



Exceptional service in the national interest

# Research Software Engineering in Practice

The Software Engineering & Research Department at Sandia

Speaker: Miranda Mundt

9 May 2023

NCAR Computational and Information Systems Lab  
Seminar Series

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.





## Outline



- **Where We Came From:** History of Research Software Engineers
- **How We Started:** Software Engineering and Research Department at Sandia National Laboratories
- **How We Engage:** The RSE Community at Sandia National Laboratories



# History of Research Software Engineers (RSEs)



## Before RSE

194 different job titles

Bioinformatics Scientist Biometric Software Systems Developer Bioprospecting Software Developer C++ / C# Graphics Software Engineer Casebooks Project Editor (Research Assistant/ Associate) Climate Researcher (Research Associate) Clinical Study Programmer CoMPLEX Research Associate Computational Biologist / Bioinformatician Computational Scientist Computational Scientist in Computational Fluid Dynamics & Industrial Applications Computational Scientist in Structural Mechanics and Industrial Applications Computer Scientist Computer Vision Researcher Content Developer/ Programmer Control Engineer-IMG - 3 posts CREATE Data Specialist Data Analyst Data Integration Coordinator Data Manager x3 Database and Software Engineer Database Manager / Researcher Database Programmer Digital Media Technician E-Learning Portal Manager (KTP Associate) e-Learning Systems Development Analyst e-Learning Systems Development Analyst (Moodle, SQL) E-Learning Web Developer E-Portfolio Learning Technologist Embedded Systems Engineer Engineering Technician Environmental Scientist EPSRC Studentship on Algorithmic Construction of Finsler-Lyapunov Functions Experimental Officer in Bioinformatics Experimental Psychologist Finance Assistant Gaia Alerts Software Developer Gaia Software Developer (Gaia UK Team) GIS Applications Specialist Graduate Programmer / Software Developer Graphics Programmer Health Data Manager / Scientist High Throughput Bioinformatician High Throughput Sequencing Bioinformatician (two posts) HIVE Manager / HIVE Co-ordinator HIVE Senior Researcher and Technical Lead Hydroinformatics Scientific Software Developer Image Analysis Machine Learning Image Processing Information Systems Developer Instrumentation Engineer Investigator Statistician IT Developer / IT Support (Language III Services Operations e-Learning Systems) IT Support Technician (Unix / Windows Systems) Knowledge Transfer Partnerships (KTP) Knowledge Transfer Partnerships (KTP) Knowledge Transfer Partnerships (KTP) Associate - Software Developer KTP Associate - Robot Vision Scientist (Research Fellow) KTP Associate (Fixed Term Contract for 24 months) KTP Associate (Precision Agriculture Data Analyst) KTP Associate â€” Graduate Web Developer KTP Associate: Electronics / Robotics Engineer Learning Technologist Leicester Respiratory BRU/IT Developer Linguist / Psycholinguist Maker Space Technician Marie Curie Early Stage Researcher Marie Curie Early Stage Researcher in Rad. Res. Cell for Integrated Water Quality Modelling Marine Earth Observation Scientists Medical Statistician Medical Statistician/Senior Medical Statistician Meteorologist Embedded Mobile Application Developer NASC IT Support - Programmer and Systems Administrator (Fixed Term) Network Research Methods Skills PDRA on EU Project on Automated Multisensor Surveillance Planning Officer Policy Modeller 2011 Post - Doctoral Research Assistant INSTRON Post Doctoral Research Worker Post Doctoral Researcher in the application of Digital Technology Post-Doctoral Research Assistant in Simulation and Visualization Post-Doctoral Research Associate Post-Doctoral Research Associate (Pathogen Genomics) Post-Doctoral Research Fellow Postdoctoral Fellow - population genetics / evolutionary genetic Postdoctoral Fellow in Bioinformatics Postdoctoral Fellow in Cancer Therapeutics Postdoctoral Research Assistant Postdoctoral Research Associate Postdoctoral Research Fellow Postdoctoral Research Scientist Postdoctoral Researcher in Declarative (Logic and Functional) Programming Postdoctoral Researcher Postdoctoral Scientist Postdoctoral statistician Postdoctoral Training Fellow - Statistical and Computational Genetics of Autism Principal / Senior Bioinformatician Principal Bioinformatician Product Development Engineer (Rail) Publishing Portal Web Developer Radio Frequency Engineer Reader in Computer Science Reporting Analyst Research (Software) Engineer Research Assistant Research Associate Research Fellow Research Image Data Manager, Biomedical Engineering Research Officer Research Officer â€” Social Protection Research postgraduate Research Programmer Research Scientist Research Scientist / Senior Research Scientist Research Scientist in Machine Learning and Computer Vision Research Software Developer Research Software Developer for the Herchel Smith Professor of Organic Chemistry Research Software Engineer Research Studentship Research Worker Researcher SAP Trainee Technical Analyst Scientific Officer with Michela Garofalo Scientist SEAHA Studentship: Extracting epidemiological data from collections SEEG Data Archive Manager Senior / Research Associate in Clinical Integration and Image Analysis for Fetal Surgery Senior Analyst Programmer (Business Analysis) Senior Analyst/Programmer Senior Bioinformatician Senior Bioinformatician / Bioinformatician Senior Computational Statistician - Spatial Models Senior Data Acquisition Scientist / Data Acquisition Scientist Senior Data Manager Senior Database Administrator Senior IT Developer Analyst Senior Mathematical Modeller Senior Media Developer Senior Postdoctoral Researcher - Evolutionary and Computational Analysis of Infectious Disease (Phylogenetics) Senior Research Assistant Senior Research Associate Senior Research Associate â€” Molecular Modelling & Simulation Senior Research Associate in Quantitative Clinical MRI Senior Research Fellow Senior Research Fellow / Research Fellow in Vibration Diagnostics and Prognostics/Digital Signal Processing Senior Research Laboratory Technician Senior Research Technician Senior Software Developer in Bioinformatics Senior Software Engineer / Software Engineer Senior Statistical Epidemiologist Senior Systems

Hettrick, Simon. "A Not-so-brief History of Research Software Engineers." *Software Sustainability Institute*, 17 Aug. 2017, [www.software.ac.uk/blog/2016-08-17-not-so-brief-history-research-software-engineers-0](http://www.software.ac.uk/blog/2016-08-17-not-so-brief-history-research-software-engineers-0).



## What is an RSE?



# RSE

"We like an inclusive definition of Research Software Engineers to encompass those who regularly use **expertise in programming to advance research**. ...

We aspire to **apply the skills and practices of software development to research** to create more robust, manageable, and sustainable research software."<sup>2</sup>

"Research Software Engineer" is a term coined by a United Kingdom group in 2010 – immortalized in the foundational paper "The research software engineer" in 2012<sup>1</sup>. There are numerous definitions – but we like the US Research Software Engineer Association one...

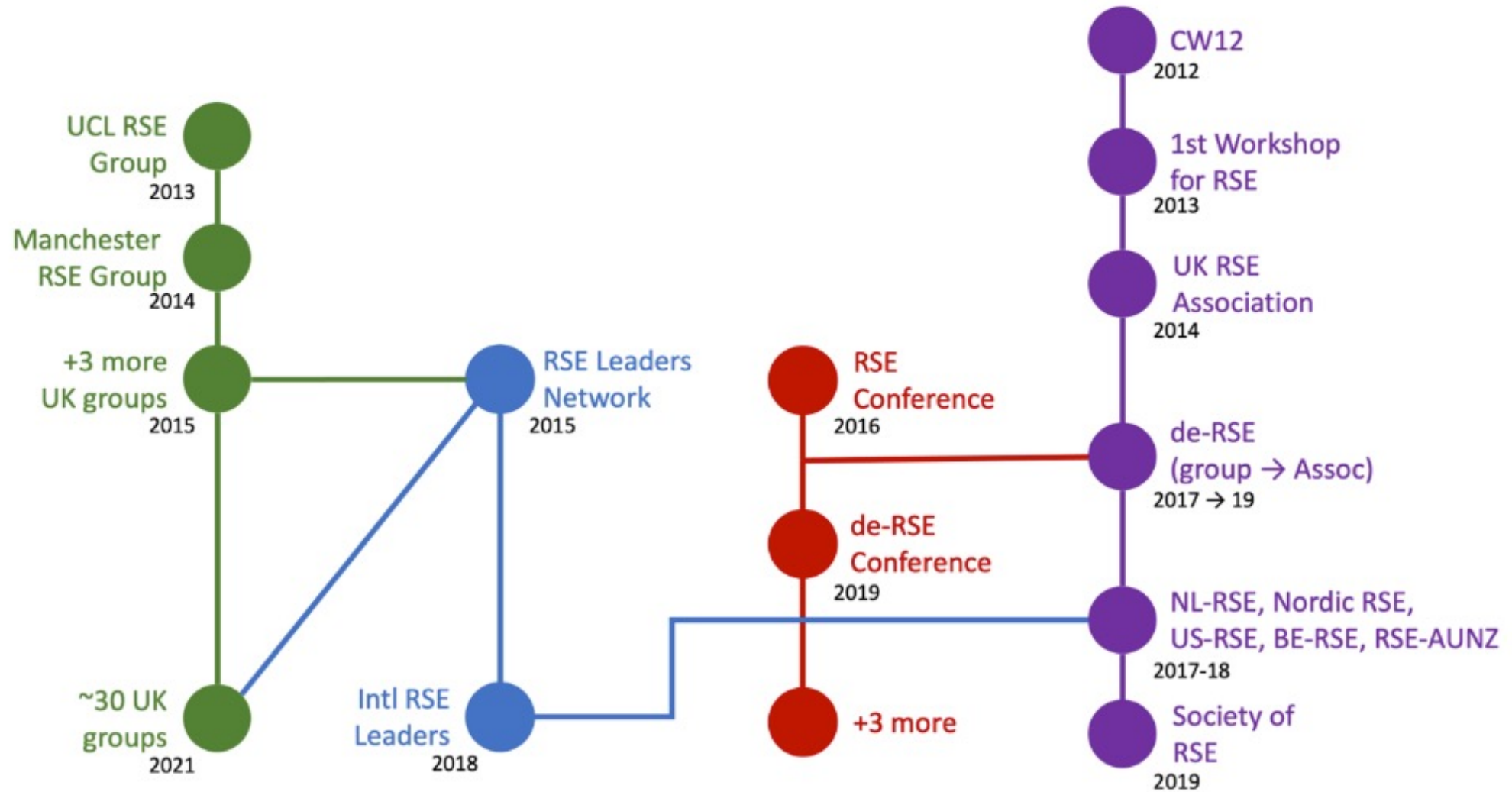
<sup>1</sup> Baxter, Rob, et al. "The research software engineer." *Digital Research Conference, Oxford*. 2012.

<sup>2</sup> "What Is an RSE?" *US Research Software Engineer Association*, 2021, <https://us-rse.org/about/what-is-an-rse/>.





# Growth of the RSE Movement

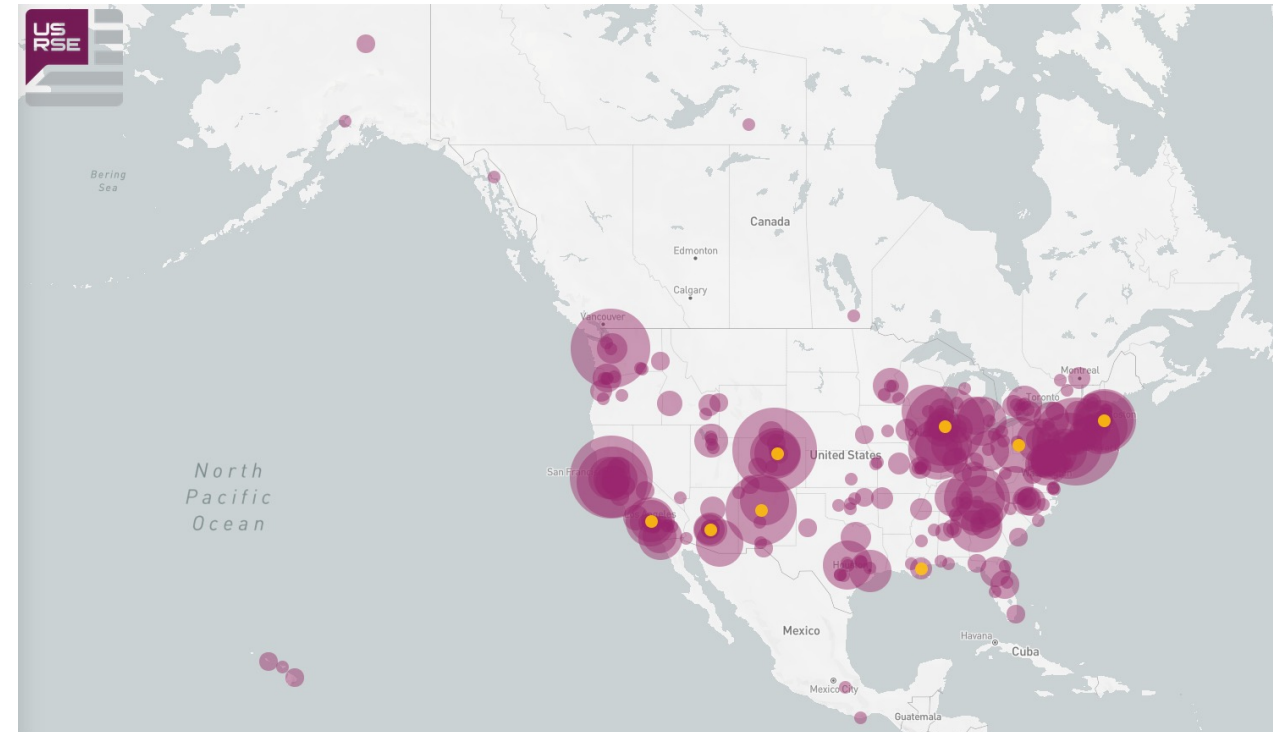




# US Research Software Engineer Association



- A community-driven organization
- Members:
  - Write and contribute research software at:
    - Universities, laboratories, knowledge institutes, companies, etc.
  - Interested in Research Software Engineer careers
  - Students, researchers, software engineers
  - Identify as RSE “allies”
  - Manage, sponsor, support





# US-RSE Mission and Activities

## Mission

- Community
- Advocacy
- Resources
- Diversity, Equality, and Inclusion



Software Enabled Discovery and Beyond  
Chicago, IL, October 16-18<sup>th</sup>

<https://us-rse.org/usrse23/>

## Activities

- Community Calls
- Education and Training
- Diversity, Equity, Inclusion Seminar
- Funder Talks
- Job Board
- Newsletters
- Workshops
- Conference





Software  
Engineering and  
Research  
Department at  
Sandia National  
Laboratories



# Who We Are

- The Department of Software Engineering and Research...
  - Is a cross-disciplinary RSE, SER, SysAdmin, and ITSM team at Sandia National Laboratories
  - Is located within the Center for Computing Research
  - Has the goal of advancing the study and practice of software engineering in the domain of scientific software



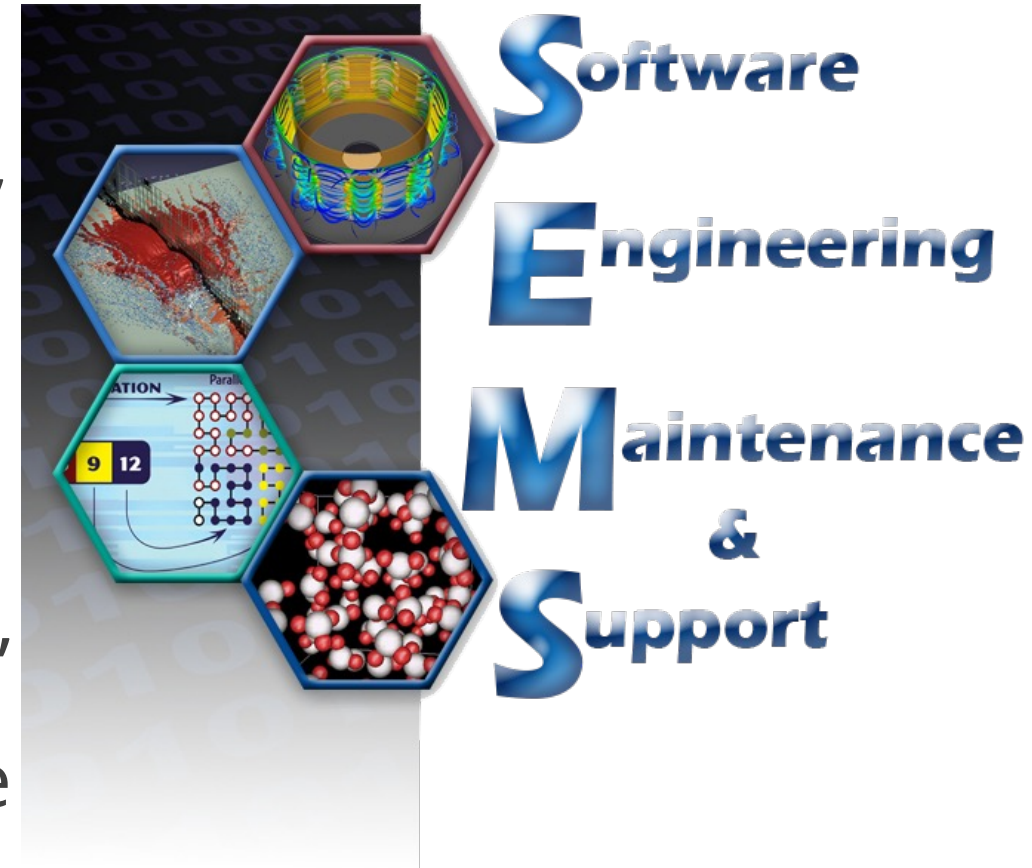
**Software Engineering  
& Research**

Department 1424



# Our Origins

- A few years prior to our department's creation, our center assembled the **Software Engineering, Maintenance, and Support (SEMS)** project.
- This was a project staffed by software practitioners split across several departments.
- SEMS met the demand for useful tools, training, and support, and this spurred motivation among leadership to create a department around the work of SEMS.







# Why Form An Official Department?



- To provide an opportunity for **software-focused staff** to interact more closely with like-minded people
- To make software engineering a **first-order concern** for the organization
- To carve out a space for people in **software-related roles** to **advance their careers**
- To expand to include a **research focus**

# Research, Develop, Deploy



- Our team structure follows what we call a **Research, Develop, and Deploy (RDD)** workflow pattern, centered around three primary areas.
- We aim to cultivate a critical mass of staff in each capability area, each of which mutually reinforces the other.
- Everyone is able to participate in one or more of these areas.



# How are they distributed?



Research

25%



Develop

65%



Deploy

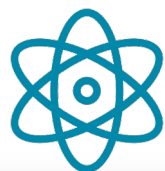
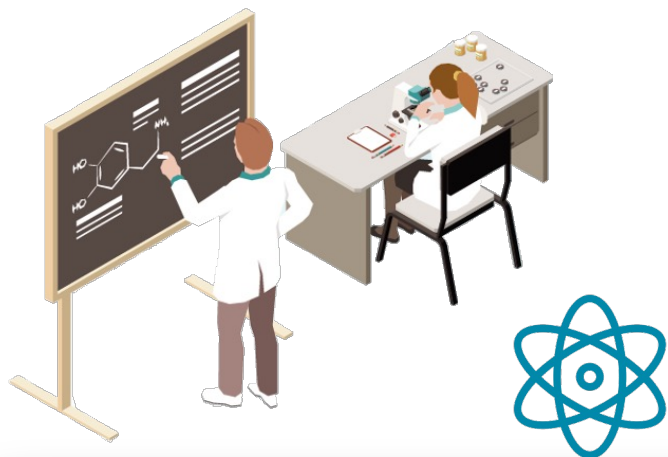
40%







# Research



- Given the importance of our work for the nation, we have a responsibility to act on the basis of the **best available evidence**. A key way in which we acquire and retain that understanding is through **rigorous, systematic investigation**.
- Specialists in this focus area tend to have a **PhD in computer science or a domain science** and a passion for software engineering.
- Specialists also pair with **practitioners** to ensure evidence-based practices are **practical** and **worthwhile**.
- *Examples:* Publications; consultations; literature reviews; rapid reviews; tutorials, workshops, and training.

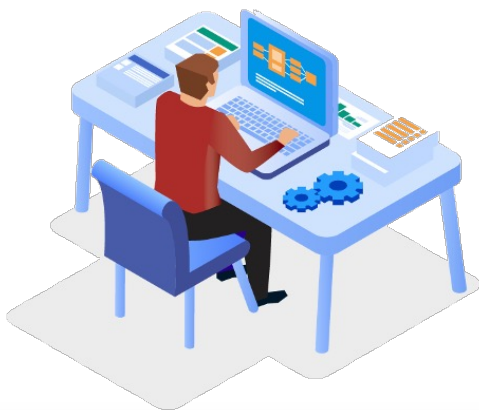


## Research Output: Evidence-Based Practice

- Software Quality Frameworks for Scientific Software Development
- Requirements Elicitation Techniques for RSEs
- Extending Software Quality Models to Include Reproducibility
- Lightweight Software Process Improvement for CSE Teams
- Data-Driven Exploration of Cross-Team Collaborations



# Develop



- A key element of our work is in **teaming with application and algorithm researchers** to provide embedded development, maintenance, and support.
- Team members who focus on development work tend to be **Research Software Engineers** of various stripes, ranging from computer science graduates to staff with a background in science or mathematics who have transitioned into a software-focused career.
  - *Examples:* Embedded software development; scientific programming; expertise with MPI, OpenMP, CUDA.

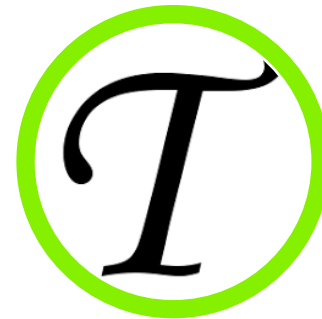
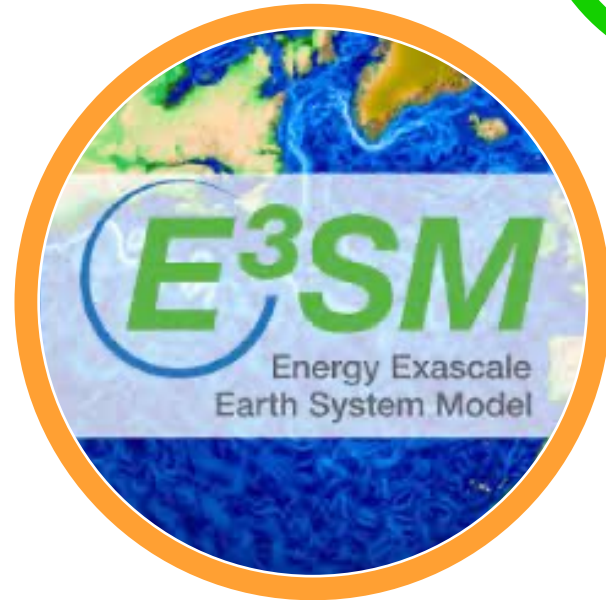


# Develop: Teaming for Success

Trilinos



E3SM



Pyomo





# Deploy



- We firmly believe that **robust, scalable, and sustainable infrastructure** for software projects is vital to the scientific computing mission of our center.
- For this reason, we have a contingent of staff who focus on **System Administration, DevOps, and/or IT Service Management** to lead these efforts.
  - *Examples:* Jenkins-based build/test farms; common dependency management system; off-the-shelf tools like Jira and Confluence; and tailored infrastructure solutions for projects.





## Deploy: Custom-Built Tools and Services



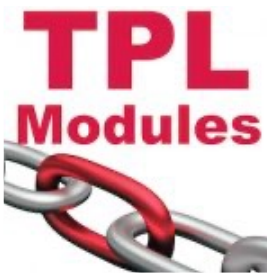
### Watchr

Jenkins plugin capable of ingesting performance data files that were captured at different points in time and generating plots



### Repometer

Tracks and stores relevant usage data from GitHub and GitLab repositories, such as stars, forks, clones, over an extended period



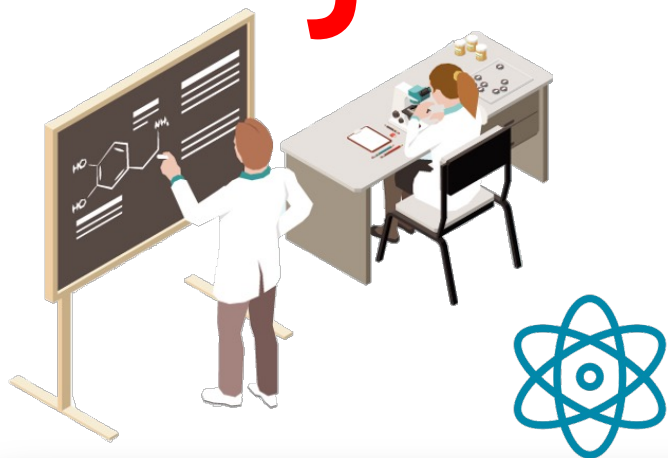
### Environment Modules

A collection of compilers, third-party libraries, and utilities that are used in the development of scientific codes





# JUST KIDDING Research



- Given the importance of our work for the nation, we have a responsibility to act on the basis of the **best available evidence**. A key way in which we acquire and retain that understanding is through **rigorous, systematic investigation**.
- Specialists in this focus area tend to have a **PhD in computer science or a domain science** and a passion for software engineering.
- Specialists also pair with **practitioners** to ensure evidence-based practices are **practical** and **worthwhile**.



# Integrating These Roles



- We believe that placing RSEs into an **interdisciplinary team** makes our team more **cohesive, creative, and productive**.
  - Our **research** work aims to create new knowledge to assist **development** and **deployment** activities.
  - Our **development** work keeps **research** grounded in real-world needs and informs **deployment** activities.
  - Our **deployment** work drives **research** priorities in tooling and support.



# Department Evolution

2022

2021



Publications: 5 papers, 6 talks

Shared tools: spack-cm, Watchr, LoadEnv, SPIFI

Projects: Pyomo, VoroCrust, IDEAS-ECP, Testbeds, Dakota, E3SM...

SEMS Operations & Maintenance

Systems Administration

"Research" – Grants, interns, academic partners

"Research" – 40+ publications: 7+ peer reviewed, 25+ presentations and tutorials; nine department authors

"Develop" – Training (staff, stakeholders)

"Develop" - Scientific Programming

"Develop" - Shared tools

"Develop" - Project support

"Develop, deploy" – Scientific Infrastructure

"Deploy" – Operations & Maintenance

"Deploy" – System Administration

2021 Built upon strong foundation of SEMS "Deploy"

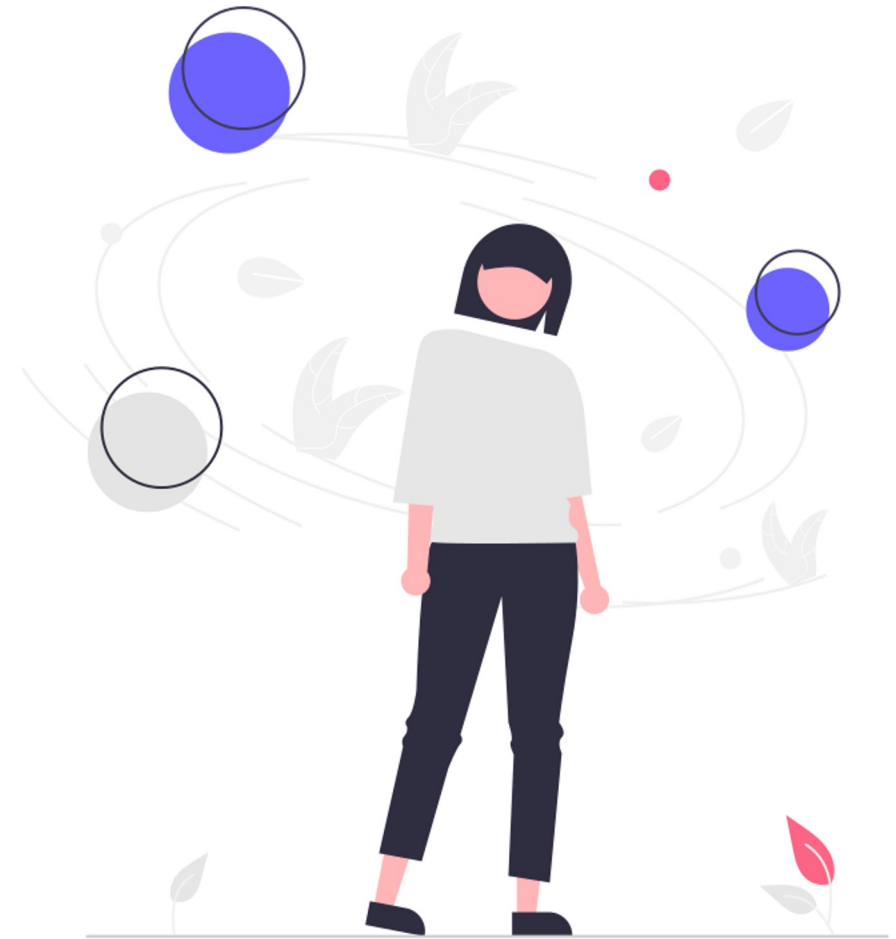
2020 Teaming - Forming, Storming, Norming

2019 Software Engineering & Research established



# What This Means for Our Customers

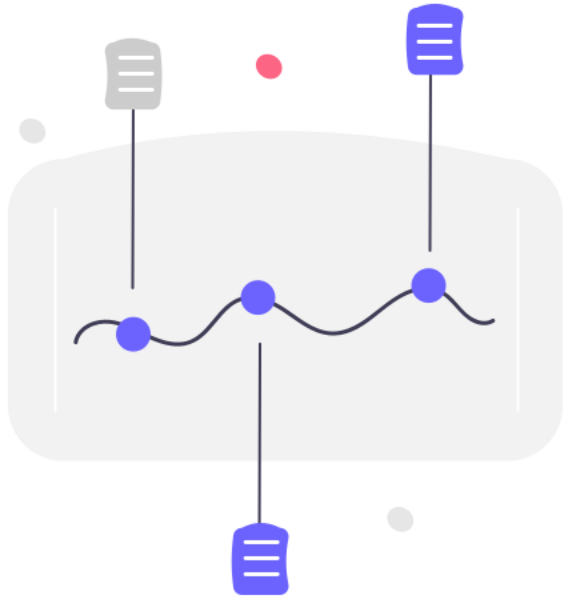
- In all of our diverse efforts, our goal is to **empower** subject matter experts to work more **efficiently** and **effectively**.
- The ways in which our team impacts developer **productivity** are numerous, but some can be subtle.
- Consider the case of a typical scientist-developer in our center...







# What This Means for Our Customers



- She is able to **track her team's work** through Jira and Confluence which we offer and maintain.
- Her ongoing research into next-generation algorithms is **directly supported** by an embedded RSE from our department.
- To manage her repository, she uses a git workflow that **we set up and trained her** on.
- Whenever she submits a pull request, our tools detect this and **launch build and test jobs** on our server farm to ensure her code performs as intended.



# What This Means for Our Customers

- Her software relies upon a complex system of library dependencies, all of which we **seamlessly provision** through our environment module system.
- Recently she reached out to our team for consultation on software design, and we are in the process of **compiling peer-reviewed literature and industry best practices** to inform her design-related decisions.







# What This Means for Our Customers



- She sends an email to a colleague about an interesting talk she attended by a leading expert in software engineering. **We invite him.**
- When she encounters a software-related problem, she submits a help ticket to the CCR Help Center. **We operate that.**
- She has a project website to advertise her research papers and document her code. **We build that.**

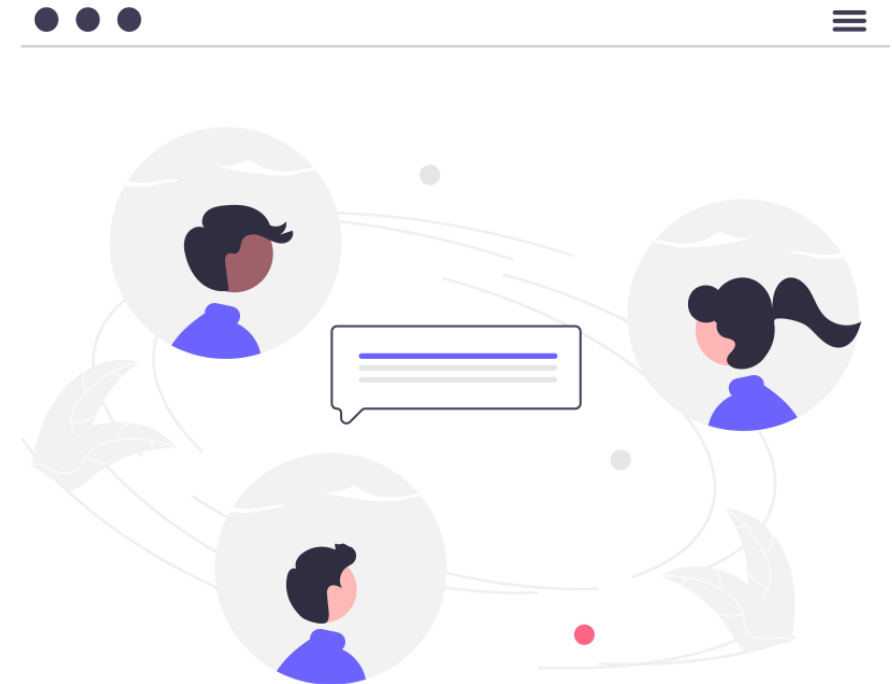


The RSE  
Community at  
Sandia National  
Laboratories



## Creating a Community of Practice

- In early 2022, members of our department began reaching out to other groups across the lab to gauge interest in a **Research Software Engineering Community of Practice**.
- Our RSEs identified those who would become the 9 founding members of the RSECOP, which began officially in September 2022.
  - Kick-off meeting had over **200 attendees**
  - Mailing list has over **100 members**
  - Meetings regularly attract **30-50 attendees**





## Goal: Community

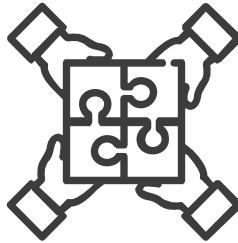


Community

- **Connect, network, collaborate** and **share**
- Embody a shared **value** system
- **Transparency**
- Right **resources** at the right time
- Effective **teaming**



## Goal: Knowledge Sharing



Knowledge  
Sharing

- Effective knowledge sharing
- Aggregate **tools** and **resources**
- Create materials for
  - **Training**
  - **FAQs**
  - **Problems** and their **solutions**
- Share **opportunities**



## Goal: Better Practices



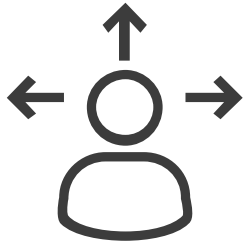
Better Practices

- **Embody** and **exercise** better software engineering **practices**
- **Awareness** and **promotion** of right-sized practices
- Solution **patterns** and knowledge **elicitation**
- **Education**, especially for those with **no formal training** in software engineering





## Goal: Advocacy, Outreach, and Impact



Advocacy,  
Outreach, and  
Impact

- **Promote** impact of Research Software Engineering
- Recognize the **role, accomplishments,** and **impact** of RSEs
- Convey impact to **funders** and **customers**
- Extend outreach into the **wider complex** of Department of Energy National Laboratories



# Community Activities

## Monthly Community Meetings

- Accessibility and Inclusivity in Software Development
- Targeting a 0% Error Rate – Good Practices for Designing a High Ribor System
- NormConf Post-Attendance Report

## Monthly Newsletters

- Changelog: Last meeting information, changes to website
- Next Meeting Information
- Related Events
- Upcoming Deadlines and Opportunities

## Calls for Collaboration

- Members choose venues or open calls to submit to
- Members identify a topic/theme
- Members put out a call to the community for collaborators

## Distribution List

- Emails sent for:
  - Submission and volunteer opportunities
  - Upcoming meetings
  - Upcoming events
  - Interesting/useful resources



# Release Notes

RESEARCH SOFTWARE ENGINEERING COMMUNITY OF PRACTICE MONTHLY NEWSLETTER



# How we engage with the wider RSE Community





Conclusion



## Summary



- **Where We Came From:** History of Research Software Engineers
- **How We Started:** Software Engineering and Research Department at Sandia National Laboratories
- **How We Engage:** The RSE Community at Sandia National Laboratories



Q&A  
and  
Discussion



## Resources / Links

- US Research Software Engineer Association: <https://us-rse.org>
- “A Tiered Approach to Scientific Software Quality Practices,” in [Proceedings of the 2022 Improving Scientific Software Conference](#)
- “Lightweight Software Process Improvement Using Productivity and Sustainability Improvement Planning (PSIP),” part of the [Communications in Computer and Information Science](#) book series
- “Towards a Data-Driven Understanding of Cross-team Collaboration” – [PDF](#)
- Trilinos: <https://trilinos.github.io/>
- E3SM: <https://e3sm.org/>
- Pyomo: <http://www.pyomo.org/>
- Repometer: <https://github.com/sandialabs/repometer>
- Watchr: <https://github.com/sandialabs/watchr-core>