Connecting people and resources to accelerate discovery by empowering the science gateway community

SGCI

Science Gateways: Sustainability via On-Campus Teams

***Sandra Gesing**, Katherine Lawrence, Nancy Wilkins-Diehr, Maytal Dahan, Michael Zentner, Marlon Pierce*

sandra.gesing@nd.edu

Gateways 2017, Ann Arbor, MI, USA

October, 24 2017

Award Number
ACI-1547611

Science Gateways Survey 2014

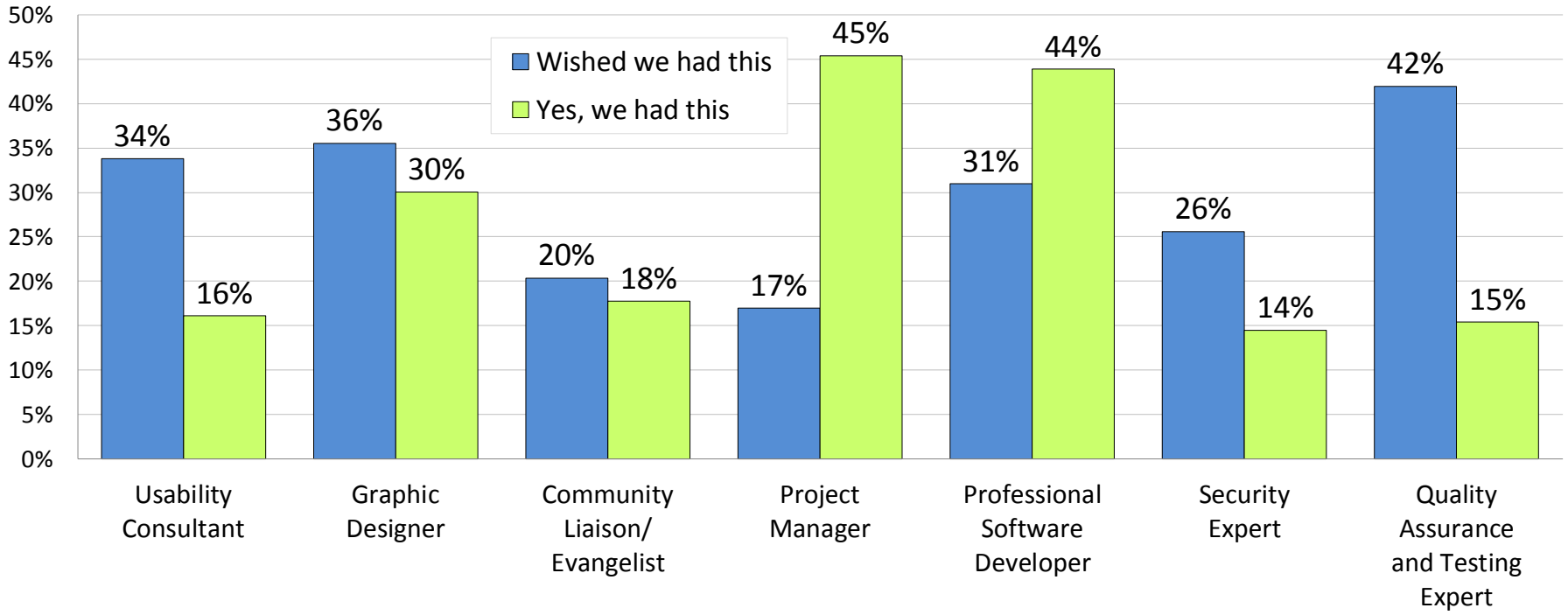
- sent out to 29,000 persons
- 4,957 responses from across domains
- 52% from life, physical or mathematical sciences
- 32% from computer and information sciences or engineering
- 45% develop data collections
- 44% develop data analysis tools

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%

Science Gateways Survey 2014

Well-designed gateways require a variety of expertise



Opportunities

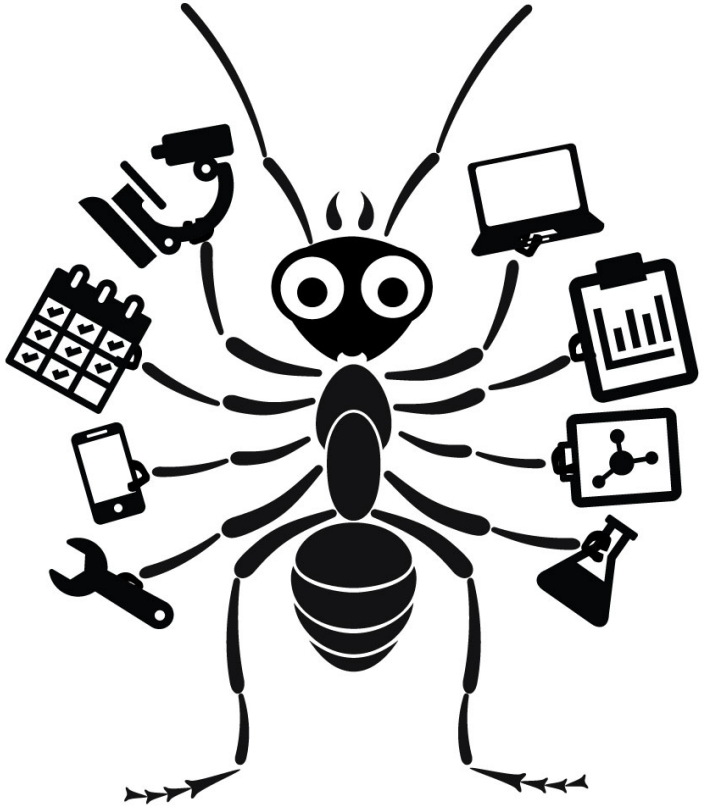
- A breadth of interesting topics in the science gateway creation process
  - Novel frameworks and web applications
  - Inter- and multidisciplinary work
  - Contributing to grand challenges, e.g., Malaria eradication
  - …
- A breadth of interesting roles
  - Designers, statisticians, librarians, machine learning experts, usability experts,…

Challenges

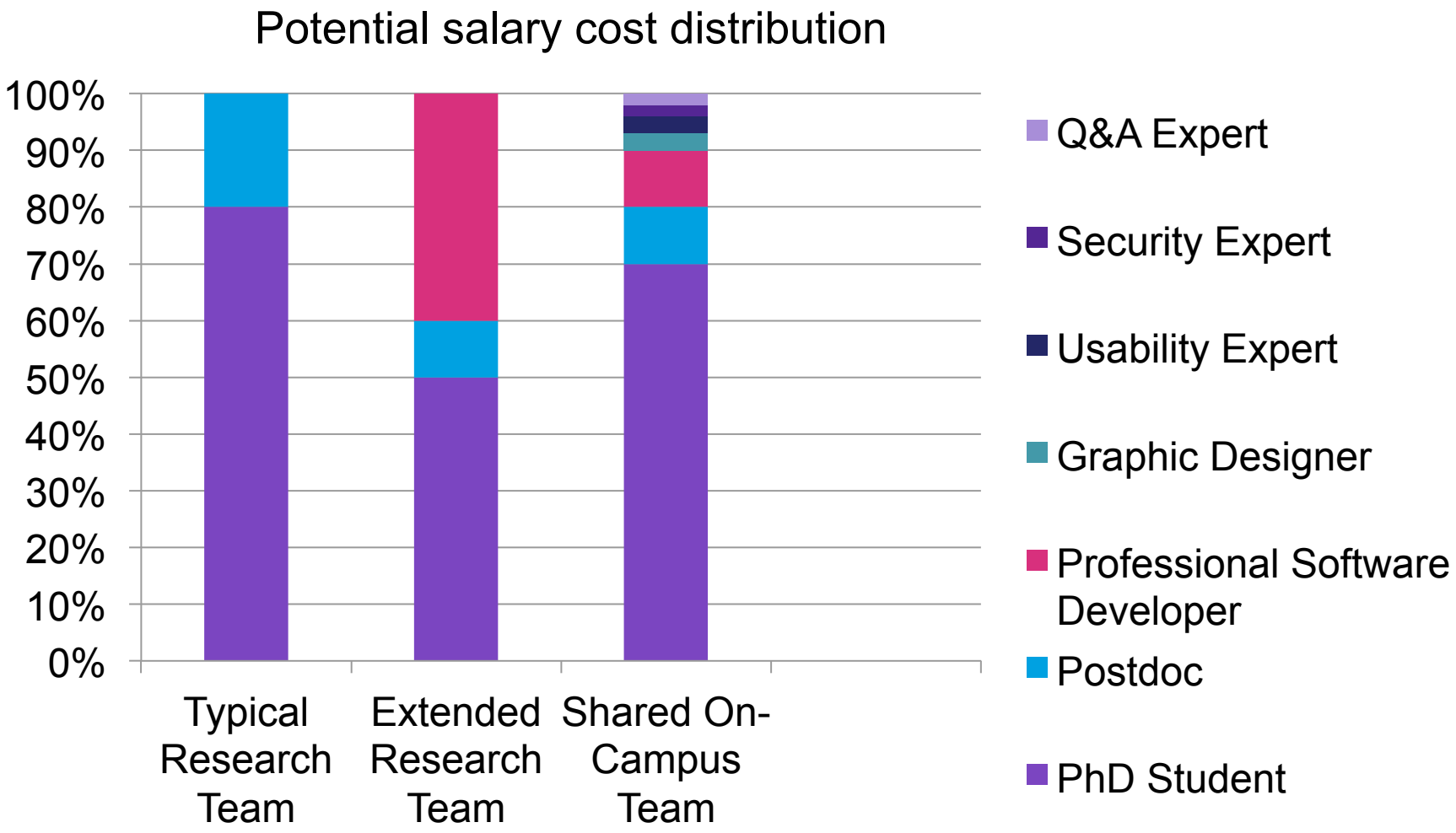
- Some topics and roles are only needed for part of the project

⇒ Not fully funded positions via one project

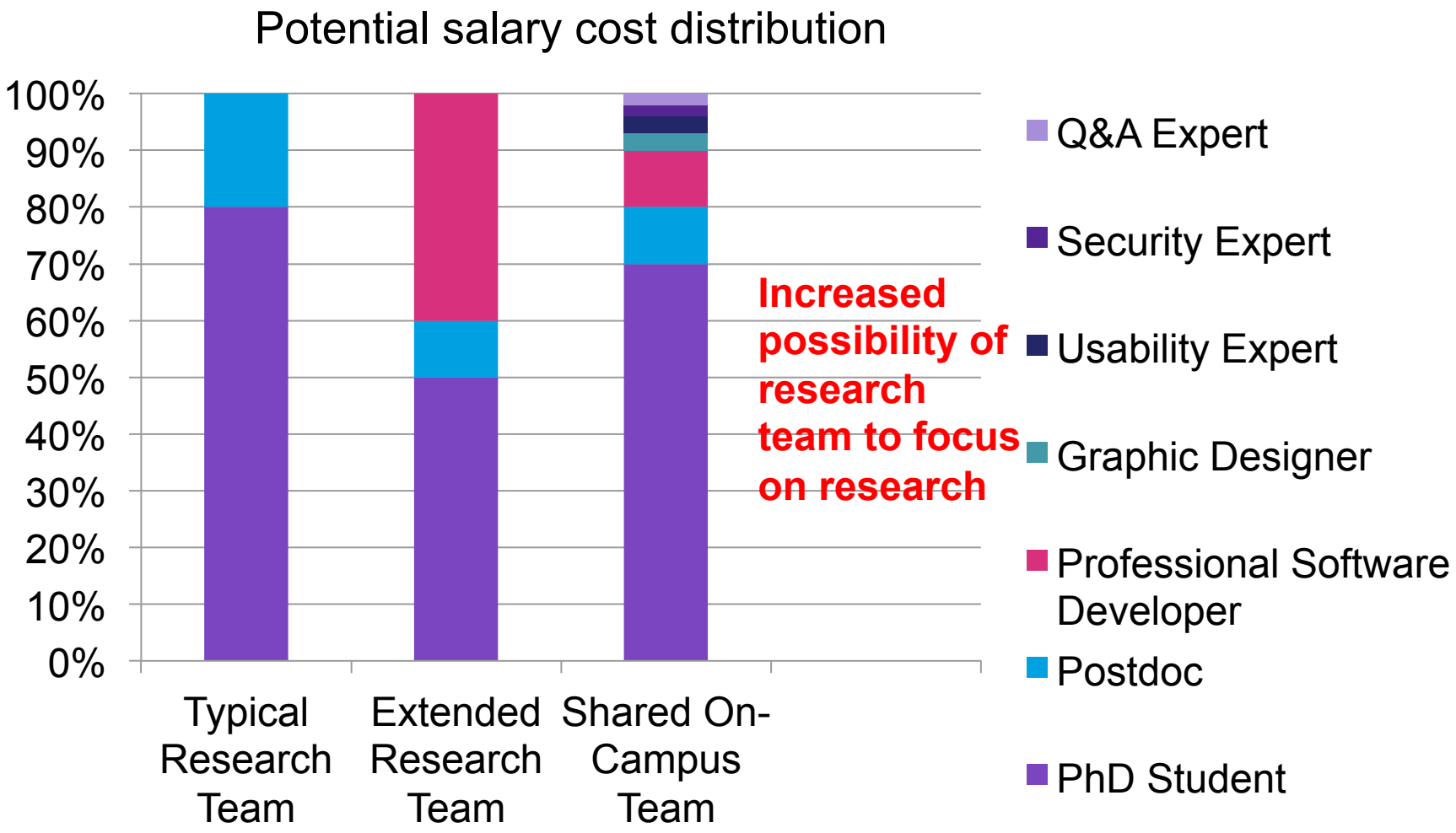
⇒ For diverse expertise, several different people are needed



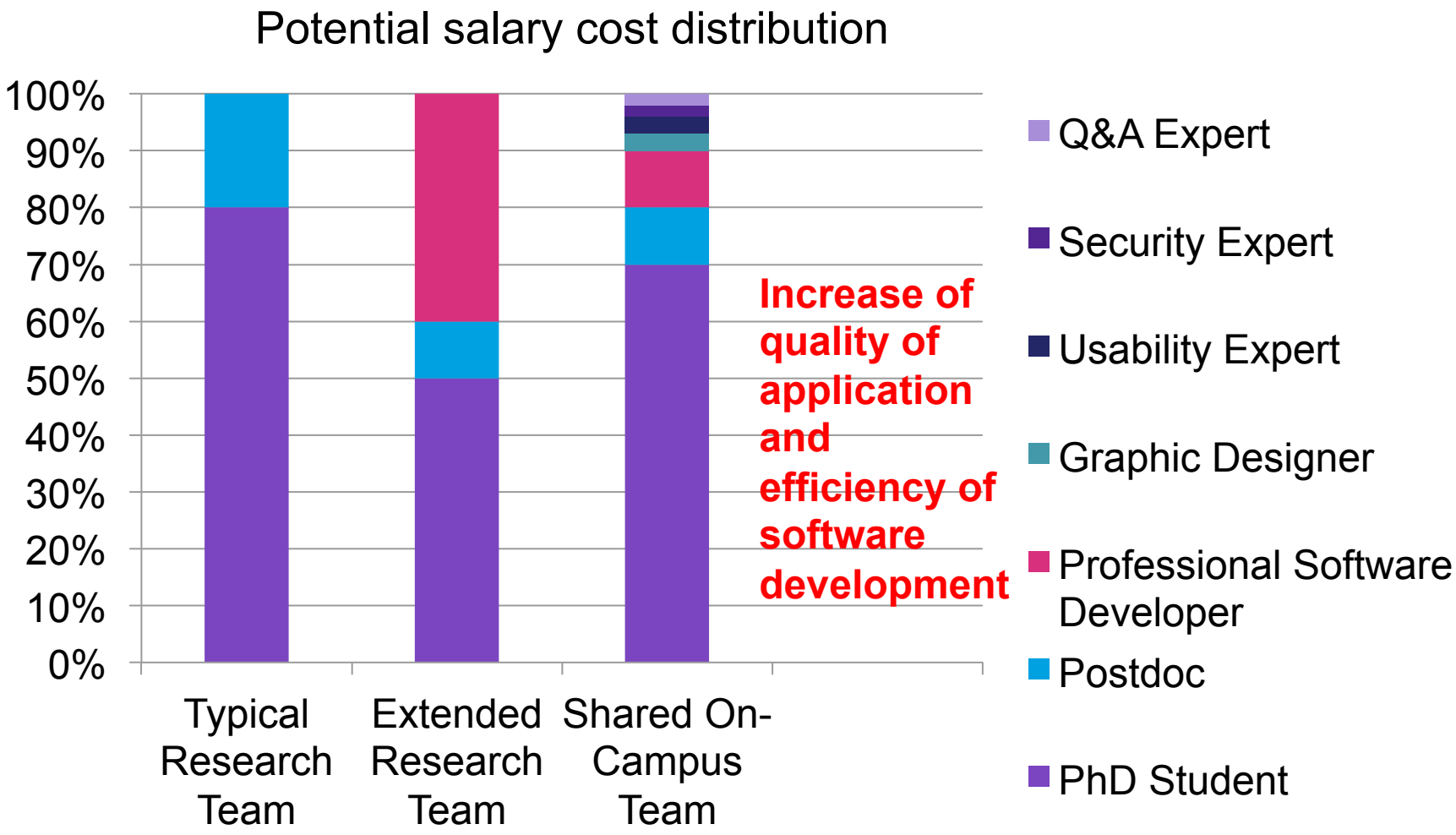
Sustainability via On-Campus Teams



Sustainability via On-Campus Teams



Sustainability via On-Campus Teams



Sustainability via On-Campus Teams

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

Success Stories

Many universities have successful centers/groups with centralized services for science gateways

- Center for Research Computing at University of Notre Dame
- HUBzero® Team at Purdue University
- Science Gateways Research Center at Indiana University
- Science Gateway Group at TACC at the University of Texas, Austin
- …

ND CRC in 2006-2008

- Effort with 7 FTEs centrally funded
- HTC and HPC Computing and basic user support
  - One centrally funded cluster plus multiple faculty funded clusters in various cabinets on campus
  - Around 300 active users
  - 80% centrally funded hardware
- No other kinds of research computing services
- Underserved social sciences and humanities for their need on science gateways
- ~1000 faculty, ~12000 students

ND CRC Vision

To become an internationally recognized multidisciplinary research computing center based upon our reputation for facilitating and accelerating discovery through effective and novel applications of cyberinfrastructure.

ND CRC Director's thoughts on Vision

- "Nice vision, but how we get there?"
- "What should we do first?"
- "Users should tell us what they need…"
- "HPC works fine, so let's focus on CI Development…"
- "We need portals and other collaborative environments"
- "We need good user support, and good, straight resource usage policies"
- "We need CI and HPC teams working together"
- "How do we fund all these with very limited resources?"
- "What is available out there?
- And so on…

Implementation by Jarek Nabrzyski

Take the risk and hire people first

➔ train people

➔ generate / bring projects

➔ assign people to projects and focus on getting more projects and more people -> etc...

➔ if not successful then exit, i.e. quit ☺

Fortunately, we had many projects, mostly science gateway projects of various kinds and difficulty.

ND CRC in 2017

- 45 FTEs with 70% of the staff salaries supported through grants and services
- HTC and HPC Computing and user support
  - 26,000 cores
  - 1,800 active users
  - 10 HPC engineers and user support staff
  - 30% of compute nodes are centrally funded
- Cyberinfrastructure development
  - ~15-20 CI projects each year with ~35 faculty from various departments including social sciences and humanities
  - supported by ~15 research programmers, ~8 computational scientists, some FTE fractions of HPC engineers, and a few (7) grad students and undergraduate interns (4-6)
- ~1000 faculty, ~12000 students

HUBzero History: Operating Model Path

- 1996: PUNCH (precursor to nanoHUB)
- 2002: NSF funding for nanoHUB
- 2007: Spin-out of HUBzero from nanoHUB
- 2015: Diversification, self-funded sustainability
- Composition:
  - Hubs hosted by Purdue (for a variety of institutions)
  - Foundation members running their own hubs
  - Open source hubs

HUBzero History: Key Team Enablers

- 1996: Vision (Mark Lundstrom)
- 2002: Vertical Expansion (Gerhard Klimeck)
- 2007: Horizontal Expansion (Michael McLennan)
- 2015: Business Transition (Michael Zentner)
- Composition:
  - Hubs hosted by Purdue (for a variety of institutions)
  - Foundation members running their own hubs
  - Open source hubs

HUBzero in 2017

- A foundation, a group at Purdue, and a software framework
- 25 full time software professionals (+5 vacancies) specializing in:
  - Cybersecurity
  - Web programming
  - User experience design
  - Scientific application development
  - Analytics
  - Middleware
  - High performance computing
  - System administration
  - Customer service
- Servicing over 2 million total visitors annually
- Entirely self funded

Even with success we must continue to innovate new business models, partnerships, delivery capabilities, and service architectures.

SGCI

IU Science Gateway Group, 2012

SGCI

Science Gateways Research Center, 2017

Marlon Pierce

marpierc@iu.edu

Suresh Marru

smarru@iu.edu

Sudhakar Pamidighantam

pamidigs@iu.edu

Eroma Abeysinghe

eabeysin@iu.edu

Jun Wang

wang208@iu.edu

Marcus Christie

machrist@iu.edu

Supun Nakandala

snakanda@iu.edu

Plus three student RA's
+
Consultants

SGCI

Lesson: Cultivate Your Pipeline

- You are going to have staff turnover.
  - The IU SGRC has only three team members in common from five years ago.
  - Team size is constant
- Plug into your talent reservoir
  - The IU SGRC has invested a lot of effort in recruiting students
  - Teaching classes in the CS department
  - Mentoring Google Summer of Code students
  - Successful students spread the word
- Put effort into your onboarding process
  - Get people contributing to your software and projects as quickly as possible
- Keep alumni involved
  - Developer mailing lists for example
  - Potentially hire as consultants

TACC Gateway Groups

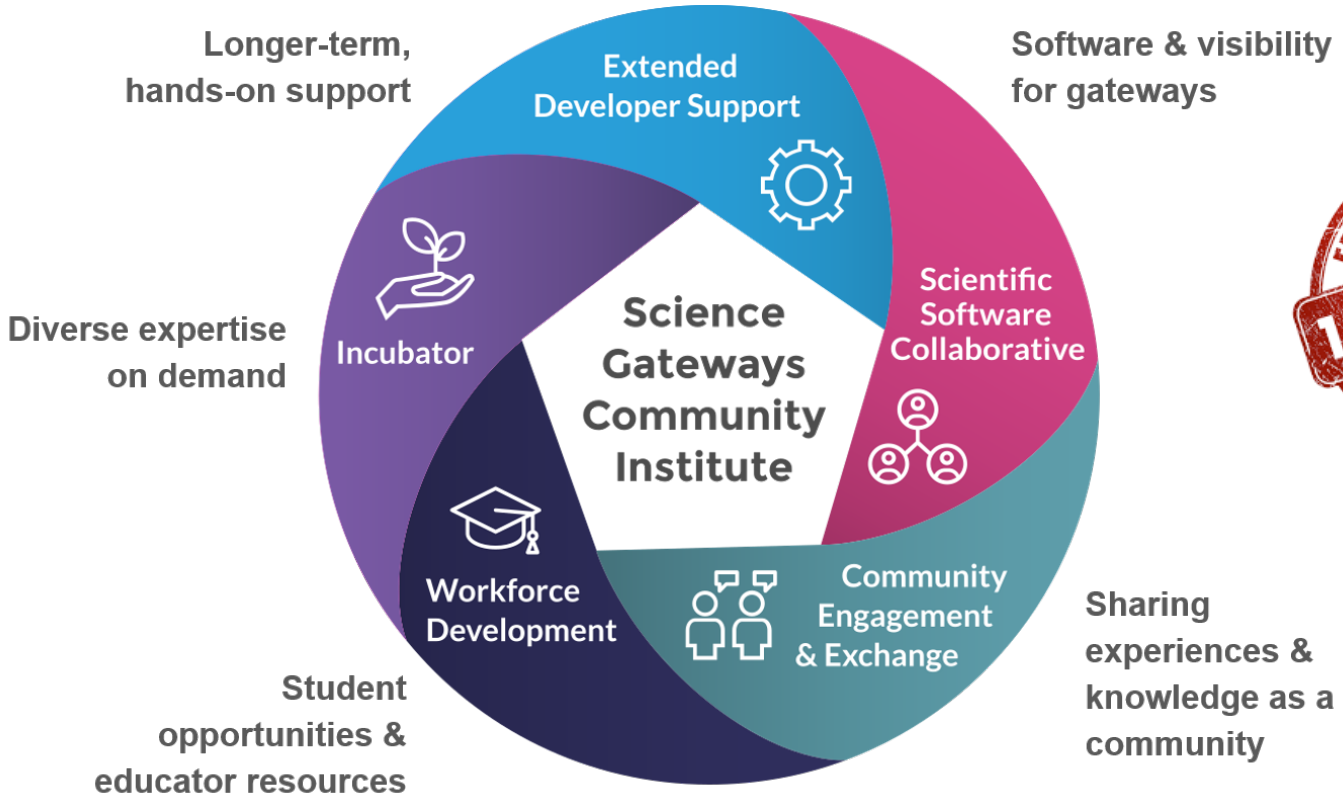
- TACC has an Advanced Computing Interfaces Department
  - 3 sub-groups that handles development of science gateways and supported backend infrastructure
  - Support multiple projects with various requirements and different platforms and development requirements
  - Wide variety of projects from cyberinfrastructure to specific scientific domains, humanities, and more.
- Started in 2002 with 3 staff members and now has 14+ full time staff members, undergraduate interns, and support from other departments.
- Participated in over 15 proposals in the past 12 months.

TACC Gateway Groups

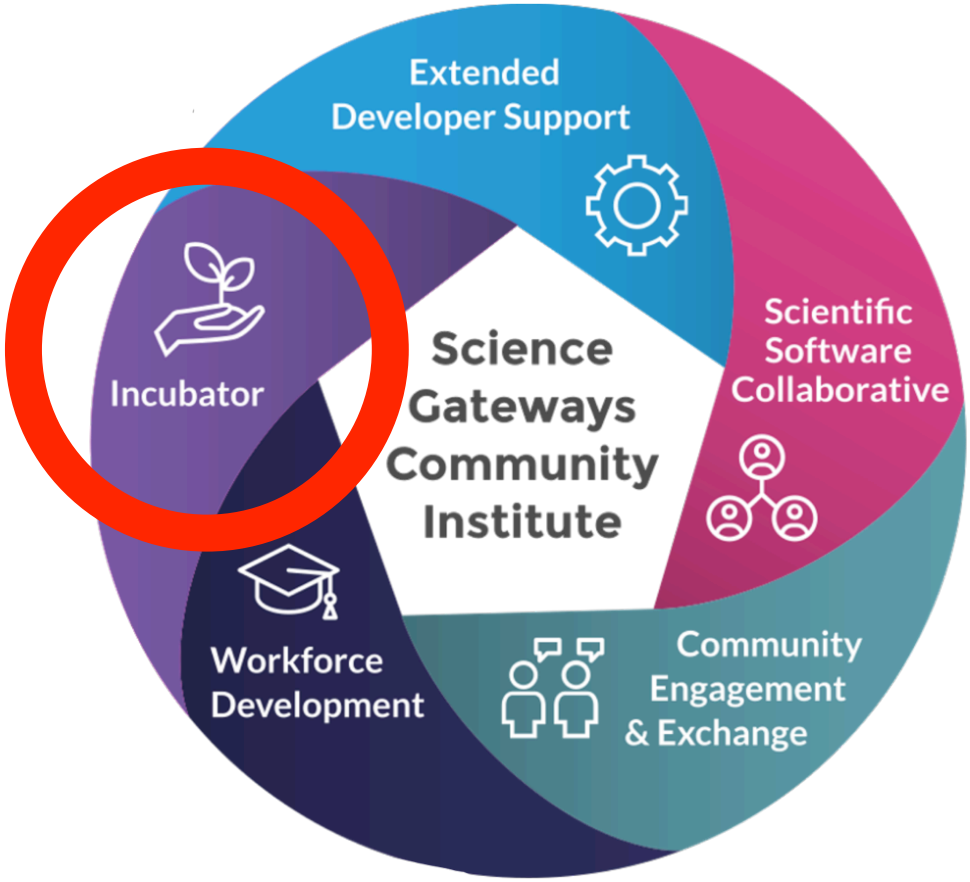
- Cuts across IT and research by handling the needs of gateway development from the resources to the User Interface
- Benefit of a team of experts in the field of gateway development
- Can more easily and effectively spin up new projects and more easily be time-sliced than faculty with a longer return on investment than traditional IT.
- Long-term investment benefits the entire organization by offering:
  - Earlier awareness of disruptive change.
  - Front-line reports about changing research tech needs
  - Non-traditional partnership opportunities
  - Built-in evangelism for standards, best practices, and software use
  - Frequent multi-domain and multi-department collaborative activities.

Funding Mechanisms – External Resources

Now is the right time – Get support from SGCI via diverse services!



Science Gateways Community Institute



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

- Diverse expertise on demand
- Longer term support engagements
- Software and visibility for gateways
- Information exchange in a community environment
- Student opportunities and more stable career paths

Incubator Service

A Framework for Decision Making

Technology Planning
- Choosing technologies
- Cybersecurity
- Software engineering
- Interfaces to compute and data

Business Planning
- Business model development
- Financial planning
- Project management
- Software licensing
- Staff and sustainability planning

Client Interaction Planning
- Usability studies
- Web/visual/graphic design
- Impact measurement
- Community engagement
- Support for education

Specialized Expertise

Security
- Center for Trustworthy Scientific Cyberinfrastructure

Sustainability
- Nancy Maron, creator of the ITHAKA S+R course on Sustaining Digital Resources

Evaluation & Impact Measurement
- Ann Zimmerman Consulting

Campus Resource Development

Network / Cohort Formation

Common Experiences
- Training sessions
- Group interactions

Continuing Engagement
- Customized structure, content, goals
- Mentoring
- Pay It Forward

An Ongoing Dispassionate Ear

SGCI

Bootcamp at a Glance

- 5 full days
- Teams on science gateways
- Knowledge dissemination
- Interactivity
- Community formation
- Putting away the normal daily routine
- Homework

Funding Mechanisms – External Resources

- Get support from SGCI via diverse services
- Contractors from mature science gateway frameworks

⇒ Means to create and/or maintain science gateway(s) successfully while working on building up internal resources

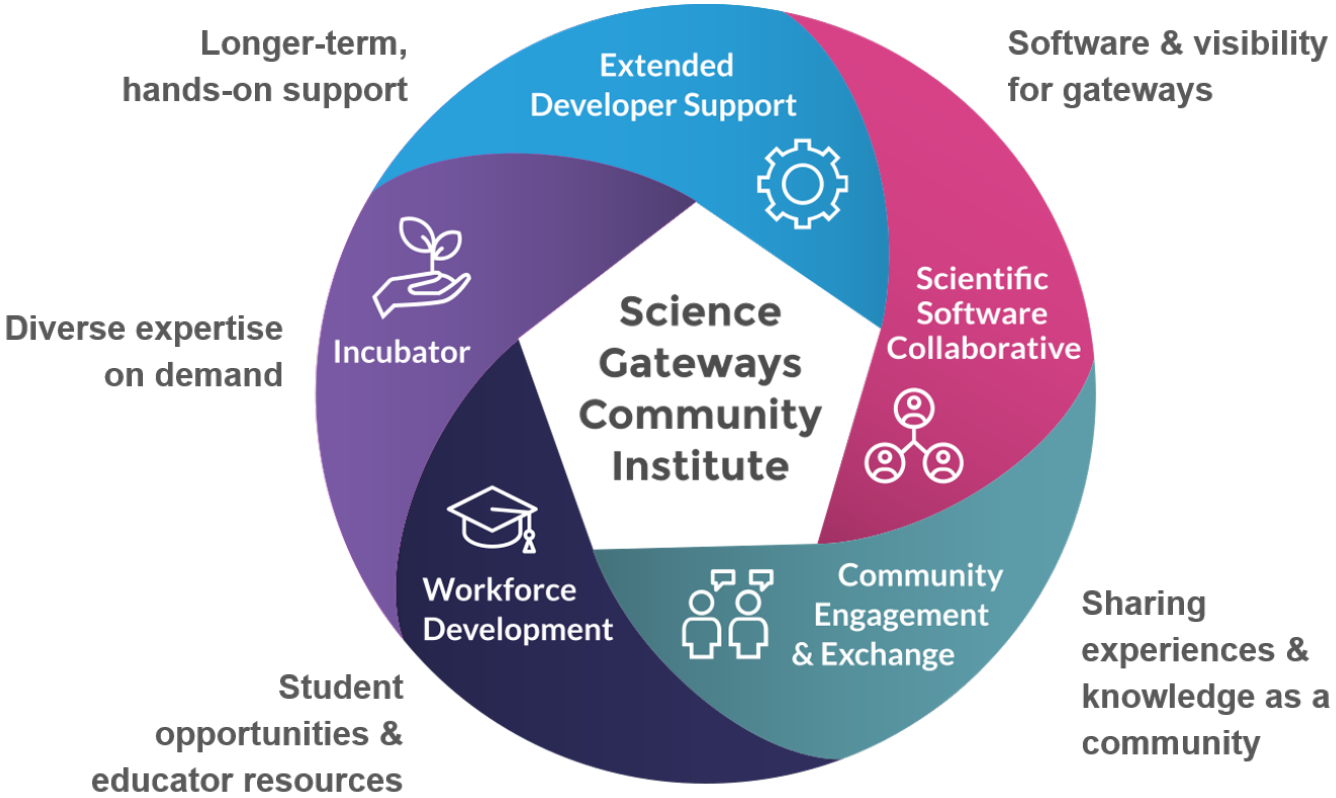
Funding Mechanisms – Internal Resources

- Funding via involvement of some person months in grants
- Funding on some hard money from universities
- Funding via re-charge
- Funding via NSF for building cyberpractitioner career-paths (under development - models like full funding the first year and increasing responsibility on the side of universities)

Using Existing Free On-Campus 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 an HPC center? They know how to access HPC resources, VMs, containerization, distributed data management …

Thanks for your attention!



help@sciencegateways.org

http://sciencegateways.org/