SGCI Science Gateways Community Institute

URSSI US Research Software Sustainability Institute

Sustainability in Research Computing

Sandra Gesing
sandra.gesing@nd.edu

Virtual Residency Summer Workshop on
Intermediate Research Computing Facilitation
University of Oklahoma
August 6, 2018

UNIVERSITY OF NOTRE DAME

CRC CENTER FOR RESEARCH COMPUTING

Sustainability

University of New Hampshire:
“We define sustainability as what sustains us as diverse people and communities—from clean air and water to healthcare, education and art—and making decisions in our individual and collective lives with this big picture in mind.

Sustainability is both local and global. It requires of us that we consider both the past and the future in terms of current and best practices.

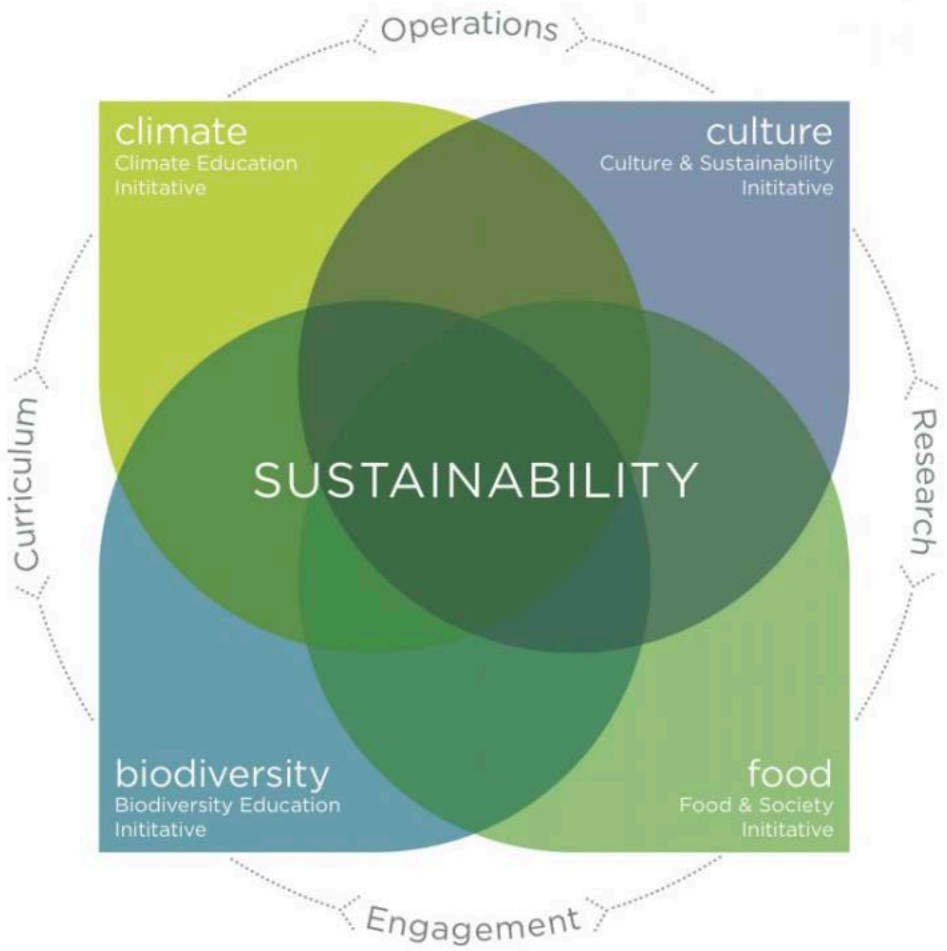
At UNH, we use the sustainable learning community model developed by Dr. Tom Kelly. Sustainability involves maintaining the long-term health of biodiversity, climate, food, and culture, and where these four systems interact. “

https://sustainableunh.unh.edu/whatissustainability

Sustainability for Cyberinfrastructure

of hardware, software, teams…

hard skills and soft skills



hardware, software, algorithms, domain research

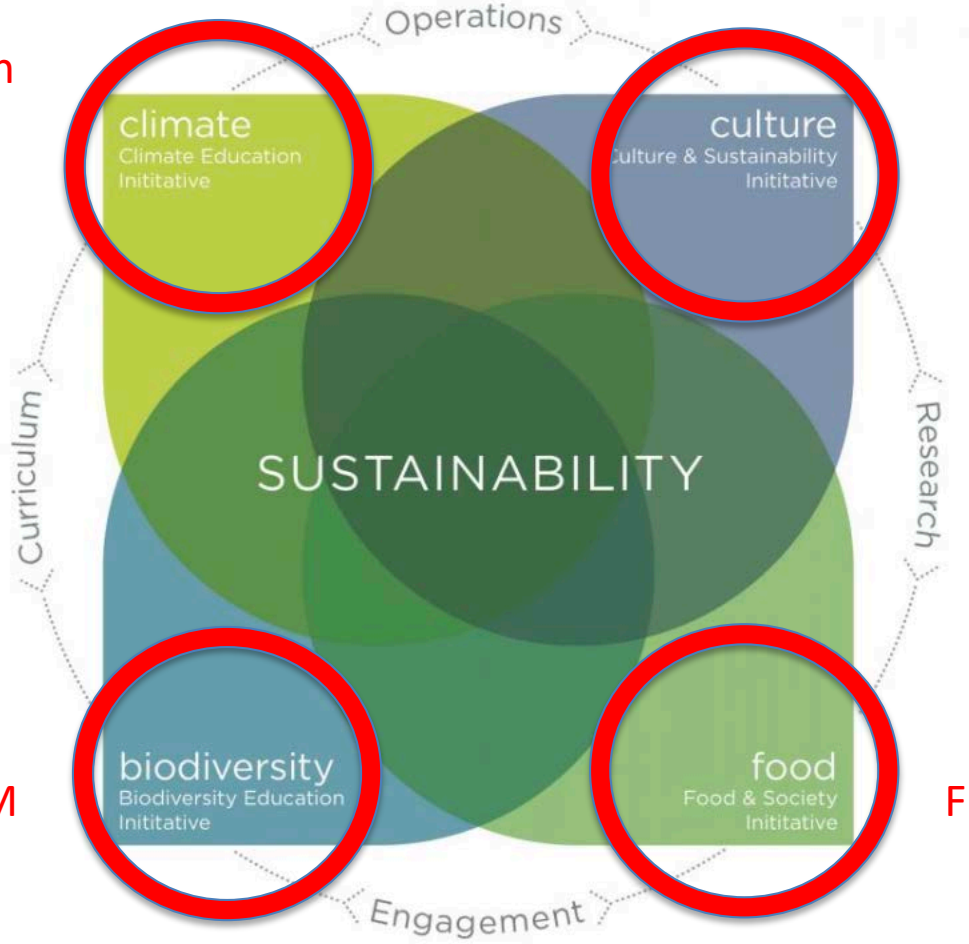
with researchers, institutions' key people, funding bodies, CI community, …

Sustainability for Cyberinfrastructure

of hardware, software, teams…

Climate between Research and Facilitation

Culture: Innovation vs. Maintenance



hard skills and soft skills

hardware, software, algorithms, domain research

Diversity in STEM

Funding and resources

with researchers, institutions' key people, funding bodies, CI community, …

Sustainability for Cyberinfrastructure

Remember the 4 facings…

Sustainability is not a local effort, it is a community effort

You are not alone!

Sustainability for Cyberinfrastructure

Get to know your

- users (diverse research domains, faculty, …)
- stakeholders (host institution, funding bodies – NSF, NIH, DoE, DoD, DARPA, Moore Foundation, etc.)
- partners (projects, initiatives, experienced IT people)
- volunteers (contributors to open-source and/or open science)

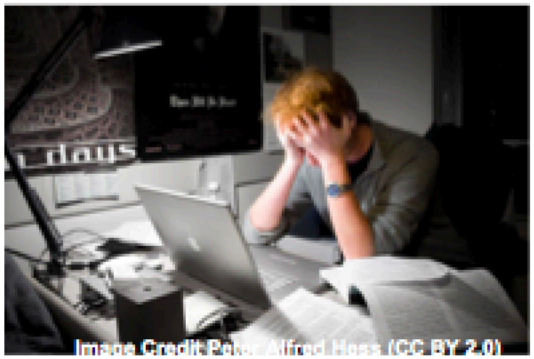
and their challenges as well as their goals – besides publications and funding.

Often their challenges are your challenges!

- Computing resources
- Data analytics
- Preservation needs

Sustainability for Cyberinfrastructure

Bridging the Gap to Data Sharing



Researchers

"the local academic community struggles to effectively manage its assets which manifested itself in a number of challenges, and as for researchers, they lacked storage capacity and data curation processes, and the institution lacked standard metadata and indexing technologies, as well as tools that would support the whole research workflow" - Digital Asset Strategy Committee, DigitalND, 2011

Libraries

Typically, data curation happens retroactively, and as a result data is either not captured at all or available metadata is incomplete.

Pressures from the Outside

"...digitally formatted scientific data resulting from unclassified research supported wholly or in part should be stored and publicly accessible to search, retrieve, and analyze." - White House OSTP Public Access Memo, Feb. 2013

https://presqt.crc.nd.edu/

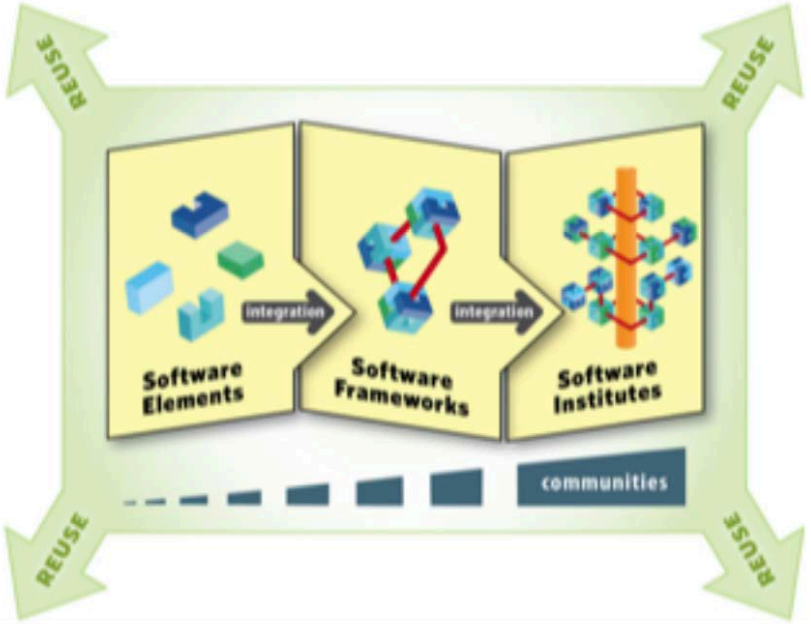
Sustainability for Cyberinfrastructure

Look at financial and non-financial support ("free" resources)

- Do you have people such as digital librarians?
  They are generally not only serving humanities and have great knowledge about data preservation, data lifecycle, programming skills, ..
- Do you have data scientists?
  They probably know about machine learning, meta-data, ontologies, statistics …
- Do you have business scientists?
  They know about marketing, financial strategies, how to build an enterprise, …

They can be partners for you to support CI projects!

Sustainability for Cyberinfrastructure - NSF



Elements: Small groups - create & deploy robust capabilities for demonstrated need to advance science & engineering.

Framework Implementations: Larger teams organized around the development and application of common infrastructure aimed at solving common research problems, resulting in a sustainable community framework serving a diverse community or communities.

Planning Grants for Community Cyberinfrastructure: Focus on long-term capabilities in cyberinfrastructure to serve a research community of substantial size and disciplinary breadth.

Community Cyberinfrastructure Implementations: Focus on long-term hubs of excellence in cyberinfrastructure and technologies, to serve a research community of substantial size and disciplinary breadth.

SI2
Software Infrastructure for Sustained innovation

CSSI
Cyberinfrastructure for Sustained Scientific Innovation

Sustainability for Cyberinfrastructure - NSF

Sustainability Institutes and Excellence Hubs are funded to support the CI and research community

Support via implemented institutes is free for you!
Your chance to influence conceptualizations!

Implementations

- Science Gateways Community Institute
- The Molecular Sciences Software Institute

Conceptualizations

- URSSI
- High-Energy Physics
- Geospatial
- …

The CI Professional Ecosystem

- Clemson-led ACI-REF project
- Coalition for Academic Scientific Computation
- Campus Research Computing Consortium (CaRCC)
- Campus Champions
- CyberAmbassadors
- Linux Clusters Institute
- SIGHPC Education Chapter
- Software & Data Carpentry
- **Science Gateways Community Institute**
- UK Research Software Engineer Association
- **US Research Software Engineer Association**
- UK Software Sustainability Institute
- **Working Toward Sustainable Software for Science Practice and Experience (WSSSPE)**
- **US Research Software Sustainability Institute**

SGCI Science Gateways Community Institute

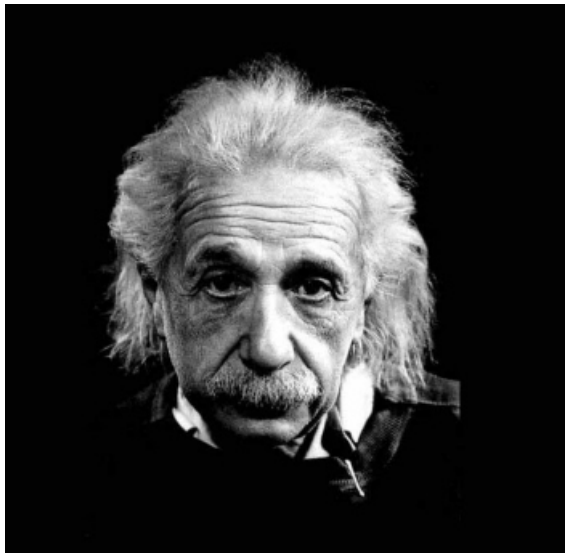
CRC CENTER for RESEARCH COMPUTING

URSSI US Research Software Sustainability Institute

State of the Art in Research

Increased complexity of

- research questions
- hardware
- software
- instruments
- data volume
- data formats

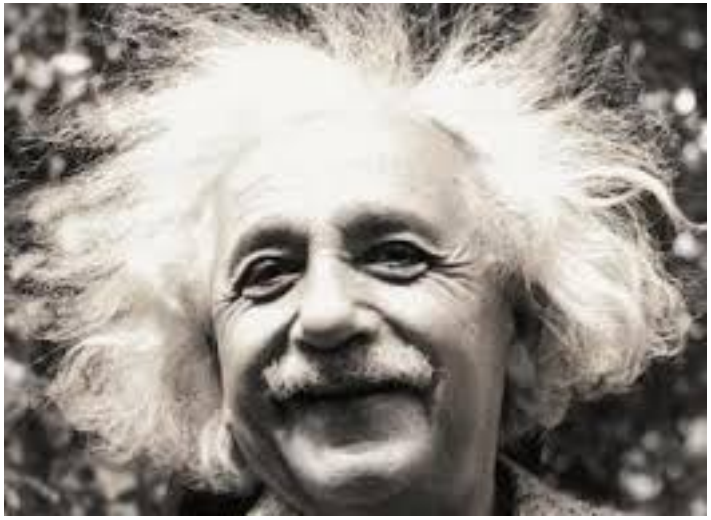


The need for end-to-end solutions for accessing data, software, computing services, and equipment specific to the needs of a science or engineering discipline

Science Gateways

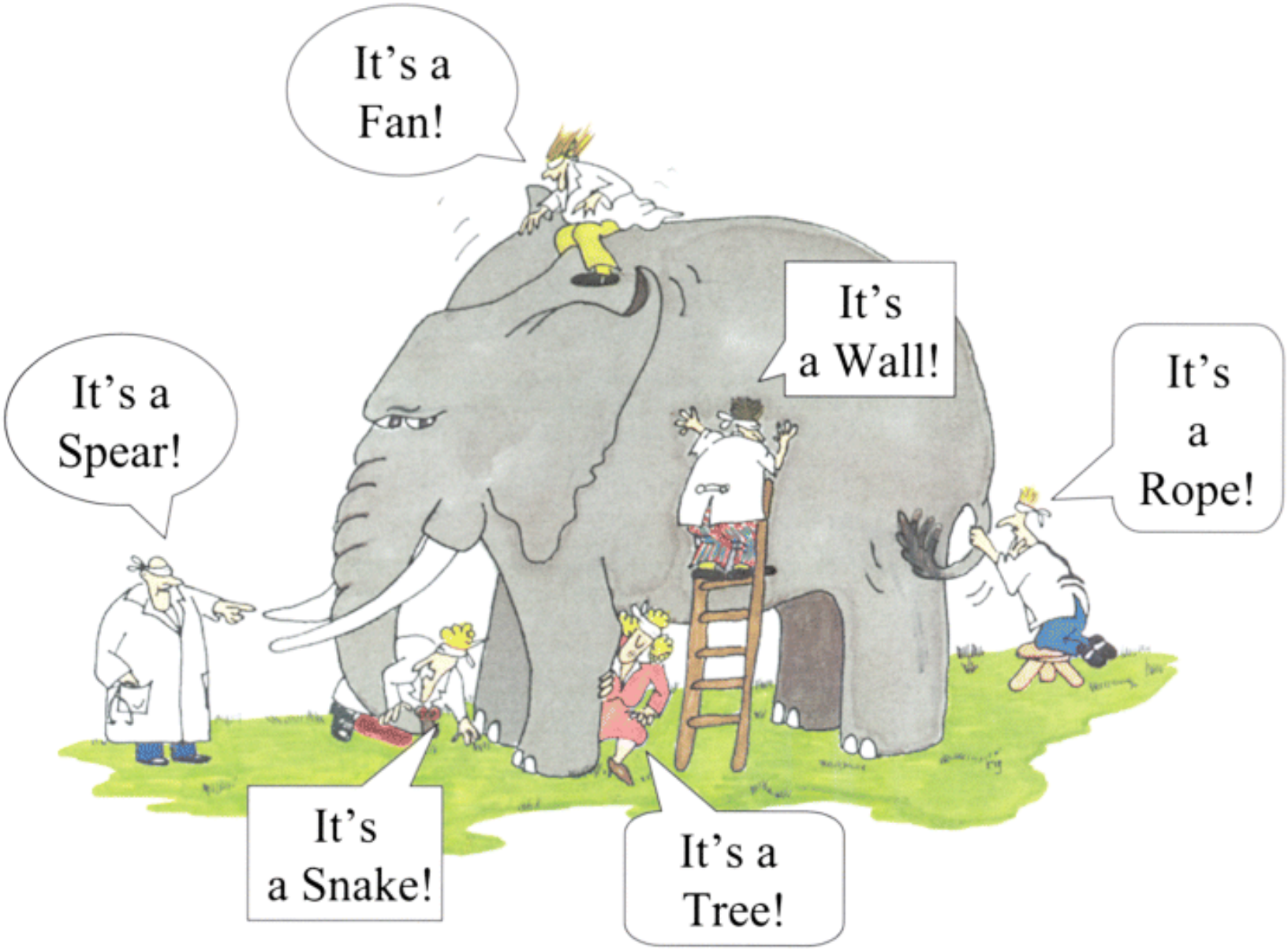
Increased complexity of

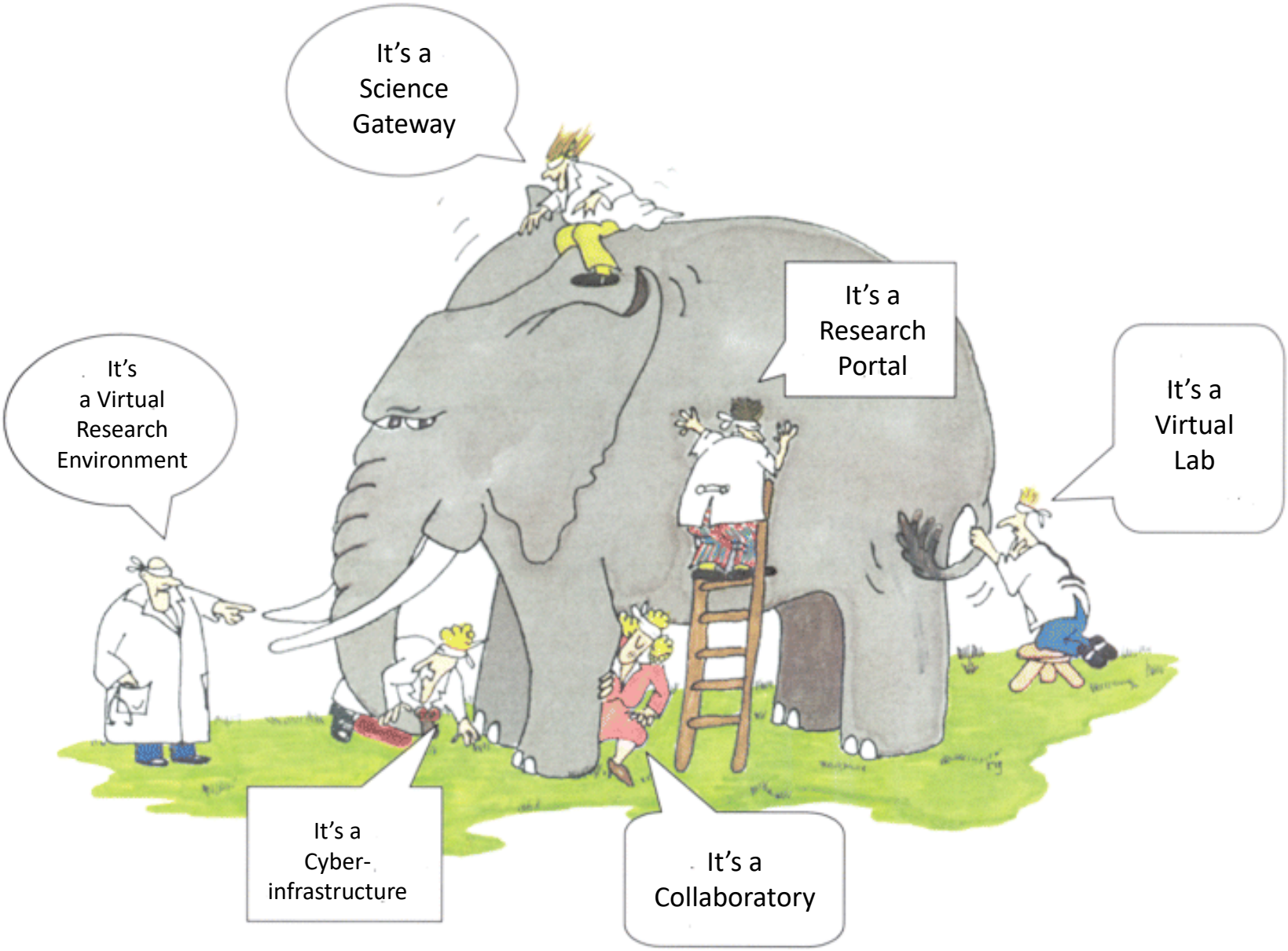
- research questions
- hardware
- software
- instruments
- data volume
- data formats



Science Gateways!

The need for end-to-end solutions for accessing data, software, computing services, and equipment specific to the needs of a science or engineering discipline

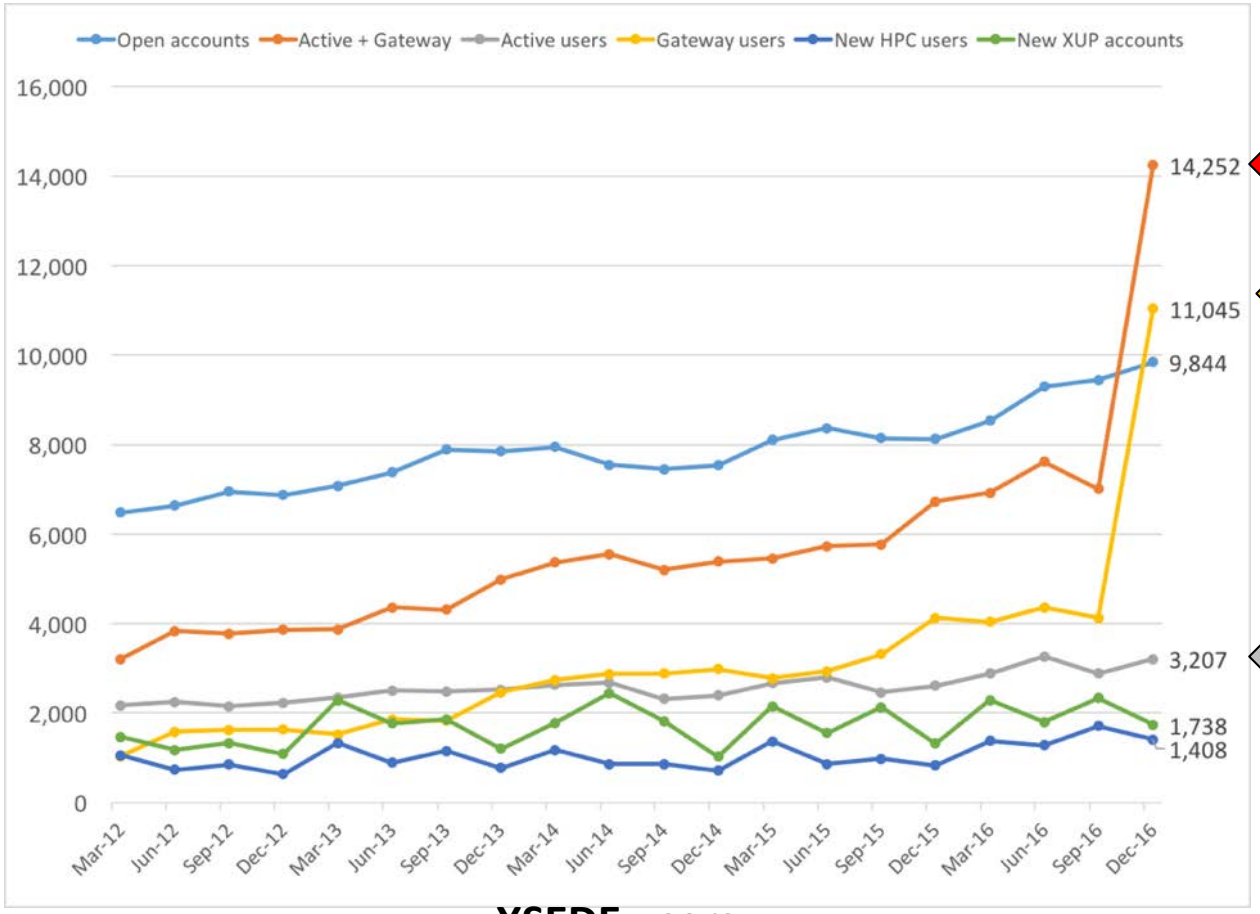

It's a Fan!
It's a Spear!
It's a Wall!
It's a Rope!
It's a Snake!
It's a Tree!



SGCI Science Gateways Community Institute

Science Gateways

Gateway users are 77% of active XSEDE users in Q4 2016



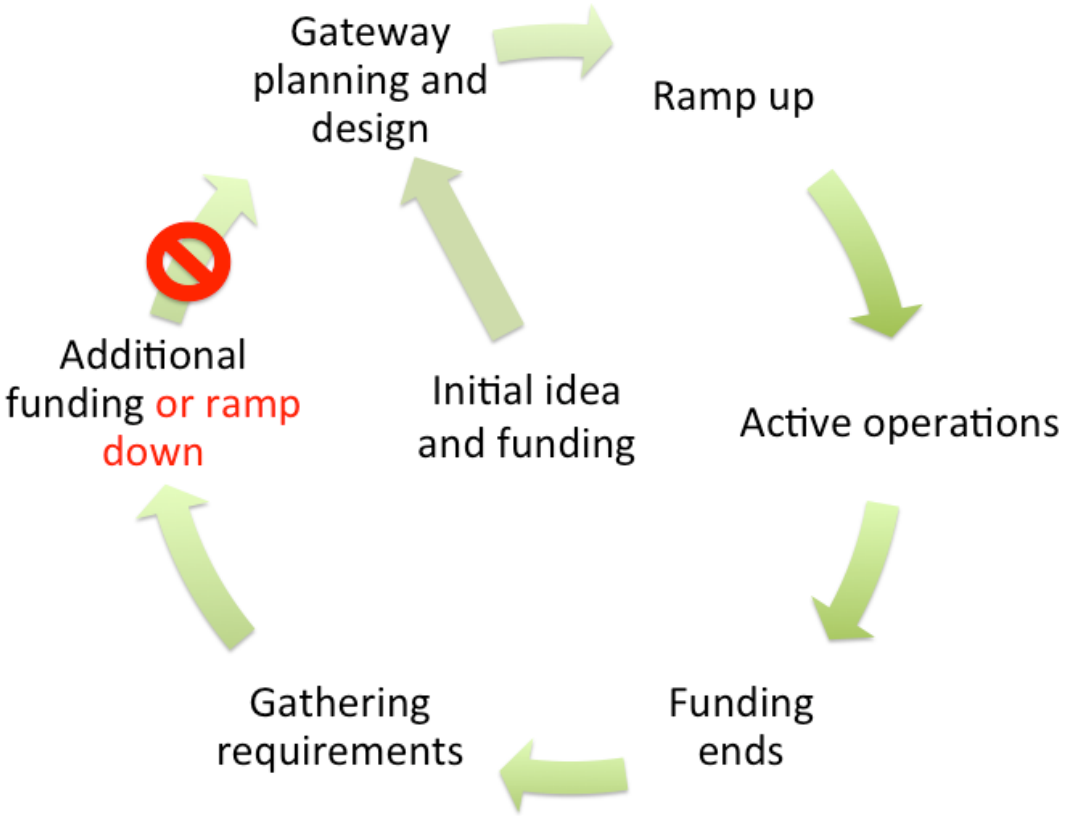
XSEDE users

This is largely due to the CIPRES and I-TASSER gateways, but others are gaining

Life Cycle of a Science Gateway

Developers typically
- work in isolation
- must bridge to variety of resources
- need building blocks in order to focus on higher-level functionality
- struggle to secure sustainable funding

Sounds familiar?



SGCI Science Gateways Community Institute

Science Gateway Survey 2014

- sent out to 29,000 persons
- 4,957 responses from across domains
- 52% from life, physical or mathematical sciences
- 32% from computer and information sciences or engineering
- 45% develop data collections
- 44% develop data analysis tools

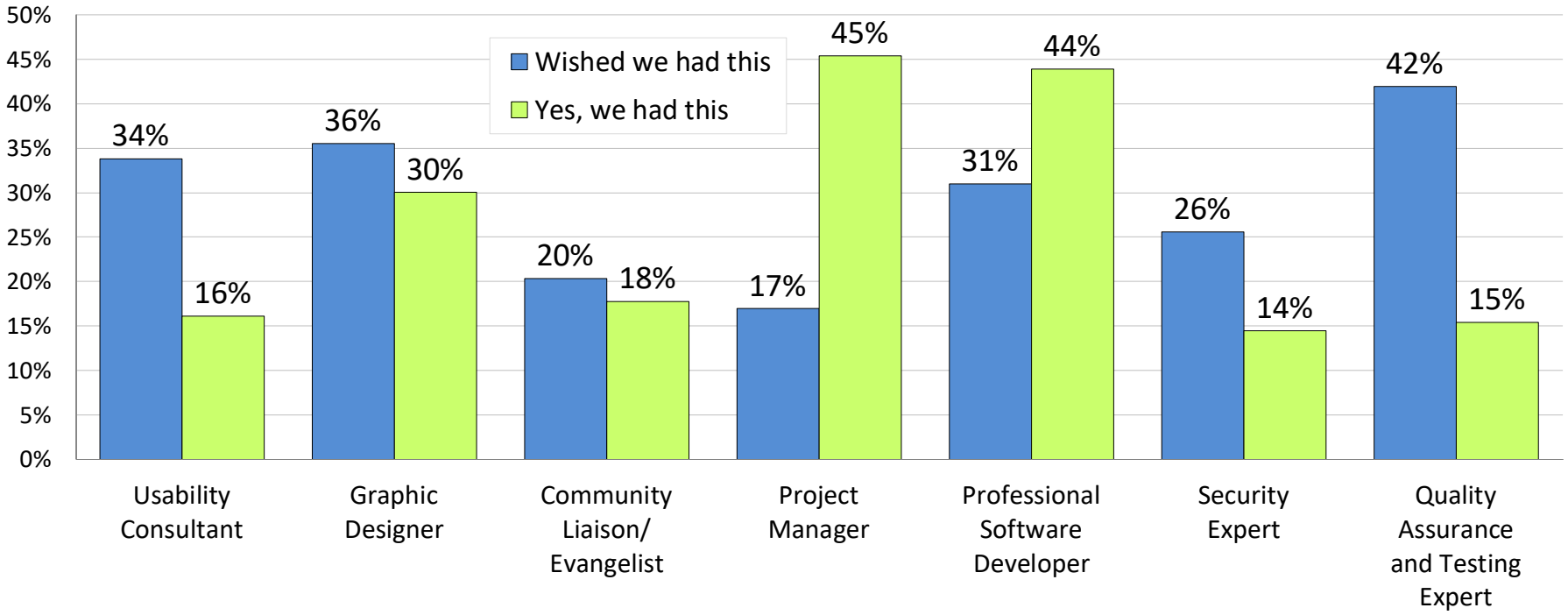
What services would be helpful?

Proposed Service	% Interest
Evaluation, impact analysis, website analytics	72%
Adapting technologies	67%
Web/visual/graphic design	67%
Choosing technologies	66%
Usability Services	66%
Visualization	65%
Developing open-source software	64%
Support for education	64%
Community engagement mechanisms	62%
Keeping your project running	62%
Legal perspectives	61%
Managing data	60%
Computational resources	59%
Mobile technology	59%
Database structure, optimization, and query expertise	59%
Data mining and analysis	58%
Cybersecurity consultation	57%
Website construction	57%
Software engineering process consultation	53%
Source code review and/or audit	51%
High-bandwidth networks	45%
Scientific instruments or data streams	44%
Management aspects of a project	38%

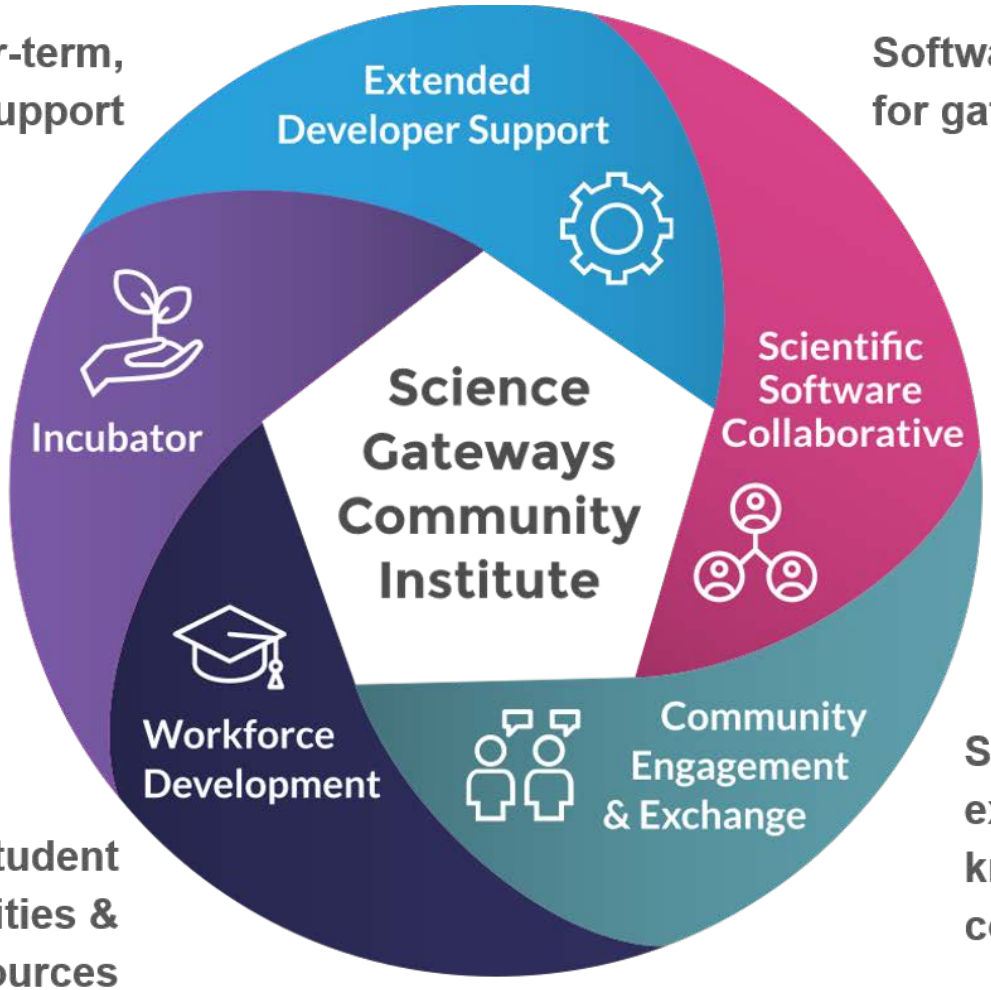
SGCI Science Gateways Community Institute

Science Gateway Survey 2014

Well-designed gateways require a variety of expertise



Science Gateways Community Institute



SGCI Science Gateways Community Institute

Sustainability for Cyberinfrastructure

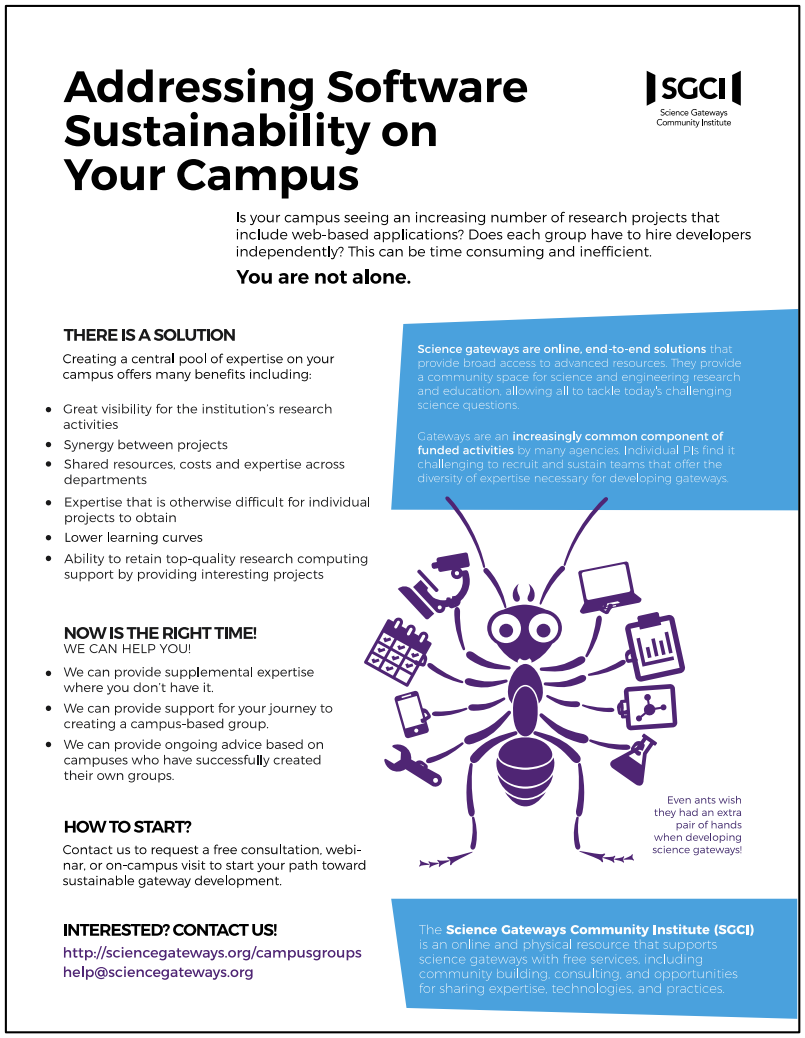
On-campus teams

It is a **centralized** team at your institution –

irrespective whether you are part of a university, a national lab, an organization, a consortium or a company…

Local teams vs. distributed and remote teams:

For local teams it is **still easier** to build more **trust**, to be more **efficient** and to create a **strong culture**.



Addressing Software Sustainability on Your Campus

SGCI
Science Gateways Community Institute

Is your campus seeing an increasing number of research projects that include web-based applications? Does each group have to hire developers independently? This can be time consuming and inefficient.

You are not alone.

THERE IS A SOLUTION

Creating a central pool of expertise on your campus offers many benefits including:

- Great visibility for the institution's research activities
- Synergy between projects
- Shared resources, costs and expertise across departments
- Expertise that is otherwise difficult for individual projects to obtain
- Lower learning curves
- Ability to retain top-quality research computing support by providing interesting projects

NOW IS THE RIGHT TIME!

WE CAN HELP YOU!

- We can provide supplemental expertise where you don't have it.
- We can provide support for your journey to creating a campus-based group.
- We can provide ongoing advice based on campuses who have successfully created their own groups.

HOW TO START?

Contact us to request a free consultation, webinar, or on-campus visit to start your path toward sustainable gateway development.

INTERESTED? CONTACT US!

http://sciencegateways.org/campusgroups
help@sciencegateways.org

Science gateways are online, end-to-end solutions that provide broad access to advanced resources. They provide a community space for science and engineering research and education, allowing all to tackle today's challenging science questions.

Gateways are an **increasingly common component of funded activities** by many agencies. Individual PIs find it challenging to recruit and sustain teams that offer the diversity of expertise necessary for developing gateways.

The **Science Gateways Community Institute (SGCI)** is an online and physical resource that supports science gateways with free services, including community building, consulting, and opportunities for sharing expertise, technologies, and practices.

https://www.codementor.io/blog/modern-engineering-teams-4ea9dpftzs

Connect with SGCI

Incubator Sustainability Bootcamp

- https://sciencegateways.org/engage/bootcamp

- 5 full days
- Teams on projects
- Interactivity
- Community formation
- Putting away the normal daily routine
- Homework

- twice per year
- additional ones can be booked (travel expenses for presenters)
- adapted to feedback

Connect with SGCI

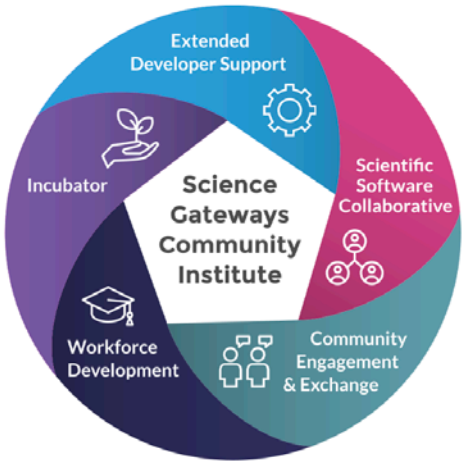
Incubator Sustainability Bootcamp

- https://sciencegateways.org/engage/bootcamp

Work with us

- https://sciencegateways.org/consulting/work-with-us

We can work with gateways at any stage of their lifecycle: a new idea, a proposal seeking funding, a gateway under construction, or an established gateway.



If you're **ready to get help** from one of our consultants, fill out our Consulting Services Request Form.

If you're **writing a proposal** and want to include our services in your proposal, request a letter of collaboration.

If you're **just getting started** and need server space for hosting your gateway, request access to our hosting service.

If you have a **general or short question,** just email us at help@sciencegateways.org.

SGCI Science Gateways Community Institute

Connect with SGCI

Incubator Sustainability Bootcamp

- https://sciencegateways.org/engage/bootcamp

Work with us

- https://sciencegateways.org/consulting/work-with-us

Yearly Conference **Early-bird registration ends today!**

- https://sciencegateways.org/engage/annual-conference

Connect with SGCI

Incubator Sustainability Bootcamp

- https://sciencegateways.org/engage/bootcamp

Work with us

- https://sciencegateways.org/consulting/work-with-us

Yearly Conference **Early-bird registration ends today!**

- https://sciencegateways.org/engage/annual-conference

Become involved as a partner or affiliate

- https://sciencegateways.org/about/partners

Connect with SGCI

Incubator Sustainability Bootcamp

- https://sciencegateways.org/engage/bootcamp

Work with us

- https://sciencegateways.org/consulting/work-with-us

Yearly Conference **Early-bird registration ends today!**

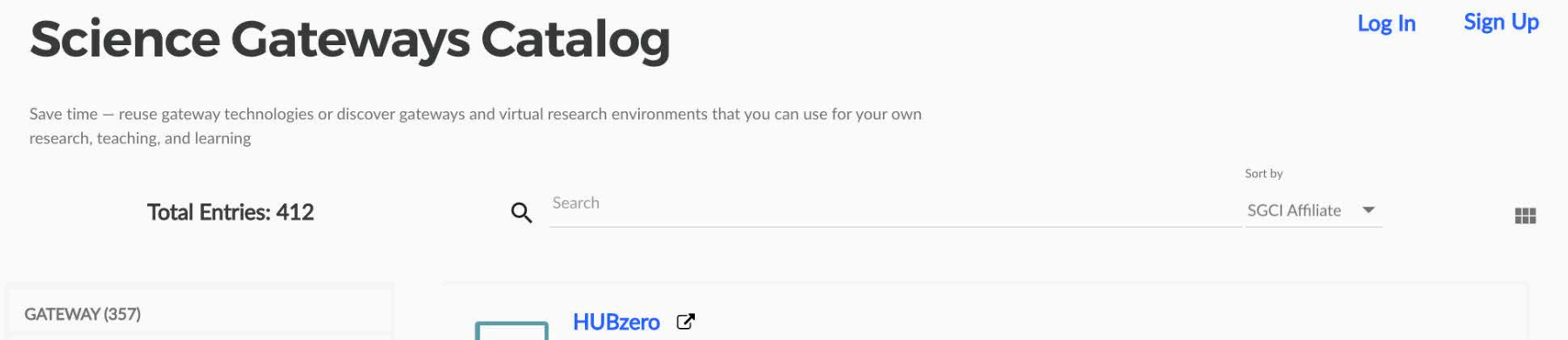
- https://sciencegateways.org/engage/annual-conference

Become involved as a partner or affiliate

- https://sciencegateways.org/about/partners

Software/Gateway Catalog

- https://catalog.sciencegateways.org/



Connect with SGCI

Incubator Sustainability Bootcamp

- https://sciencegateways.org/engage/bootcamp

Work with us

- https://sciencegateways.org/consulting/work-with-us

Yearly Conference **Early-bird registration ends today!**

- https://sciencegateways.org/engage/annual-conference

Become involved as a partner or affiliate

- https://sciencegateways.org/about/partners

Software/Gateway Catalog

- https://catalog.sciencegateways.org/

Train students in internships

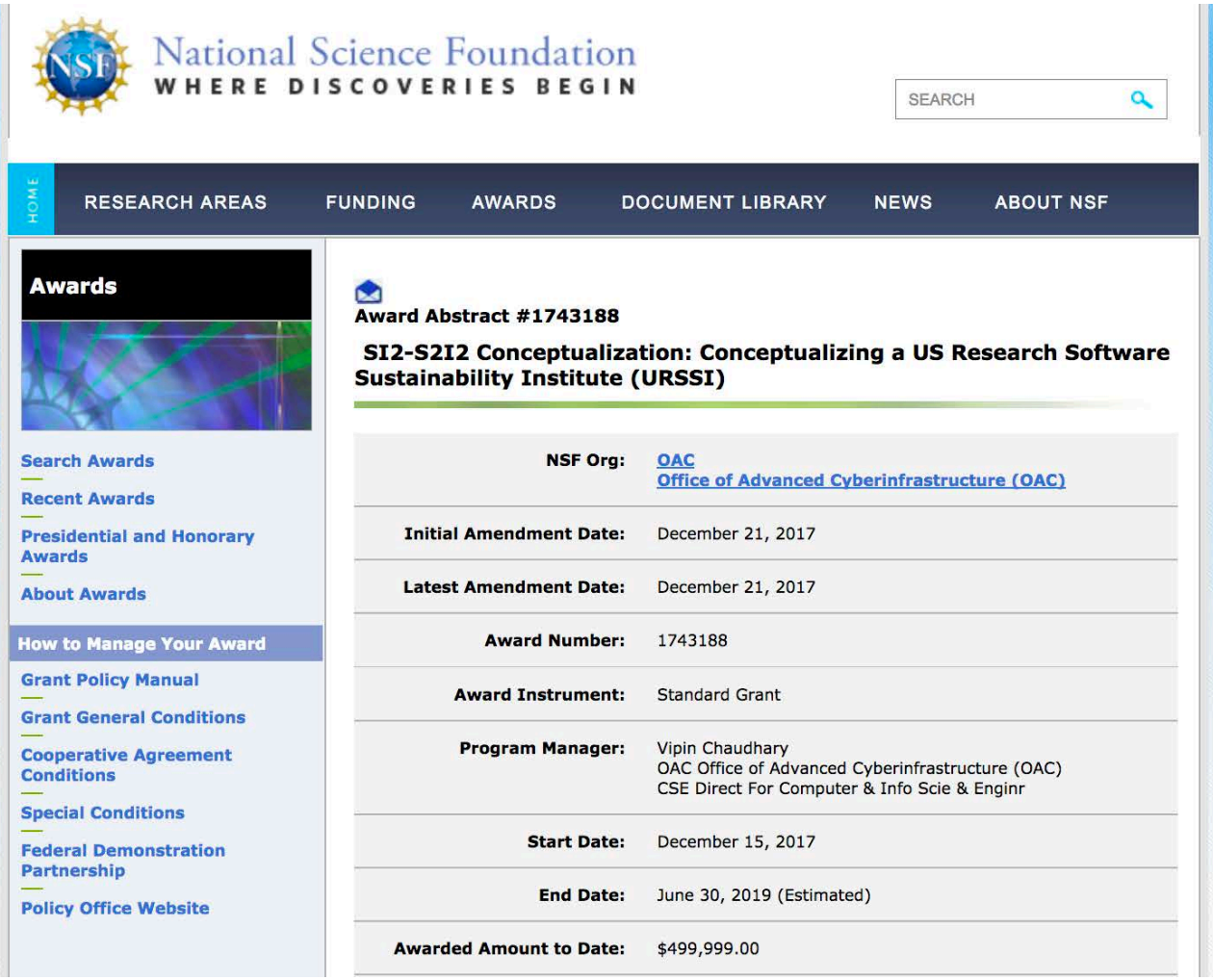
- https://sciencegateways.org/engage/student-focused

Webinars, blogs, newsletter, Twitter, LinkedIn etc.

https://sciencegateways.org

SGCI Science Gateways Community Institute

Remember Henry's Question: How much time Do You Get to Learn Your Job as Facilitator?



National Science Foundation
WHERE DISCOVERIES BEGIN

SEARCH

HOME | RESEARCH AREAS | FUNDING | AWARDS | DOCUMENT LIBRARY | NEWS | ABOUT NSF

Awards

Search Awards
Recent Awards
Presidential and Honorary Awards
About Awards

How to Manage Your Award

Grant Policy Manual
Grant General Conditions
Cooperative Agreement Conditions
Special Conditions
Federal Demonstration Partnership
Policy Office Website

Award Abstract #1743188

SI2-S2I2 Conceptualization: Conceptualizing a US Research Software Sustainability Institute (URSSI)

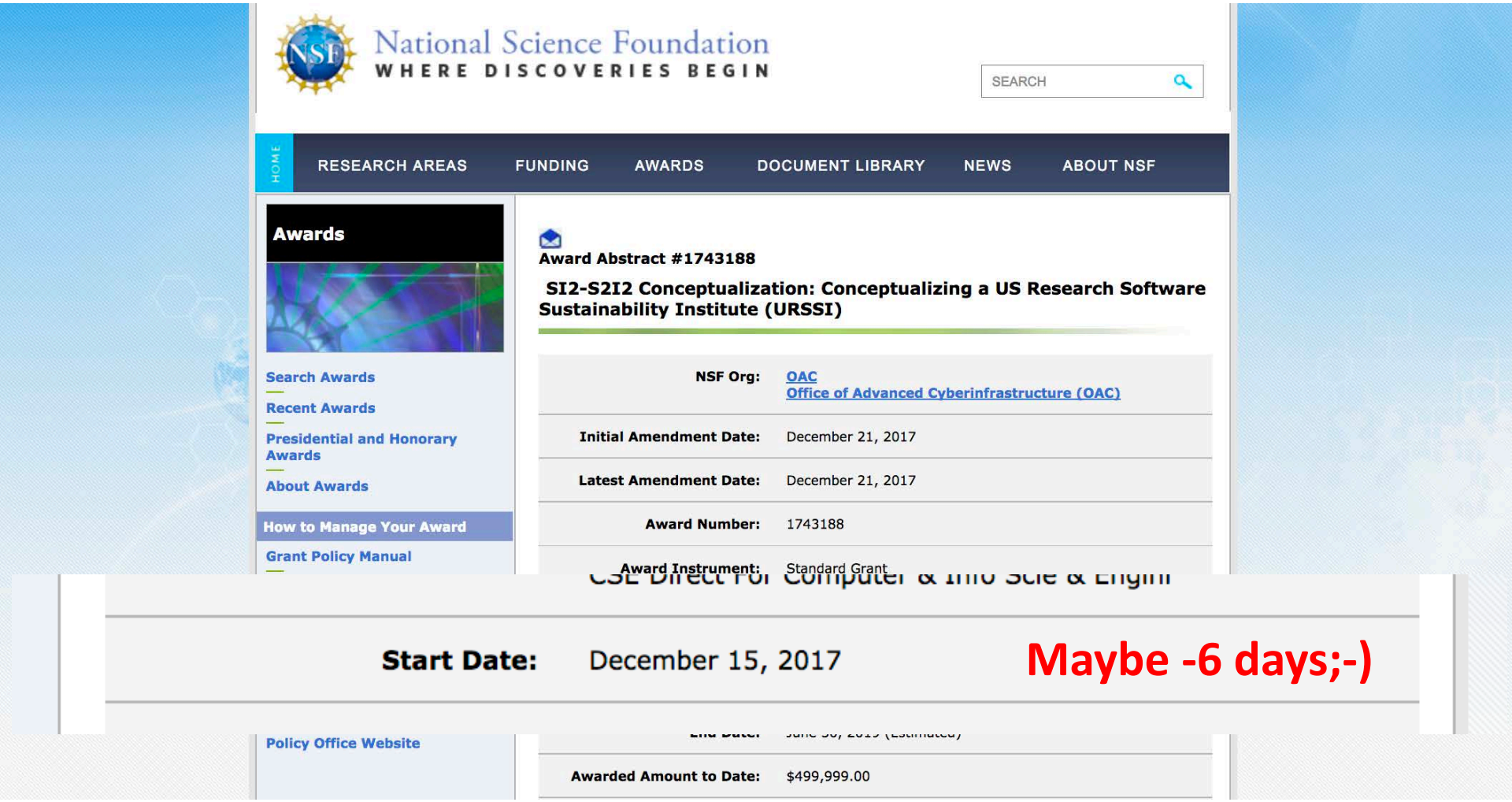
NSF Org:	OAC Office of Advanced Cyberinfrastructure (OAC)
Initial Amendment Date:	December 21, 2017
Latest Amendment Date:	December 21, 2017
Award Number:	1743188
Award Instrument:	Standard Grant
Program Manager:	Vipin Chaudhary OAC Office of Advanced Cyberinfrastructure (OAC) CSE Direct For Computer & Info Scie & Enginr
Start Date:	December 15, 2017
End Date:	June 30, 2019 (Estimated)
Awarded Amount to Date:	$499,999.00

SGCI Science Gateways Community Institute

CRC CENTER for RESEARCH COMPUTING

URSSI US Research Software Sustainability Institute

Remember Henry's Question: How much time Do You Get to Learn Your Job as Facilitator?



National Science Foundation
WHERE DISCOVERIES BEGIN

SEARCH

HOME | RESEARCH AREAS | FUNDING | AWARDS | DOCUMENT LIBRARY | NEWS | ABOUT NSF

Awards

Search Awards
Recent Awards
Presidential and Honorary Awards
About Awards

How to Manage Your Award
Grant Policy Manual

Award Abstract #1743188

SI2-S2I2 Conceptualization: Conceptualizing a US Research Software Sustainability Institute (URSSI)

NSF Org:	OAC Office of Advanced Cyberinfrastructure (OAC)
Initial Amendment Date:	December 21, 2017
Latest Amendment Date:	December 21, 2017
Award Number:	1743188
Award Instrument:	Standard Grant

CSE Direct for Computer & Info Scie & Engrng

Start Date: December 15, 2017

Maybe -6 days;-)

Policy Office Website

Awarded Amount to Date:	$499,999.00

SGCI Science Gateways Community Institute

CRC CENTER for RESEARCH COMPUTING

URSSI US Research Software Sustainability Institute

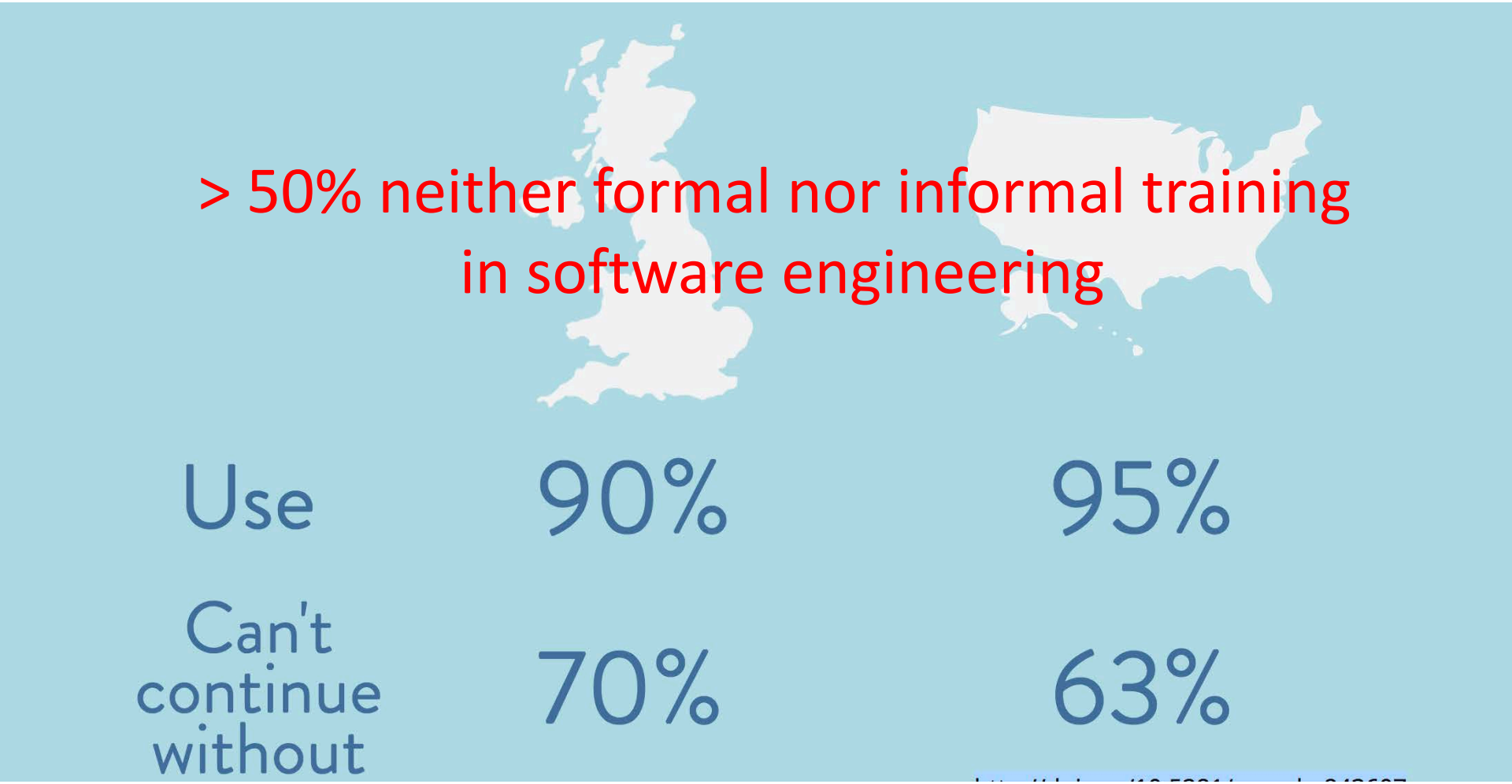
Research Software

Use	90%	95%
Can't continue without	70%	63%

http://doi.org/10.5281/zenodo.843607

URSSI US Research Software Sustainability Institute

Research Software



http://doi.org/10.5281/zenodo.843607

US Research Software Sustainability Institute

Research Software

Lack of career paths

Use	90%	95%
Can't continue without	70%	63%

http://doi.org/10.5281/zenodo.843607

US Research Software Sustainability Institute

Research Software

How to cite software?

Use	90%	95%
Can't continue without	70%	63%

http://doi.org/10.5281/zenodo.843607

US Research Software Sustainability Institute

Areas of Concern

- Functioning of the individual and team
- Functioning of the research software
- Functioning of the research field itself

US Research Software Sustainability Institute

Functioning of the Individual and Team

- Training & education
- Ensuring appropriate credit for software development
- Enabling publication pathways for research software
- Fostering satisfactory and rewarding career paths for people who develop and maintain software
- Increasing the participation of underrepresented groups in software engineering

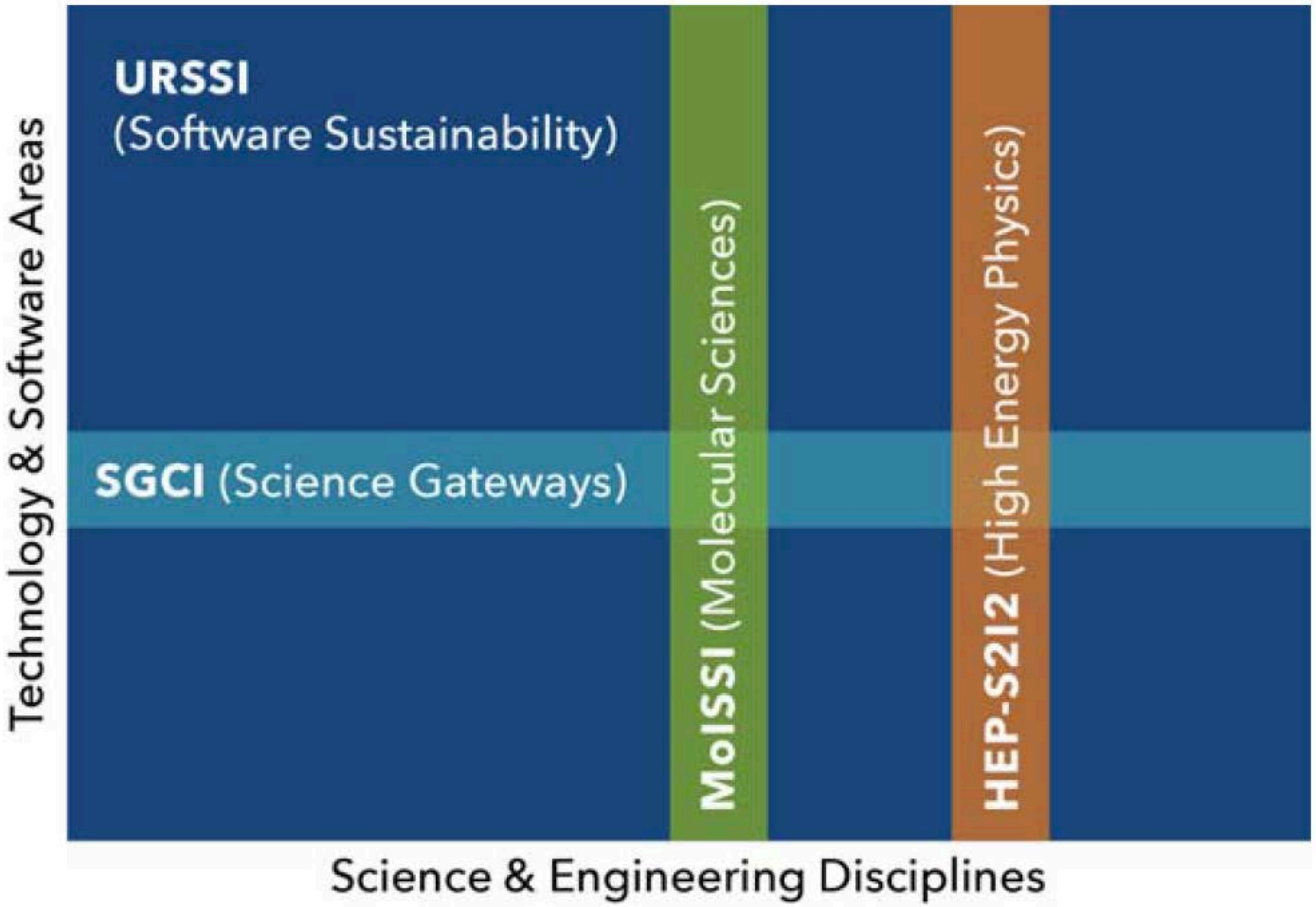
Functioning of Research Software

- Supporting sustainability of the software
- Growing community, evolving governance, and developing relationships between organizations, both academic and industrial
- Fostering both testing and reproducibility
- Supporting new models and developments (e.g., agile web frameworks, Software-as-a-Service)
- Supporting contributions of transient contributors (e.g., students)

Functioning of the Research Field Itself

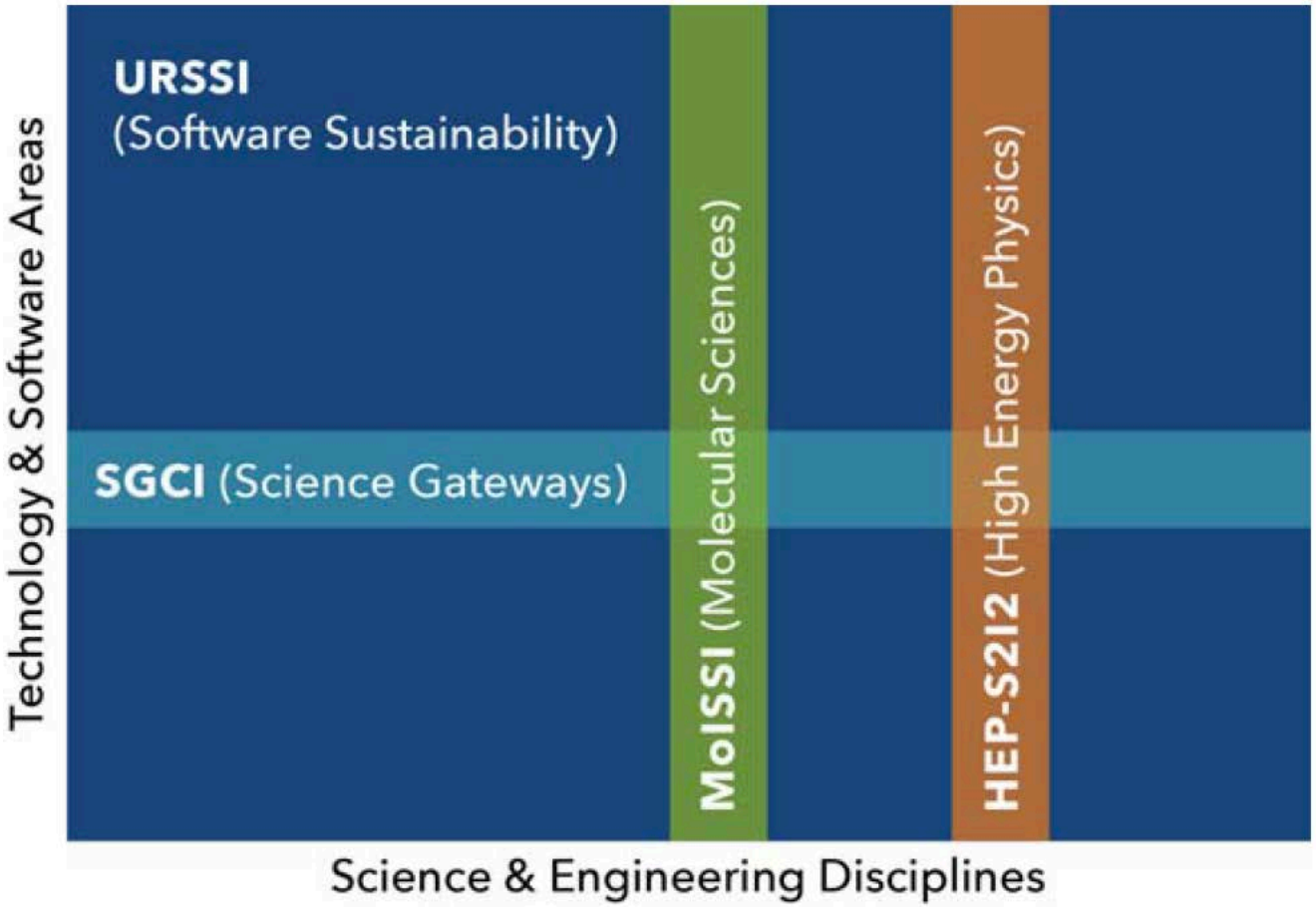
- Growing communities around research software and disparate user requirements
- Cataloging extant and necessary software
- Disseminating new developments
- Training researchers in the usage of software
- Understanding and improving pipelines of diverse developers and maintainers

URSSI and Other S2I2 Projects



US Research Software Sustainability Institute

URSSI and Other S2I2 Projects



Goal: Close collaboration and fill in gaps on each axis

URSSI US Research Software Sustainability Institute

Conceptualization

- Workshops
  - First workshop took place in April in Berkeley

US Research Software Sustainability Institute

Conceptualization

- Workshops
  - First workshop took place in April in Berkeley
  - Next workshop will take place in October in Chicago
  - Software credit workshop
  - Incubator workshop
- Survey
- Ethnographic studies
- Mission and vision working group

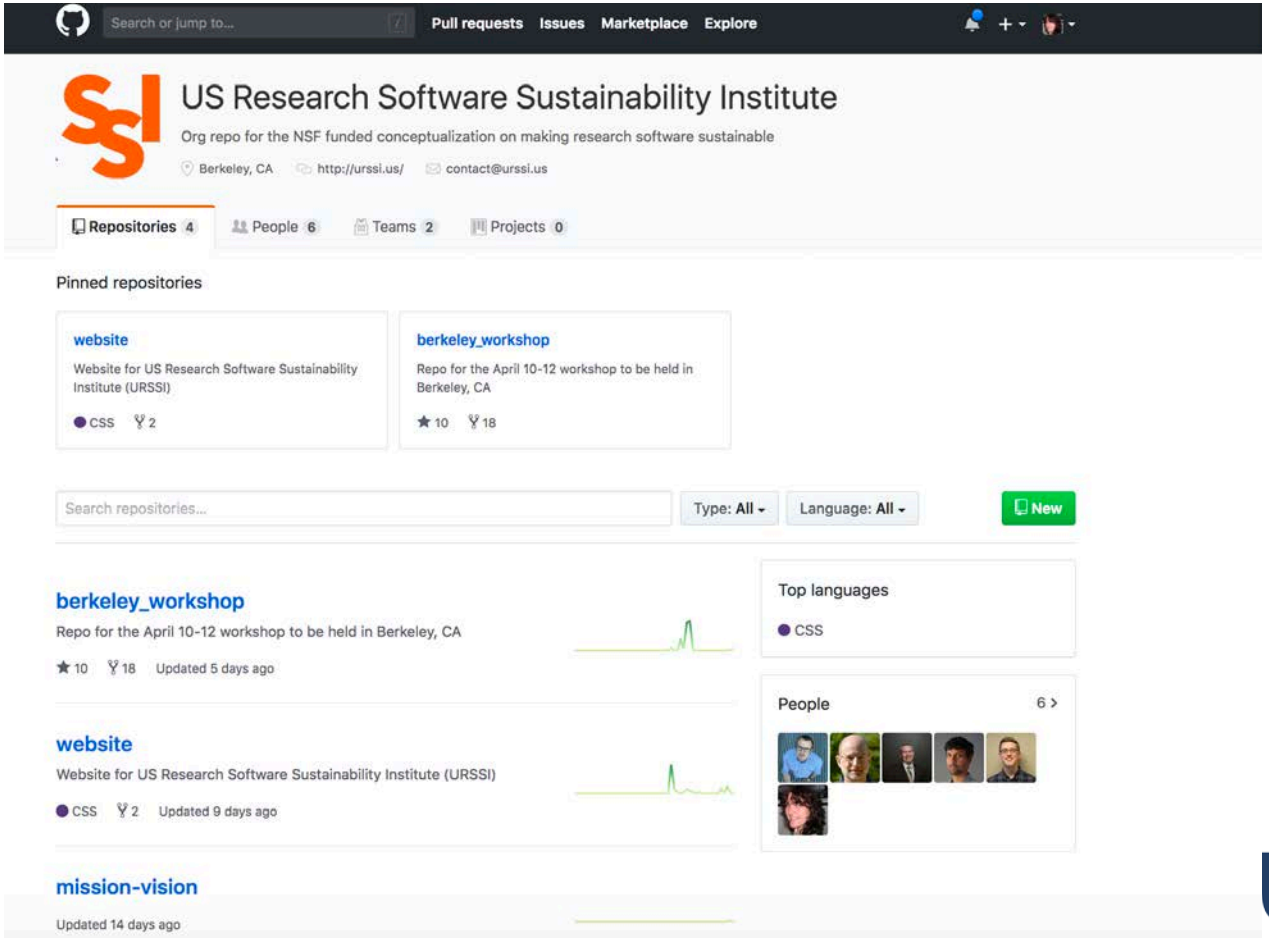
How to Connect

- Website http://urssi.us/

US Research Software Sustainability Institute

How to Connect

- Website http://urssi.us/
- Materials https://github.com/si2-urssi



US Research Software Sustainability Institute

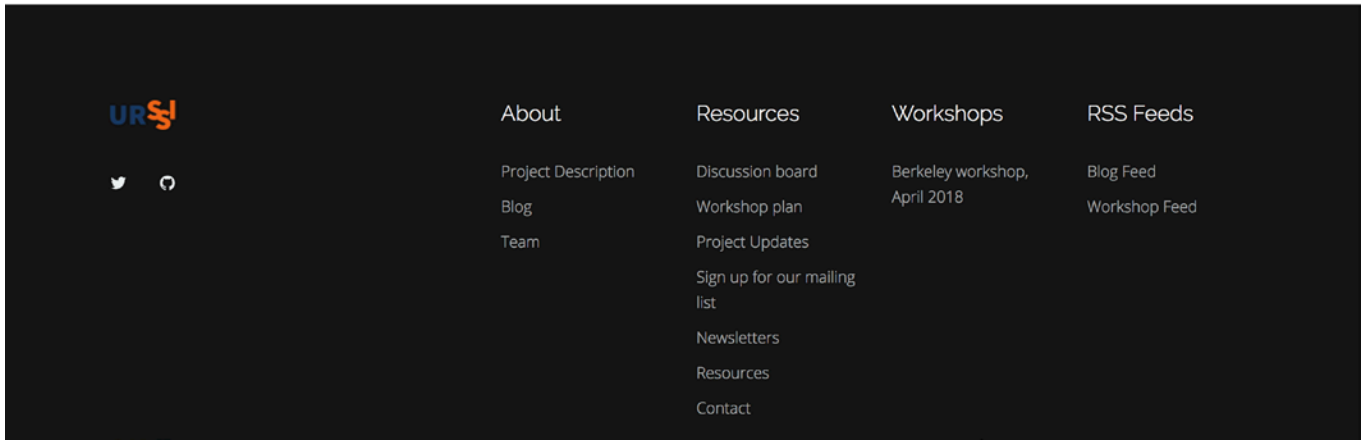
How to Connect

- Website http://urssi.us/
- Materials https://github.com/si2-urssi
- Blog posts http://urssi.us/blog/

URSSI US Research Software Sustainability Institute

How to Connect

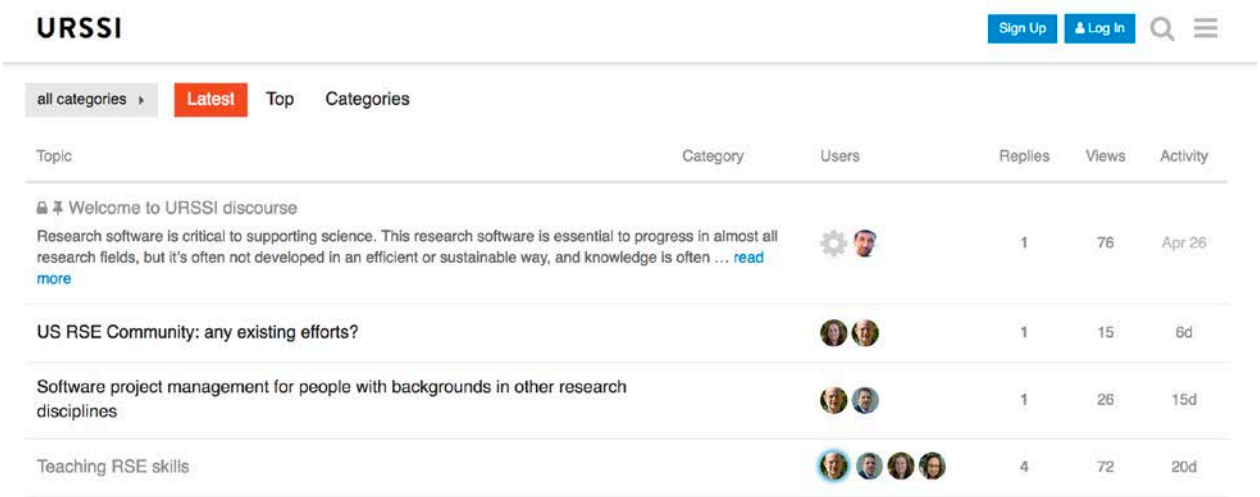
- Website http://urssi.us/
- Materials https://github.com/si2-urssi
- Blog posts http://urssi.us/blog/
- Mailing list http://urssi.us/



US Research Software Sustainability Institute

How to Connect

- Website http://urssi.us/
- Materials https://github.com/si2-urssi
- Blog posts http://urssi.us/blog/
- Mailing list http://urssi.us/
- Discuss https://discuss.urssi.us/



US Research Software Sustainability Institute

How to Connect

- Website http://urssi.us/
- Materials https://github.com/si2-urssi
- Blog posts http://urssi.us/blog/
- Mailing list http://urssi.us/
- Discuss https://discuss.urssi.us/
- Twitter https://twitter.com/si2urssi

How to Connect

- Website http://urssi.us/
- Materials https://github.com/si2-urssi
- Blog posts http://urssi.us/blog/
- Mailing list http://urssi.us/
- Discuss https://discuss.urssi.us/
- Twitter https://twitter.com/si2urssi
- Workshops http://urssi.us/workshops/

Join us for our next workshop in Chicago

October 23-24!

URSSI US Research Software Sustainability Institute

Lessons Learned on International Level

UK SSI and UK Research Software Engineer Association

- Buy-in from universities
- Viable career path
- Large community

Software Sustainability

Software Sustainability Institute

About | Programmes and Events

The importance of sustainability

Sustainability means that the software you use today will be available - and continue to be improved and supported - in the future.

Better science through superior software

Our work is focussed around four themes we believe are fundamental to doing research correctly in the digital age. These are related to **our manifesto**.

The first of these is **Skills and Training**: creating a capable research software community by enabling access to software development training for all researchers and teaching them methods to advance their research.

Recognition and Reward promotes and contributes to systems of credit for good software development and reuse practice.

Career Paths recognises and champions the varied job roles associated with research software; with a primary focus on the academic sector but suggesting industrial practice where applicable.

Finally, **Reproducible Research** promotes the fundamental place of software in supporting confidence in the research process and its results.

Taken together, these enable the efficient and effective use of software to tackle both the grand challenges that push the boundaries of human knowledge to day-to-day research software tasks.

https://www.software.ac.uk/about

SGCI Science Gateways Community Institute

CRC CENTER for RESEARCH COMPUTING

URSSI US Research Software Sustainability Institute

Sustainability

The Ecosystem offers a lot of opportunities

But

- Another mail list
- Another newsletter
- Another discussion list
- Another slack channel
- Plenty of workshops
- Plenty of surveys

Overwhelming? Too many scattered approaches?
How to sustain knowledge on and collaboration with sustainability approaches?

Henry's Travel Schedule

Date	Location	Event
Wed Feb 21 - Fri Feb 23	Dallas TX	Project meeting
Mon March 12 - Wed March 14	Alexandria VA	CaRC CI Professionalization Workshop
Thu March 15 - Fri March 16	Washington DC area	Review panel
Wed March 14 - Fri March 16	Alexandria VA	Coalition for Academic Scientific Computation spring meeting
Fri Apr 10	Oklahoma City OK	Cameron University Computing Technology Day Speaker
Tue Apr 17 - Wed Apr 18	Stillwater OK	Coalition for Advancing Digital Research & Education Conference
Wed May 16	Dallas TX	University of Texas Dallas Innovation & Technology Summit Speaker
Wed May 30 - Fri June 1	Kansas City MO	Great Plains Network GPN Annual Meeting
Sun July 22 - Thu July 27	Pittsburgh PA	PEARC 2018 conference
Sun Aug 5 - Fri Aug 10	Norman OK	Advanced Cyberinfrastructure Research & Education Facilitators Virtual Residency Workshop Facilitator
Mon Sep 17	Rolla MO	Missouri U of Science & Technology Research Technology Day Speaker
Tue Sep 25 - Wed Sep 26	Norman OK	Oklahoma Supercomputing Symposium 2018 Conference Chair/Speaker
Sun Nov 11 - Thu Nov 15	Dallas TX	Supercomputing 2018

Lessons Learned on International Level

What is different in the UK?

- It's a smaller country - not a characteristics we can change for the US
- The culture between research, software engineering and facilitation is different (SE and facilitation are more appreciated by researchers) – we can work on this, the more positive examples, white papers, metrics, etc. the more buy-in…

Lessons Learned from Sustained Teams/Projects

Commonalities

- Evangelist
- Diverse mechanisms of funding
- Community building
- Open source and open science
- Collaboration, collaboration, collaboration

What Are Our Next Steps?

- Evangelists for diverse initiatives – not every approach suits all
- Define a roadmap for collaboration and community building

What Are Our Next Steps?

How to make it less scattered and without duplicating effort? How to change research culture?

- Catalog on sustainability projects?
- Catalog on events?
- Catalog on metrics?
- Catalog on success stories?
- Evangelist and an outreach specialist?
- “Exchange” instructors between events?
- Involve faculty, HR people, stakeholders in events?

What Are Our Next Steps?

Let's make the next steps together!

Thanks!

sandra.gesing@nd.edu