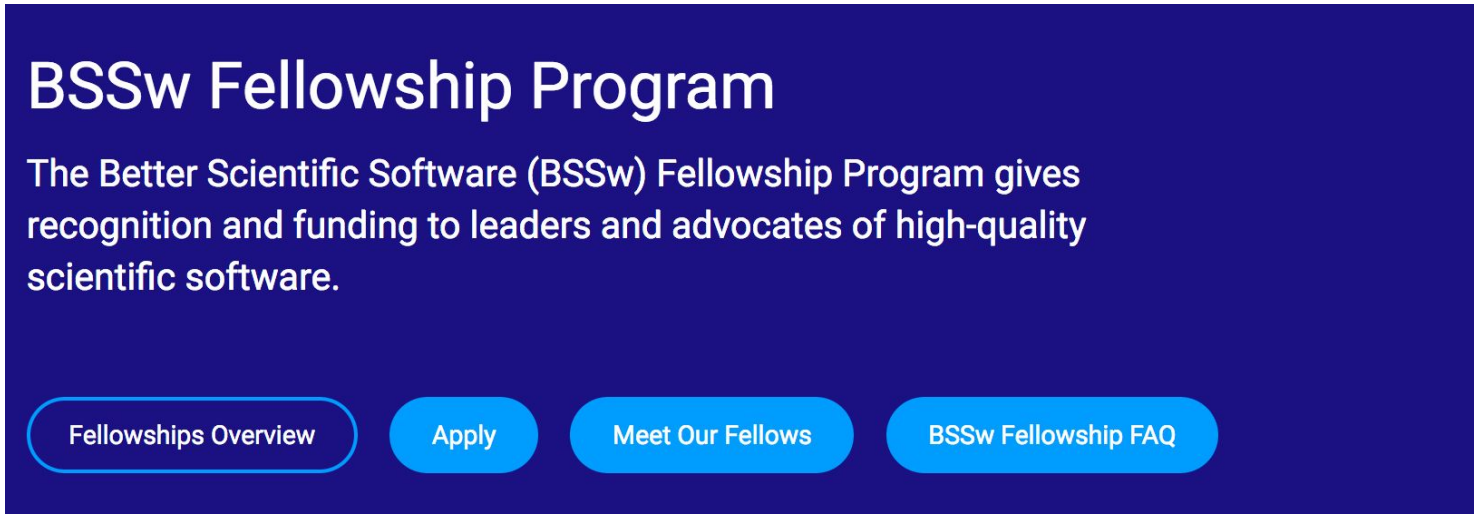
June 13, 2024

4

BSSw Fellowship Program

better scientific software

https://bssw.io/fellowship



ECP EXASCALE COMPUTING PROJECT

U.S. DEPARTMENT OF ENERGY

NSF

SUSTAINABLE HORIZONS INSTITUTE

GOAL:

Foster and promote practices, processes, and tools to improve developer productivity and software sustainability of scientific codes

AWARD:

We select at least three Fellows per year and honorable mentions as appropriate. Each 2024 BSSw Fellow will receive up to $25,000 for an activity that promotes better scientific software. Activities can include organizing a workshop, preparing a tutorial, or creating content to engage the scientific software community.

BSSw Fellowship: Meet the Fellows

better scientific software

RECOGNIZING LEADERS since 2018

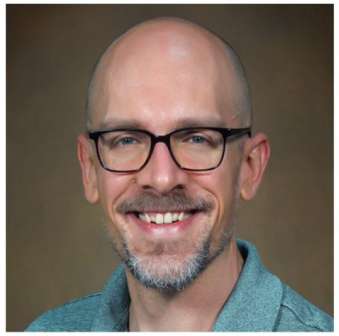
Meet Our Fellows

Meet the Fellows and Honorable Mentions and learn more about how they impact Better Scientific Software.

https://bssw.io/pages/meet-our-fellows

6

Meet the 2024 Fellowship Class



David Bunten

University of Colorado Anschutz Medical Campus

Software gardening almanac: Cultivating sustainable software development in the generative era

Dorota Jarecka

MIT McGovern Institute for Brain Research

Best practices for reproducibility and testing in scientific software

Olivia Newton

University of Central Florida

Team learning in scientific software projects

Ken Raffenetti

Argonne National Laboratory

Guidelines for improving MPI performance

Ryan Richard

Ames National Laboratory

Sustainable scientific software through multi-project CI/CD

Leah Wasser

pyOpenSci

Essential collaboration skills for contributing to open source software

Antigoni Georgiadou

Oak Ridge National Laboratory

Computational Scientist, National Center of Computational Sciences

Jack Marquez

University of Tennessee, Knoxville

Research Assistant Professor, Electrical Engineering and Computer Science

Drew Paine

Lawrence Berkeley National Laboratory

User Experience Researcher, Scientific Data Division

Noam Ross

EcoHealthAlliance

Principal Scientist, Computational Research

Matthew Scarpino

Purdue University

Lead Research Software Engineer, Rosen Center for Advanced Computing

Aristana Scourtas

Globus Labs, University of Chicago and Argonne National Laboratory

Project Manager and Research Software Engineer

7

2024 BSSw Fellow: **Dave Bunten**

- **Working on:** *Software Gardening Almanack*, handbook and tools for long-term sustainable software development.
- **Fascinated by:** bioinformatic software, information theory, software quality, in-memory data engineering, learning systems.
- **Why BSSw?** strong community and chance to innovate scientific software sustainability through knowledge and applied tooling.
- **Interest in ECP + DOE:** increasing scientific software quality and helping scale large data scientific discovery (esp. in the health domain).

Let's get in touch! @d33bs (LinkedIn, email, etc.)

Dorota Jareka

MIT McGovern Institute
for Brain Research

BSSw Project: Best practices for reproducibility and testing in scientific software.

9

Team Learning for Better Scientific Software

Olivia B. Newton, Ph.D.
Postdoctoral Scholar
Institute for Simulation and Training
University of Central Florida

Software development and interdisciplinary collaboration in scientific projects necessitate integrative learning

- ➔ Critical for performance outcomes but challenging given the breadth of expertise that must be synthesized to effectively build knowledge

2-day workshop focused on developing a set of best practices for team learning in scientific software development

- ➔ Produce open knowledge resource and toolkit in GH repository for understanding, evaluating, and improving team learning in scientific software projects
- ➔ Disseminate products and findings via bssw.io and publications

More info + express interest in participating: https://tinyurl.com/tl4bssw
Contact me: olivia.newton@ucf.edu

GUIDELINES FOR GETTING BETTER MPI PERFORMANCE

Ken Raffenetti – Research Software Engineer

Challenge

- MPI is a large, complex programming standard. Key features and usage patterns not always obvious to users.
- Documentation for configuration and settings often implementation-specific

Proposal

- Website covering MPI performance topics commonly encountered by users
  - — Informational pages with code examples
  - — Video walkthroughs
  - — Implementation agnostic

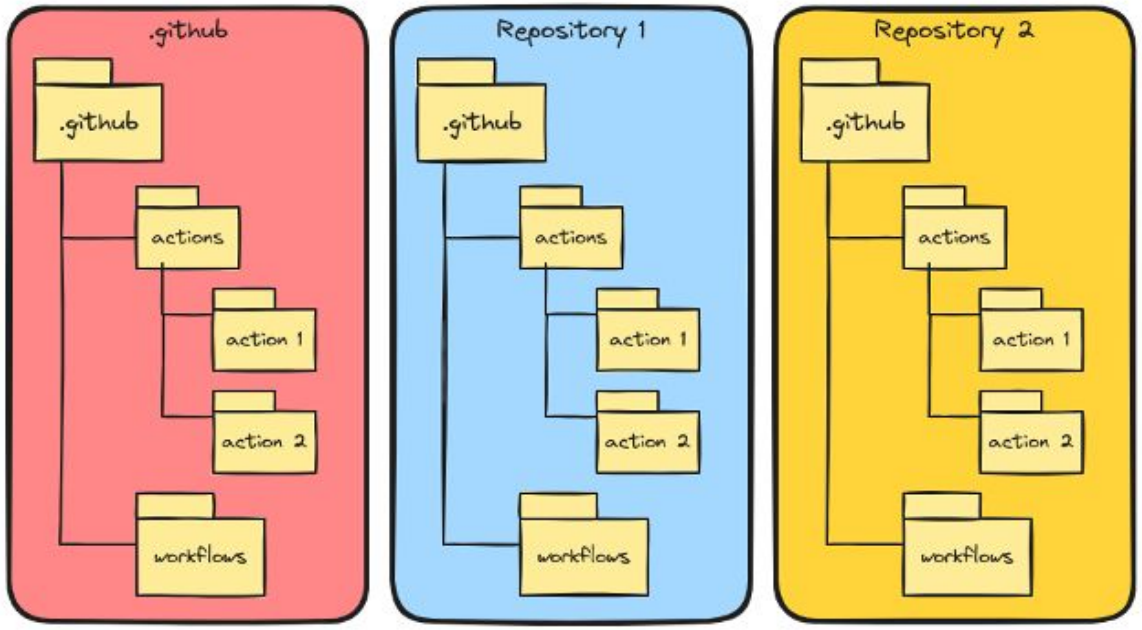
Focus Areas

- Better Performance
  - High-Performance Computing
  - Performance Portability
  - Performance at Leadership Computing Facilities
- Better Planning
  - Design
  - Software Interoperability
- Better Development
  - Documentation
  - Configuration and Builds

Sustainable software through multi-project CI/CD

Ryan Richard

- Scientists are increasingly developing multiple projects.
- Projects usually have similar continuous integration/continuous deployment needs.
- **Challenge**: How do we avoid duplicating CI/CD infrastructure?
- **Research**:
- Workshops to assess state of the art practices.
- Create tutorials based on emerging best practices.
- Website: https://multiprojectdevops.github.io/



Leah Wasser

pyOpenSci

BSSw Project: Essential collaboration skills for contributing to open source software.

13

Antigoni Georgiadou, PhD

- Computational Scientist in Mathematics at the Oak Ridge National Laboratory& a member of the Algorithms & Performance Analysis Group in the Science Engagement Section.
- Part of the FLASH-X astrophysics collaboration and a partner with the HACC team to perform cosmology simulations.
- Been serving as the OLCF liaison for the XGC fusion collaboration.
- Committee member of INCITS for Fortran and the parent Programming Languages (PL22). The technical committee works on the interpretations of the current standard for Fortran and develops the list of new features for the following revisions.
- BSSw Concept
  - As power consumption of data centers is becoming a hot topic given the growth of AI and exascale systems are requiring 30+ MW, enough to power sizable towns, there is a universal effort towards green software development with academia, government, and industry spearheading the discussion on how software developers will produce better software in terms of more energy efficient approaches. I am bound to create and hosting a series of videos in Youtube to encourage the community to lead in energy efficiency.
- Scope
  - My goals include raising awareness of good software practices, improve scientific productivity, and inform for ongoing efforts for sustainability planning.

Jack Marquez

University of Tennessee, Knoxville

Research Assistant Professor, Electrical Engineering and Computer Science

Honorable Mention

15

Drew Paine, PhD

My focus is to help improve the state of practice for the user experience (UX) of scientific software tools, systems, and facilities

Current Activities

STRUDEL

https://strudel.science
Building tools to enable more usable scientific software

HPDF
High Performance Data Facility

https://hpdf.science
User experience & engagement for ASCR's newest facility project

US RSE

https://us-rse.org/wg/ux/
Co-chairing UX Working Group fostering community of practice for UX in science

Background

PhD, Human Centered Design & Engineering, University of Washington
BS, Software Engineering, Rose-Hulman Institute of Technology

Professional experience at LBNL & Google Cloud

Lawrence Berkeley National Laboratory

Honorable Mention

Reach out! pained@lbl.gov

16

Noam Ross

EcoHealthAlliance

Principal Scientist,

Computational Research

Honorable Mention

17

Matthew Scarpino

Purdue University

Lead Research Software Engineer, Rosen Center for Advanced Computing

Honorable Mention

18

Aristana Scourtas

Globus Labs, University of Chicago and Argonne National Laboratory

Project Manager and Research Software Engineer

Honorable Mention

19

Meet the 2023 Fellowship Class

Nicole Brewer

Arizona State University

Improving accessibility of data and software with scientific web apps

Myra Cohen

Iowa State University

Techniques for scientific software testing

Johannes Doerfert

Lawrence Livermore National Laboratory

Demystifying the compiler black box

William Hart

Sandia National Laboratories

Best practices for software supply chain security

Helen Kershaw

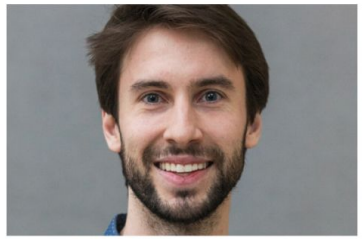
National Center for Atmospheric Research

Improving code review skills for scientific software developers

Rafael Mudafort

National Renewable Energy Laboratory

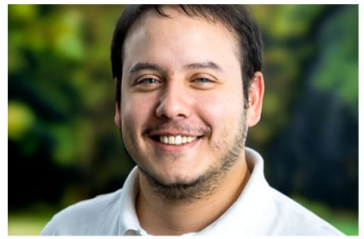
Effective communication of software design



Jean Luca Bez

Lawrence Berkeley National Laboratory

Scientific Data Division, Postdoctoral Researcher



Jose Monsalve Diaz

Argonne National Laboratory

Postdoctoral Researcher, Mathematics & Computer Science Division

Xu Liu

North Carolina State University

Associate Professor, Computer Science

Alisa Neeman

Muskingum University

Assistant Professor, Mathematics and Computer Science

Kristina Riemer

University of Arizona

Scientific Programmer, Data Science Institute

Brigitta Sipőcz

California Institute of Technology

Applications Developer

20

Nicole Brewer

Arizona State University

BSSw Project: Improving accessibility of data and software with scientific web apps

21

Myra Cohen

Iowa State University

BSSw Project: Techniques for scientific software testing

22

Johannes Doerfert

Lawrence Livermore National Laboratory

BSSw Project: Demystifying the compiler black box

23

William Hart

Sandia National Laboratory

BSSw Project: Best practices for software supply chain security

24

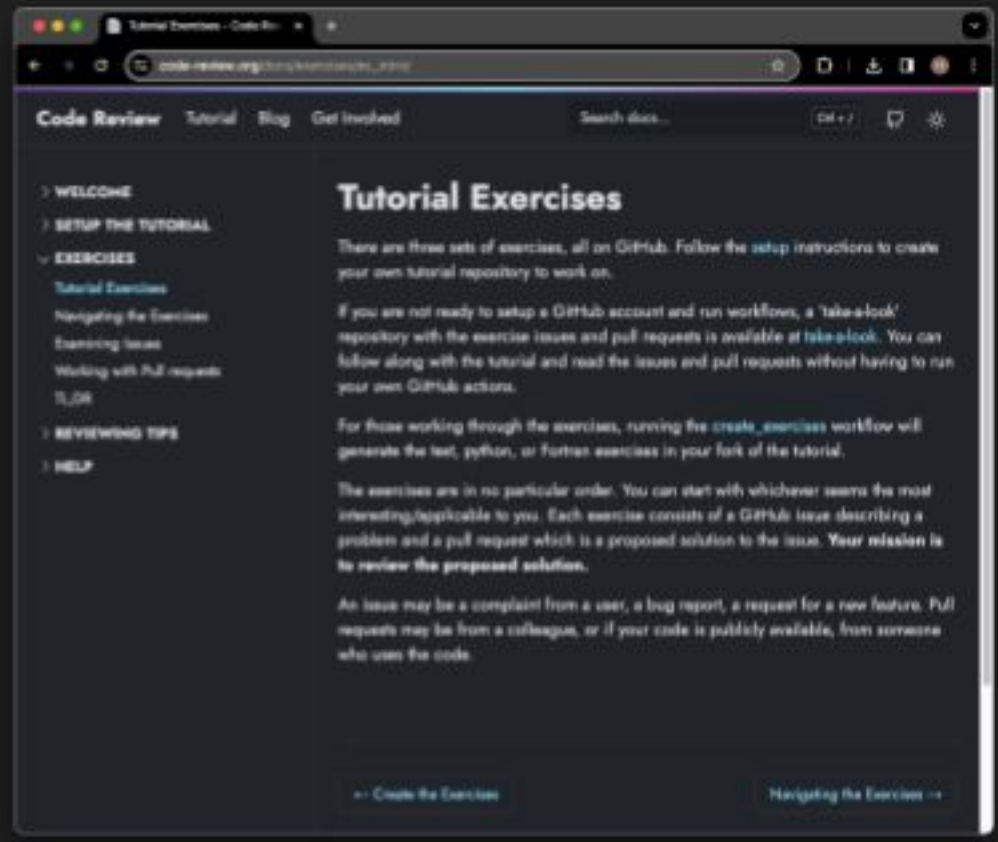
What I did: Code Review

An online tutorial for code review
- Hands on exercises & website
- “No code” examples

What I’m interested in:

How you write code with other people
Getting better at reviewing and being reviewed
The joy and pain of Open Source Software
Reading more code than I write!

Helen Kershaw
Data Assimilation Research Section
NSF NCAR
hkershaw@ucar.edu
USRSE slack
BSSw slack



code-review.org

https://github.com/scientific-software-reviewers/tutorial

“Increasing developer impact through effective communication of software design”

Deliverables:

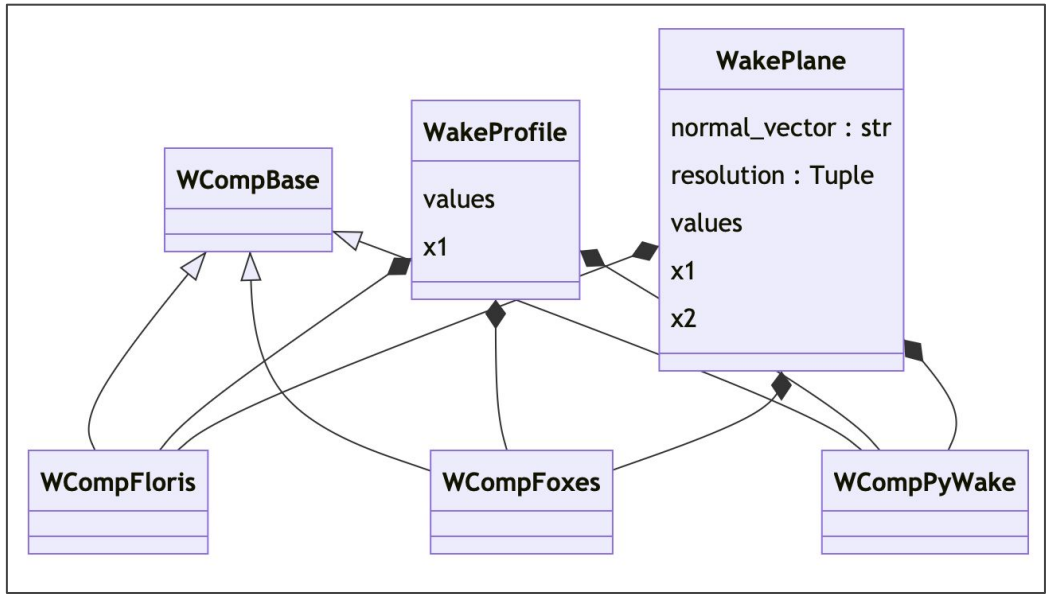
- Software diagramming resources in an online dashboard
- BSSw blog post
- Conference presentation
- IDEAS HPC Best Practices Webinar

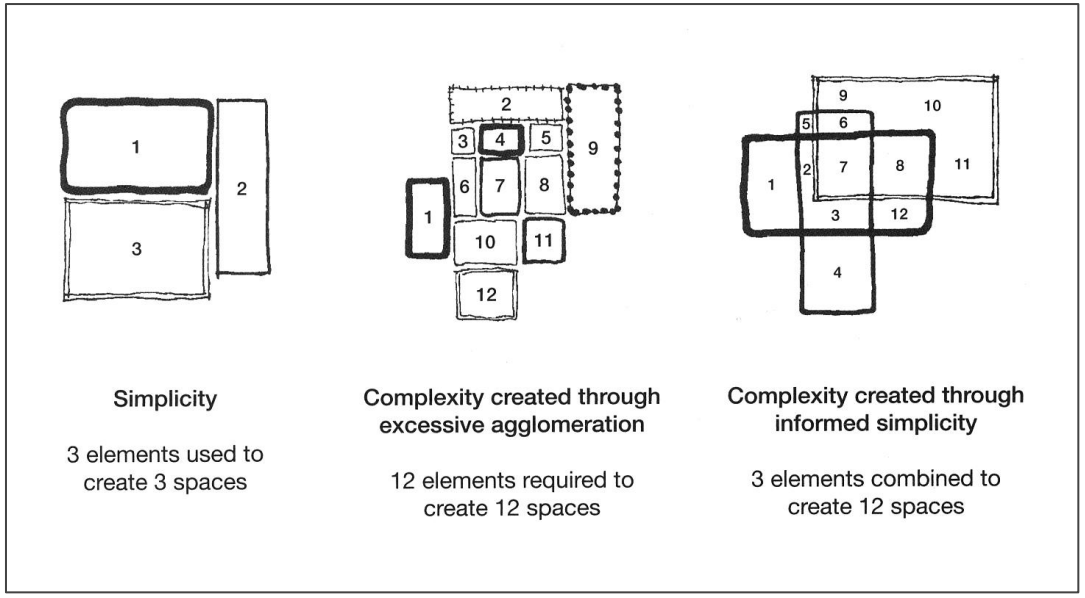
Rafael Mudafort

rafael.mudafort@nrel.gov

@rafmudaf

NREL Transforming ENERGY





101 Things I Learned in Architecture School

Jean Luca Bez

Lawrence Berkeley National Laboratory
Postdoctoral Researcher,
Scientific Data Division

Honorable Mention

27

Jose Monsalve Diaz

Argonne National Laboratory
Postdoctoral Researcher,
Mathematics & Computer
Science Division

Honorable Mention

28

Xu Liu

North Carolina State University
Associate Professor, Computer Science

Honorable Mention

29

Alisa Neeman

Muskingum University
Assistant Professor,
Mathematics and Computer Science

Honorable Mention

30

Kristina Riemer

University of Arizona
Scientific Programmer,
Data Science Institute

Honorable Mention

31

Brigitta Sipöcz

California Institute of Technology
Applications Developer

Honorable Mention

32

2022

Ritu Arora

University of Texas at San Antonio

Optimizing I/O for better performance

Rob Latham

Argonne National Laboratory

I/O sleuthing: a tour of I/O challenges and solutions

Julia Stewart Lowndes

National Center for Ecological Analysis and Synthesis (NCEAS), UC Santa Barbara

Openscapes: Open data

Amiya K. Maji

Purdue University

Simplifying scientific Python package management

Nitin Sukhija

Slippery Rock University of Pennsylvania

Secure scientific software development

Karan Vahi

USC Information Sciences Institute

Scientific workflows for high efficiency HPC

Sarah Bratt

Syracuse University

Ph.D. Student, School of Information Studies

William Godoy

Oak Ridge National Laboratory

Computer Scientist

Brittany Johnson-Matthews

George Mason University

Assistant Professor, Computer Science Department

Max Jones

University of Hawai'i at Mānoa

Postdoctoral Researcher, Department of Earth Sciences, School of Ocean & Earth Science & Technology

Rafael Mudafort

National Renewable Energy Laboratory

Research Software Engineer, National Wind Technology Center

Qiusheng Wu

University of Tennessee, Knoxville

Assistant Professor, Department of Geography

33

2022 BSSw Fellow - RITU ARORA

University of Texas at San Antonio
Assistant Vice President, Research Computing

Email: ritua.arora@utsa.edu, ritua2@yahoo.com

WHAT ATTRACTED ME TO THE FELLOWSHIP?

Opportunity to network with other colleagues in the community who care about developing better scientific software

Bringing visibility to the topic of optimizing I/O

BSSw FOCUS

OPTIMIZING I/O FOR BETTER PERFORMANCE

Creating videos, articles/blogs, and examples/exercises to demonstrate how to optimize I/O in scientific applications including the area of AI/machine learning

Developing applications that support "checkpointing and restart" mechanism also involves careful consideration of doing optimal I/O, and this topic will also be covered in the articles/blogs/exercises

CONTINUED ENGAGEMENT

Would love to find collaborators to partner on future work related to fault-tolerance and optimizing I/O

Rob Latham

I/O Sleuthing

- Long history of I/O tutorials at SC, ATPESC
  - MPI-IO, I/O libraries
- Time for something a bit more “in the trenches”
  - A collection of hands-on examples
  - A framework for collecting and evaluating benchmark data
- BSSW fellowship funds a “first draft”
  - Code, materials will be available for ongoing refinement, community contributions

Hi, I'm a marine ecologist and Openscapes founder

Julia Stewart Lowndes, PhD

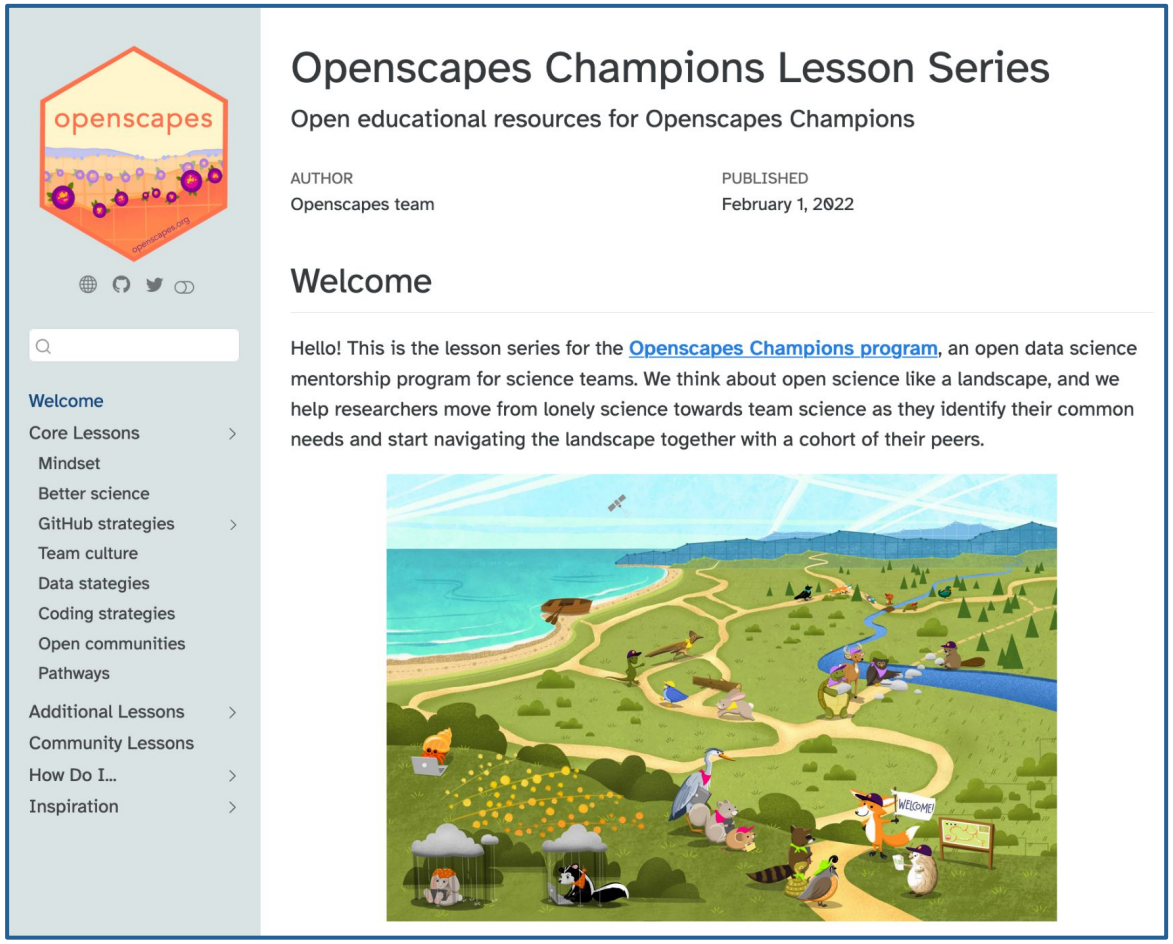
Senior Fellow, NCEAS, University of California Santa Barbara
Open Science Community Member, entryway through R
Mozilla Fellow, Open Science
Better Scientific Software Fellow, NSF
lowndes@nceas.ucsb.edu; @juliesquid

Openscapes helps research teams transition to inclusive open data science workflows

We mentor teams to better tackle their questions by strengthening shared practices, underpinned by existing tools. Supercharge your research (Lowndes et al. 2019); Open software means kinder science (Lowndes 2019); Our path to better science in less time using open data science tools (Lowndes et al. 2017)

During the fellowship

I improved open educational resources for our Champions program in part by upgrading to Quarto, led the program >five times with government staff and academic teams, and keynoted the global launch of Quarto at the RStudio conference (slides,



https://openscapes.org/series

openscapes.org
nasa-openscapes.github.io
nmfs-openscapes.github.io

36

Simplifying scientific Python package installation and environment management

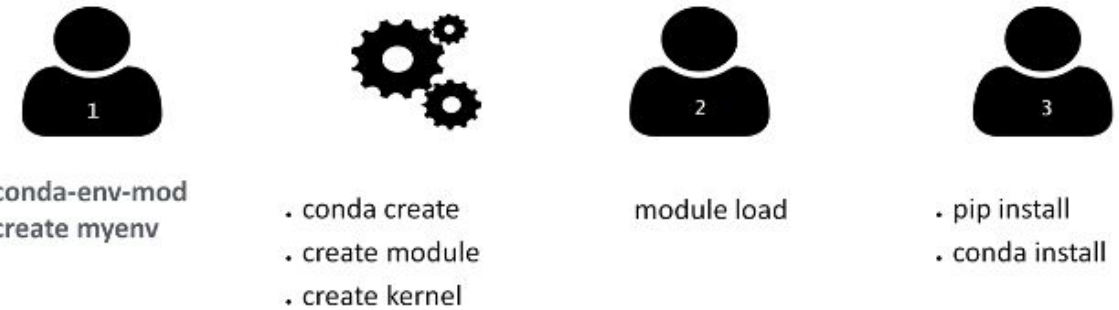
Amiya K Maji
amaji@purdue.edu

Motivation

- Installing and sharing Python packages is a difficult task on HPC systems
- How can we simplify environment management and reduce errors

Goals

- Document best practices
  - https://hpc-python-solutions.readthedocs.io/
- Simplify installation through automation and scripting
  - conda-env-mod
  - https://github.com/amaji/conda-env-mod



Solution approach

- Codify best practices using scripts
- Simplify environment management using module files

Future goal

- HPC Python forum

PURDUE UNIVERSITY

Securing Scientific Software Development

Mitigating the risk of software vulnerabilities with best practices and tools for secure scientific software development

Tutorial on Best Practices and Tools For Secure Scientific Software Development

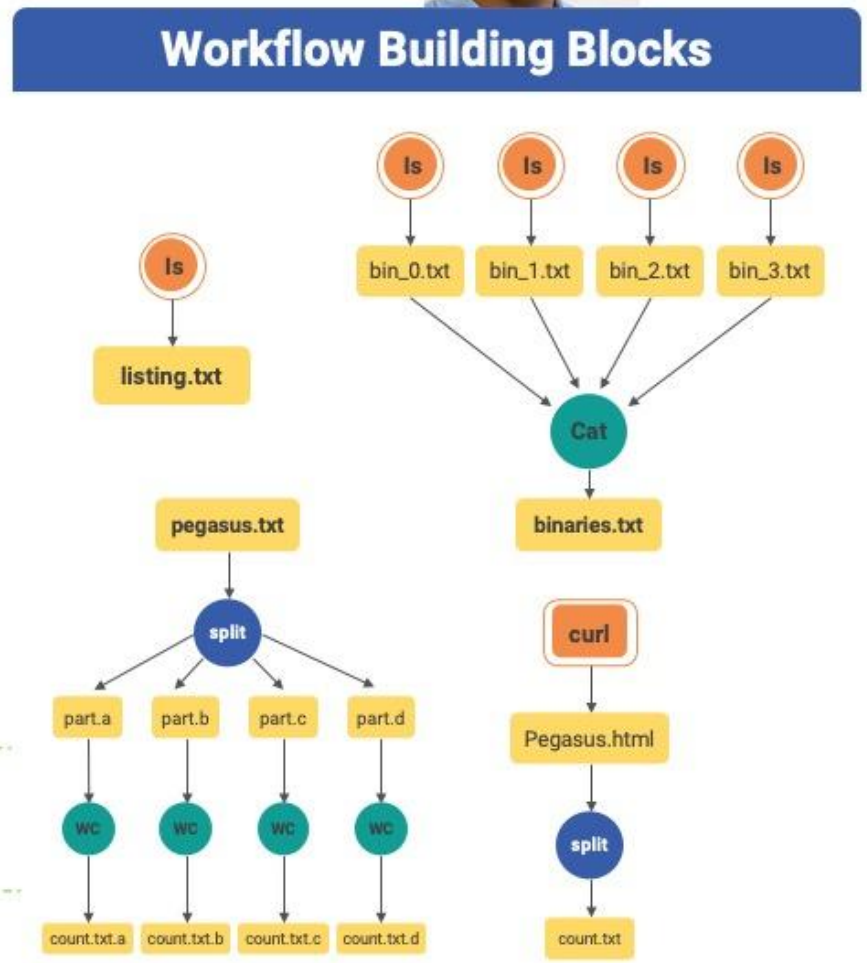
SIAM MDS 2022

Nitin Sukhija

Research Areas:
Resilience, Cybersecurity and High Performance Data Analytics

Enabling complex scientific computations and efficient use of HPC resources with scientific workflows

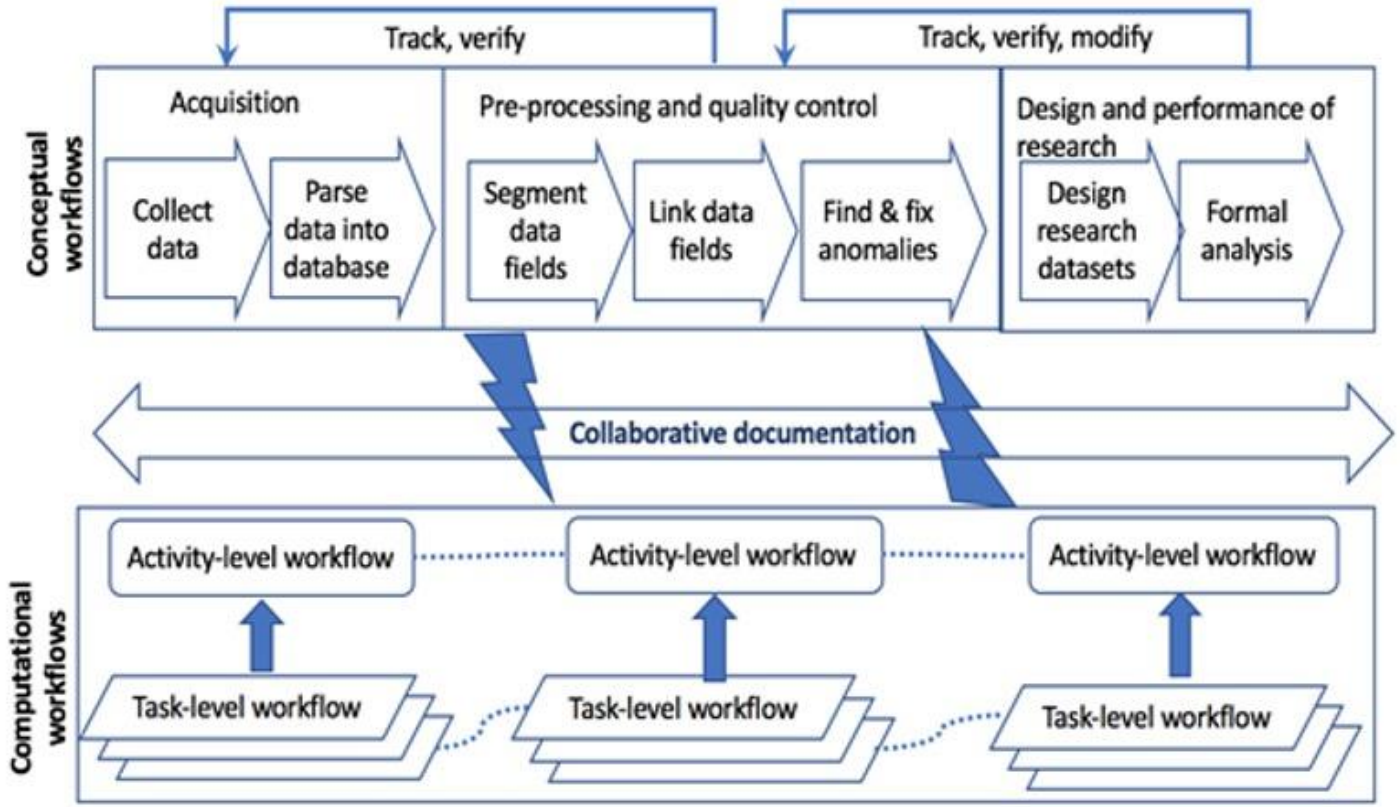
- **Scientific Workflows**
  - **Conducts a series of computational tasks.**
    - Resources distributed across Internet.
  - **Chaining (outputs become inputs) replaces manual hand-offs.**
    - Accelerated creation of products.
- **Fellowship Goals**
  - **Develop training materials**
    - Focus on workflow lifecycle and challenges
    - Easy to understand
    - Jupyter notebooks that leverage Pegasus WMS to submit resources such as OSG, HPC centers
    - Package codes into application containers and deploy them in workflows
  - **Target Audience**
    - Domain scientists that currently use or are considering using HPC
  - **Dissemination**
    - Tutorials conducted at various confereces such as PEARC22 (accepted) and EScience22 (proposal submitted)
    - Seminar at NERSC followed by development of NERSC specific training materials



Karan Vahi

USC Viterbi School of Engineering

Workflow Model for Big Metadata



- **Conceptual workflows** define the goals and strategies of activities
- **Computational workflows** codify the conceptual workflows for machine execution
- **Future work:** Codifying **paradata**

Bratt, S., Hemsley, J., Qin, J., & Costa, M. (2017). Big data, big metadata and quantitative study of science: A workflow model for big scientometrics. *Proceedings of the Association for Information Science and Technology*, *54*(1), 36-45.

Qin, J., Hemsley, J., & Bratt, S. The Structural Shift and Collaboration Capacity in GenBank Networks: A Longitudinal Study. *Quantitative Science Studies.* (*2022*)

Syracuse University
School of Information Studies

THE UNIVERSITY OF ARIZONA.

Sarah Bratt
PhD Candidate
G' May 2022

William Godoy

Oak Ridge National Laboratory

Honorable Mention

41

Brittany Johnson-Matthews

George Mason University

Assistant Professor,
Computer Science
Department

Honorable Mention

42

Max Jones

University of Hawa’i at Mānoa

Postdoctoral Researcher,

Department of Earth Sciences, School of Ocean & Earth Science & Technology

Honorable Mention

43

Qiusheng Wu

qwu18@utk.edu

- Profiles: Google Scholar | GitHub | YouTube | Twitter | LinkedIn
- Expertise: Geospatial Data Science, Open Source Software Development
- Projects: geemap | leafmap | lidar | whitebox | whiteboxR | whiteboxgui | geospatial | geospatial-ml | pypackage

Assistant Professor

University of Tennessee, Knoxville

44

2021

better scientific software

https://bssw.io

SO MY CODE WILL SEE THE FUTURE

2021 Class

Fellows

Marisol García-Reyes
Farallon Institute

Increasing accessibility of data & cloud technologies

Mary Ann Leung
Sustainable Horizons Institute

Increasing developer productivity and innovation through diversity

Chase Million
Million Concepts

Project management best practices for research software

Amy Roberts
University of Colorado Denver

Enabling collaboration through version control user stories

Honorable Mentions

Keith Beattie
Lawrence Berkeley National Laboratory

Computational Research Division, Computer Systems Engineer

Julia Stewart Lowndes
National Center for Ecological Analysis and Synthesis (NCEAS), UC Santa Barbara

Openscapes Director

Jonathan Madsen
Lawrence Berkeley National Laboratory

NERSC, Application Performance Specialist

Addi Thakur Malviya
Oak Ridge National Laboratory

Software Engineering Group, Group Leader

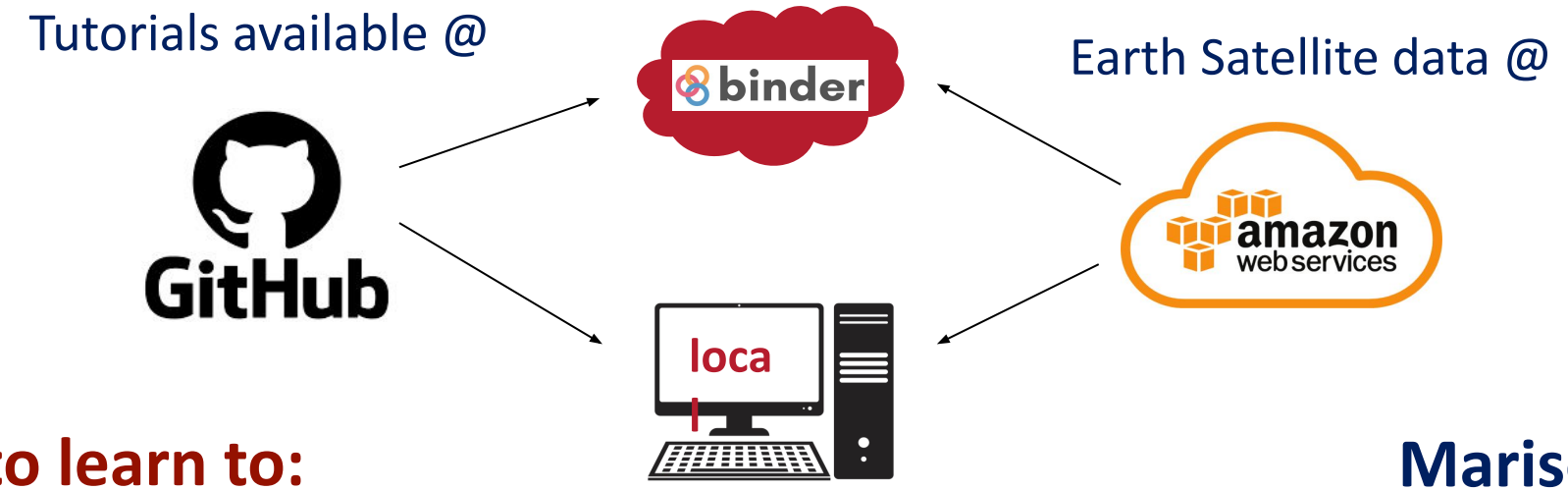
45

github.com/marisolgr/python_sat_tutorials

Tutorial: Timeseries of Satellite Data using Python

launch binder

Tutorial to learn how to access and analyze time series of Satellite and Satellite-based Data using Python and JupyterLab in the Cloud



Code to learn to:

python & xarray
Acquire data from cloud
Analyze timeseries

FARALLON INSTITUTE

https://bssw.io
better scientific software

IMPACT
Interagency Implementation and Advanced Concepts Team

Marisol García-Reyes, Ph.D.
Oceanographer
marisolgr@faralloninstitute.org
github: marisolgr

Mary Ann Leung, Sustainable Horizons Institute
mleung@shinstitute.org

Increasing developer productivity and innovation through diversity and normalizing inclusion

Motivation

- Diversity increases innovation
- Critical workforce challenges
- Normalizing inclusion → Sustainable workforce

A Few Current Projects

- Consortium for Advanced Scientific Software (CASS) – Broadening Participation Initiative
- Sustainable Research Pathways
- Broader Engagement Program
- BSSw Fellowship Administration
- Intro to HPC through Energy Justice Projects

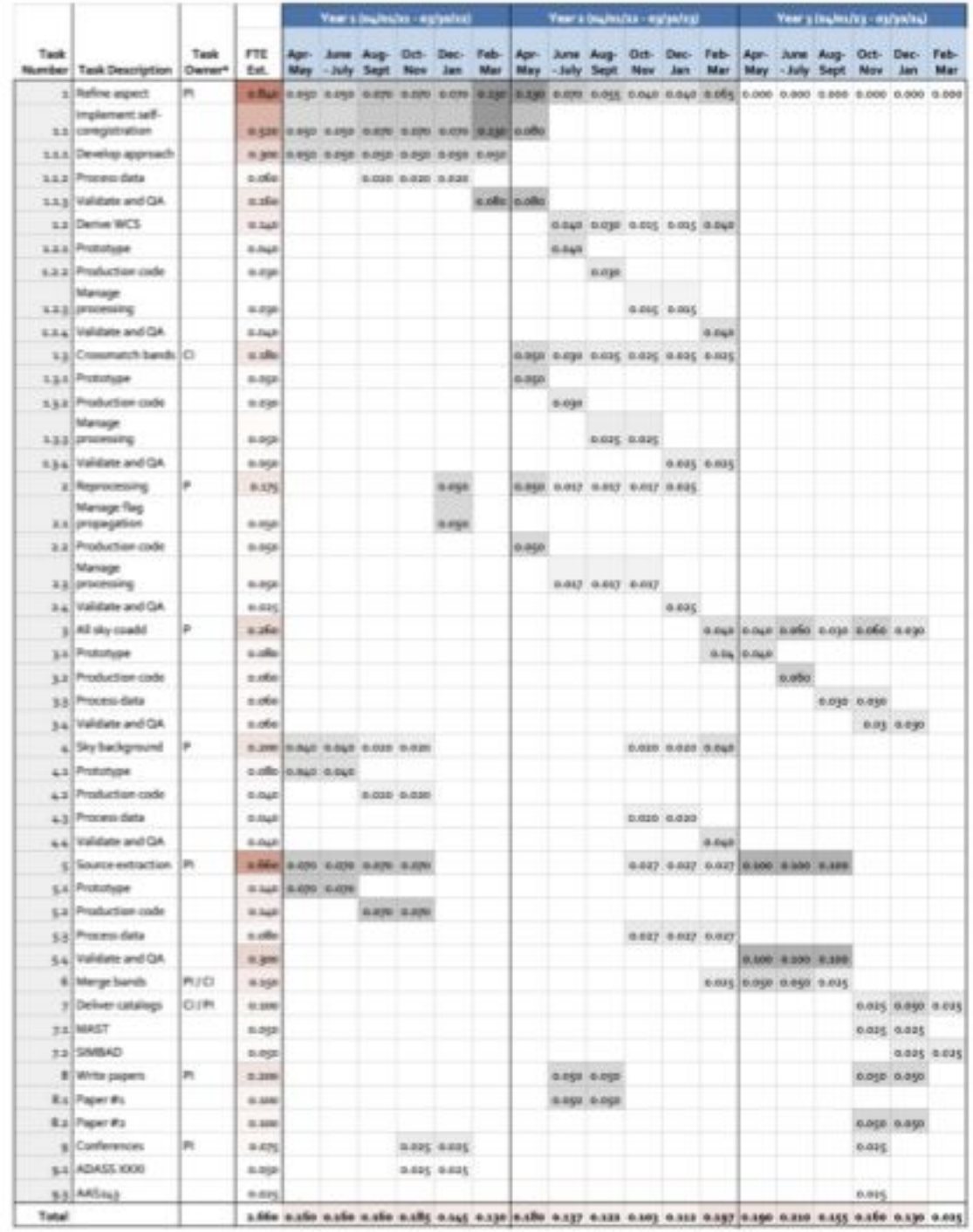
"A Practical Guide to Research Software Project Estimation"

https://github.com/MillionConcepts/software_project_management

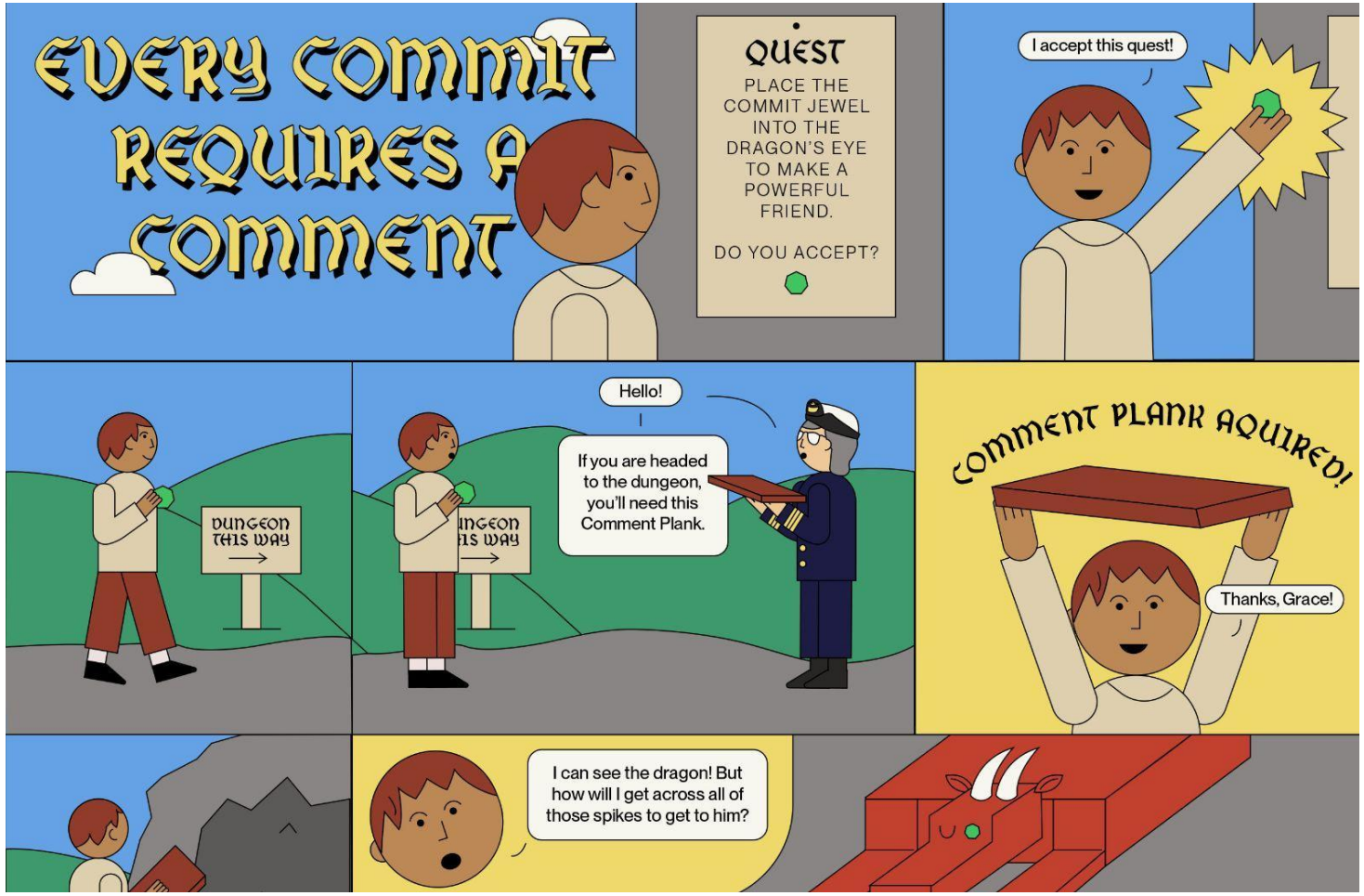
Chase Million
2021 BSSw Fellow

10^6C Million Concepts

The pursuit of an accurate estimate for research software development is foolhardy. Accurate estimates are extremely difficult even for expert estimators in fields like business development that often produce mountains of historical data on which to base their calculations. But estimating the envelope of viability — the minimum necessary time and resources on the low end, and those necessary to have reasonable confidence in project success on the high end — is fortunately both useful and somewhat straightforward. This book teaches a process for estimating such envelopes and how to use them as the bases for software research project work plans or budgets that can be incorporated into proposals.



Story-based git training for new developers



- Looking for help with testing, material development, and sharing!
- [git-stories · GitLab]
- amy.roberts@ucdenver.edu

Keith Beattie

- **Background:** BA in Math, MS in Comp Sci, ~20 years at LBL as a CSE (Computer Systems Engineer). Worked in industry for 5 years, returned to the lab to work on science
- **Axe to Grind:** For us to have sustainable scientific software, we need to support sustainable scientific software development careers.
- **I've worked on some great projects at LBL**
  - IceCube Neutrino detector - DAQ development: got to go to the south pole a few times
  - Helped Chem-Es develop carbon capture models, a process systems engineering framework and now water recovery systems with reverse osmosis, desalination and full plant models.
  - LuxZepplin Dark Matter Experiment: data movement for physicists
  - Lots of being the lone, or one a just a few, software developers in a team of scientists of other disciplines

I now want to be part of the RSE movement recognizing the unique and powerful role software engineers (and the like) play in contributing to scientific advancement.

Lawrence Berkeley National Laboratory

50

Jonathan Madsen

Lawrence Berkeley National Laboratory

Application Performance Specialist

NERSC

Honorable Mention

51

Addi Malviya-Thakur
malviyaa@ornl.gov
Group Leader,
Software Engineering,

Advanced Computing Systems Research Section,

Computer Science and Mathematics Division,

Oak Ridge National Laboratory

Software engineering, software architecture, data architecture, design patterns, test-driven processes for the development and maintenance of scientific software applications.

Automated and scalable machine learning and artificial intelligence engineering for data-intensive applications

Software quality assurance, integration, validation, and verification for high-performance computing applications.

Incorporating reliability and security in software life cycle management & maintenance.

Human-Computer Interaction and usability research for emerging context-aware interactive computing platforms.

Platforms for Inter-connected Science to enable the future of science by connecting edge instruments to other edge instruments and centralized computing resources.

2020

better scientific software

https://bssw.io

SO MY CODE WILL SEE THE FUTURE

2020 Class

Fellows

Nasir Eisty

University of Alabama

Automating testing in scientific software

Damian Rouson

Sustainable Horizons Institute, Sourcery Institute

Introducing agile scientific software development to underrepresented groups

Cindy Rubio-Gonzalez

University of California, Davis

Improving the reliability and performance of numerical software

Honorable Mentions

David Boehme

Lawrence Livermore National Laboratory

Research Staff, Center for Applied Scientific Computing

Sumana Harihareswara

Changeset Consulting

Founder and Principal, Open source software management and collaboration

David Rogers

National Center for Computational Sciences, Oak Ridge National Lab

Computational Scientist

53

Nasir Eisty

Assistant Professor
Department of Computer Science
Boise State University
Boise, ID, USA
nasireisty@boisestate.edu

My role:
I lead the Software Engineering & Analytics Lab (SEAL) at Boise State University. I teach Software Engineering courses and conduct software engineering research.

BOISE STATE UNIVERSITY
COLLEGE OF ENGINEERING
Department of Computer Science

I'm interested in:
Software Engineering research for Research Software.

A problem I'm grappling with:
How to build high-quality research software through rigorous testing and other software quality practices.

My background:
My Ph.D. dissertation was about software quality assurance practices in Research Software.

I would like to see:
Long-term collaborations between the BSSw fellows and the ECP team to make great contributions to the community.

Damian Rouson

Sustainable Horizons Institute, Sourcery Institute

BSSw Project: Introducing agile scientific software development to underrepresented groups

55

Cindy Rubio-Gonzalez

University of California Davis

BSSw Project: Improving the reliability and performance of numerical software

56

David Boehme

Lawrence Livermore National Laboratory

Honorable Mention

57

Sumana Harihareswara
Honorable Mention

- Working on fixing the leadership gap for legacy projects with a book, workshops, and coaching
- Seeking reviewer readers for early chapters & participants for trainings
- Email sh@changeset.nyc or follow @brainwane@social.coop on Mastodon

Changeset Consulting
New York City

David Rogers

Advanced Computing for Chemistry and Materials

Oak Ridge National Laboratory

David M. Rogers is a Computational Scientist in the Advanced Computing for Chemistry and Materials at ORNL's National Center for Computational Science. He obtained his Ph.D. in Physical Chemistry from University of Cincinnati in 2009 where he worked on applying Bayes' theorem to the free energy problem with applications to the statistical mechanics of liquids. After a Post-doctoral fellowship at Sandia National Labs 2009-2012, he was a faculty member of the USF Department of Chemistry 2013-2019. He has worked on theory and application of statistical mechanical methods in liquids, biomolecules, and quantum models.

His current research interests include computation, mathematics, and theory enabling HPC methods for electron- to fluid-scale modeling. Especially interesting open problems include applications to small nonequilibrium systems, dielectric friction, dispersion, hydration, and its role in nanoscale devices. David's focus at ORNL is gathering and communicating community needs for developing more powerful and general libraries and interfaces for modeling these systems at scale.

59

2019

better scientific software

https://bssw.io

SO MY CODE WILL SEE THE FUTURE

2019 Class

Fellows

Rene Gassmoeller

University of California, Davis

Guiding your scientific software project from inception to long-term sustaintability

Ignacio Laguna

Lawrence Livermore National Laboratory

Improving the reliability of scientific applications by analyzing and debugging floating-point software

Tanu Malik

DePaul University

Reducing technical debt in scientific software through reproducible containers

Kyle Niemeyer

Oregon State University

Educating scientists on best practices for developing research software

Honorable Mentions

Stephen Andrews

Los Alamos National Laboratory

Staff Scientist, XCP-8: Verification and Analysis

Nasir Eisty

University of Alabama

Ph.D. Student, Computer Science

Benjamin Pritchard

Virginia Tech

Software Scientist, Molecular Sciences Software Institute

Vanessa Sochat

Stanford University

Research Software Engineer, Stanford Research Computing Center

Rene Gassmoeller

University of California, Davis

BSSw Project: Guiding your scientific software project from inception to long-term sustaintability

61

Ignacio Laguna

Lawrence Livermore National Laboratory

BSSw Project: Improving the reliability of scientific applications by analyzing and debugging floating-point software

62

2019 BSSW Fellowship

Reducing Technical Debt with Reproducible Containers

Tanu Malik
Associate Professor
School of Computing
Director, DICE Lab
DePaul University, Chicago IL
w: https://dice.cs.depaul.edu
e: tanu.malik@depaul.edu

Acknowledgements: NSF CNS-1846418, ICER-1639759, ICER-1661918, NASA AIST 00021-95

BSSw Fellowship, DePaul Seed Grants

Sciunit (https://sciunit.run): Creates reproducible containers. Is based on application containerization and provenance

- **Code debt:** Arises when unnecessary code and libraries are included
  - Sciunit[1,2,3,4] prunes away unnecessary files not used by an application

Addressed 3 types of technical debt for scientific computing and ML workloads:

- **Environment debt**: Arises when environments are incompatible.
  - Sciunit[5,6,7] simplifies environment translation and exports environments.
- **Testing debt**: Arises when software has no compatible testing suite.
  - Sciunit[8,9] makes it easier to test programs on multiple inputs and reports differences.
- **Evaluation debt**: Arises when multiple versions of a software are run. Typical during Artifact Evaluation[10].
  - Sciunit[11] with a fixed cache size makes if 50% faster to evaluate multiple versions of a software
- **Data debt**: When containers must download more data than required.

Kyle Niemeyer

Oregon State University

BSSw Project: Educating scientists on best practices for developing research software

64

Stephen Andrews

Los Alamos National Laboratory

Staff Scientist,

XCP-8: Verification and Analysis

Honorable Mention

65

Benjamin Pritchard

Virginia Tech

Software Scientist,

Molecular Sciences Software Institute

Honorable Mention

66

Vanessa Sochat

Lawrence Livermore National Laboratory

Honorable Mention

67

2018

better scientific software

https://bssw.io

SO MY CODE WILL SEE THE FUTURE

2018 Class

Fellows

Jeffrey Carver

University of Alabama

Improving code quality through modern peer code review

Ivo Jimenez

University of California, Santa Cruz

Enabling reproducible research through automated computational experimentation

Daniel S. Katz

University of Illinois at Urbana-Champaign, National Center for Supercomputing Applications

Giving software developers long-overdue credit through principles for software citation

Andrew Lumsdaine

Pacific Northwest National Laboratory, University of Washington, Northwest Institute for Advanced Computing

Guiding efficient use of modern C++ for high-performance computing

Honorable Mentions

Neal Davis

University of Illinois at Urbana-Champaign

Teaching Assistant Professor, Computer Science

Marc Henry de Frahan

National Renewable Energy Laboratory

Postdoctoral Researcher

Elsa Gonsiorowski

Lawrence Livermore National Laboratory

HPC I/O Specialist, Livermore Computing

Ying Li

Argonne National Laboratory

Argonne Scholar, Argonne Leadership Computing Facility

68

Jeff Carver

carver@cs.ua.edu

BSSw Project: Contemporary Peer Code Review for Research Software

- Developed an interactive tutorial on peer code review
- Interested in giving the tutorial to various audiences
- Interested in interacting directly with teams on peer code review

University of Alabama

69

Ivo Jimenez

University of California, Santa Cruz

BSSw Project: Enabling reproducible research through automated computational experimentation

70

Daniel S. Katz
Chief Scientist, NCSA
Research Associate Professor, CS
Research Associate Professor, ECE
Research Associate Professor, iSchool
University of Illinois at Urbana-Champaign
d.katz@ieee.org
https://danielskatz.org
@danielskatz

- BSSw inaugural fellow, Class of 2018
- Project goal: giving software developers long-overdue credit through principles for software citation
  - Implement Software Citation Principles (10.7717/peerj-cs.86, 2016)
- Status: making progress; more work needed
  - Software citation implementation challenges: Katz et al., 2019, "Software Citation Implementation Challenges," arXiv 1905.08674 [cs.CY]; upcoming 2022 IMLS-funded workshop
  - Checklists for (paper) authors and (software) developers: Chue Hong et al., 2019, "Software Citation Checklist for Authors," 10.5281/zenodo.3479198; Chue Hong et al., 2019, "Software Citation Checklist for Developers," 10.5281/zenodo.3482768
  - Best practices for software repositories and registries: Task Force on Best Practices for Software Registries et al., 2020, "Nine Best Practices for Research Software Registries and Repositories: A Concise Guide," arXiv 2012.13117 [cs.DL]
  - Guidance for journals: Katz et al., 2021, "Recognizing the value of software: a software citation guide [version 2; peer review: 2 approved]," 10.12688/f1000research.26932.2
  - Metadata progress in CITATION.cff (adopted by GitHub, Zenodo, Zotero, …) & CodeMeta (being worked into schema.org)
  - Related to FAIR for Research Software (FAIR4RS): Chue Hong et al., 2022, "FAIR Principles for Research Software version 1.0. (FAIR4RS Principles v1.0)," Research Data Alliance, 10.15497/RDA00068

Andrew Lumsdaine

Pacific Northwest National Laboratory, University of Washington, Northwest Institute for Advanced Computing

BSSw Project: Guiding efficient use of modern C++ for high-performance computing

72

Neal Davis

University of Illinois at Urbana-Champaign

Teaching Assistant Professor, Computer Science

Honorable Mention

73

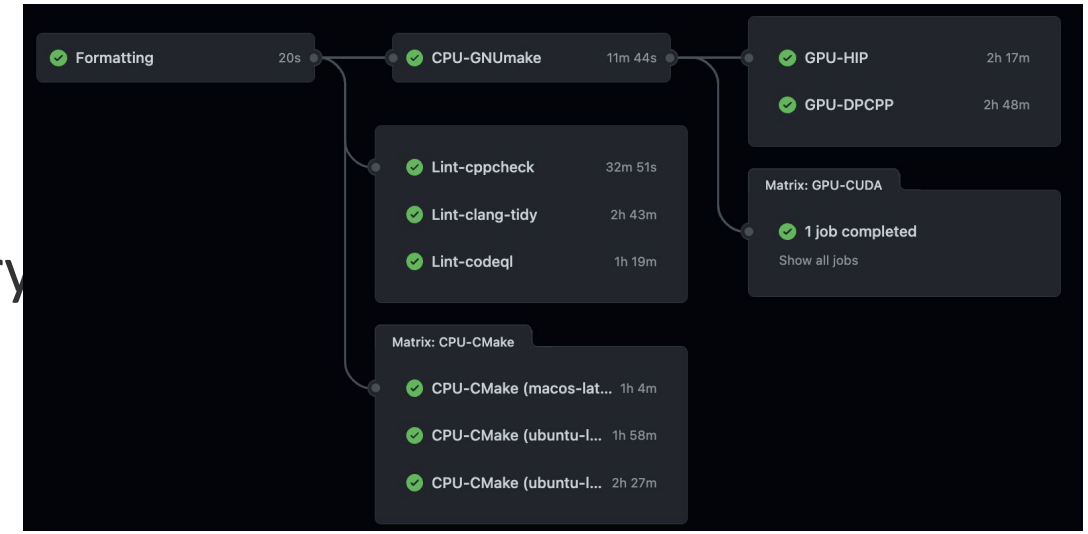
Marc Henry de Frahan

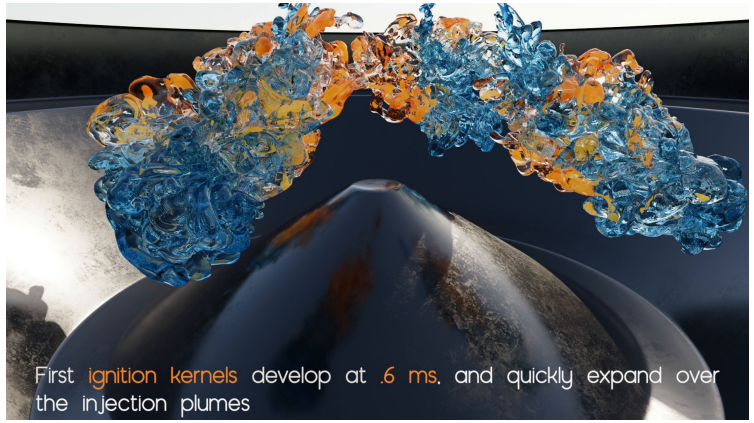
Computational Scientist

National Renewable Energy Laboratory

marc.henrydefrahan@nrel.gov

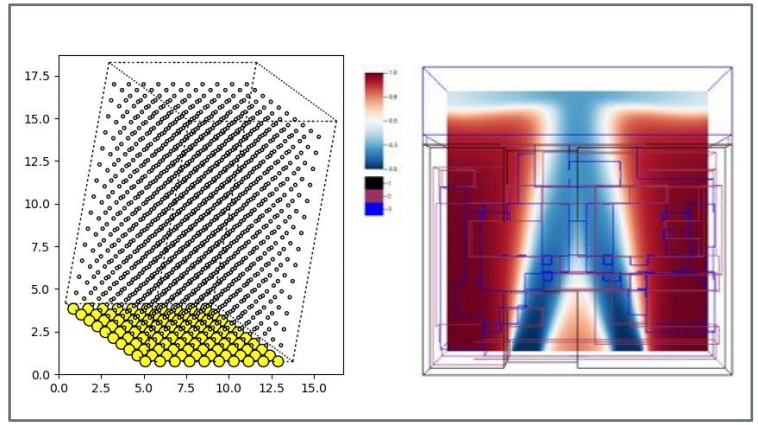
marchdf





PeleC, https://github.com/AMReX-Combustion/PeleC, Henry de Frahan, Marc T., et al. Int. J. High Perf. Comp. App. (2022): 10943420221121151. Image credit: N. Brunhart-Lupo

ExaWind, https://github.com/Exawind. Image credit: N. Brunhart-Lupo



Adaptive Mesh and Algorithm Refinement – coupling continuum flow physics to kinetic Monte-Carlo for material growth. Image credit: R. Larsen, H. Sitaraman.

Elsa Gonsiorowski

gonsiorowski1@llnl.gov

- Joined the IDEAS-ECP project around 2019
- Worked on the Productivity and Sustainability Improvement Planning subteam to develop a light-weight agile software improvement methodology
- Transitioned to the BSSw Fellows team as community organizer at the start of FY22
- Focused on building the community within the fellowship program and building connections to the broader EPC community

Lawrence Livermore National Laboratory

75

Ying Li

Argonne National Laboratory

Argonne Scholar, Argonne Leadership Computing Facility

Honorable Mention

76