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

Science Gateways Community Institute: Subsidized Services and Consultancy to Facilitate Research on Your Campus

Sandra Gesing
Center for Research Computing, University of Notre Dame
sandra.gesing@nd.edu

Advanced Cyberinfrastructure Research & Education Facilitators Virtual Residency 2017, University of Oklahoma
August 4, 2017

Award Number
ACI-1547611
Science Gateways

- Increased complexity of today’s research questions
  - hardware and software
  - skills required
- Greater need for openness and reproducibility
  - Science increasingly driving policy questions
- Opportunity to integrate research with teaching
  - Better workforce preparation

We need end-to-end solutions that provide broad access to advanced resources and allow all to tackle today’s challenging science questions.
It’s a Fan!

It’s a Spear!

It’s a Wall!

It’s a Snake!

It’s a Tree!

It’s a Rope!
It's a Science Gateway

It's a Research Portal

It's a Virtual Research Environment

It's a Cyber-infrastructure

It's a Collaboratory

It's a Virtual Lab
Gateway users are 77% of active XSEDE users in Q4 2016

This is largely due to the CIPRES and I-TASSER gateways, but others are gaining
Typical Lifecycle of a Science Gateway

Early adopters
Publicity
Wider adoption
Funding ends
Scientists disillusioned
New project prototype
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?

<table>
<thead>
<tr>
<th>Proposed Service</th>
<th>% Interest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evaluation, impact analysis, website analytics</td>
<td>72%</td>
</tr>
<tr>
<td>Adapting technologies</td>
<td>67%</td>
</tr>
<tr>
<td>Web/visual/graphic design</td>
<td>67%</td>
</tr>
<tr>
<td>Choosing technologies</td>
<td>66%</td>
</tr>
<tr>
<td>Usability Services</td>
<td>66%</td>
</tr>
<tr>
<td>Visualization</td>
<td>65%</td>
</tr>
<tr>
<td>Developing open-source software</td>
<td>64%</td>
</tr>
<tr>
<td>Support for education</td>
<td>64%</td>
</tr>
<tr>
<td>Community engagement mechanisms</td>
<td>62%</td>
</tr>
<tr>
<td>Keeping your project running</td>
<td>62%</td>
</tr>
<tr>
<td>Legal perspectives</td>
<td>61%</td>
</tr>
<tr>
<td>Managing data</td>
<td>60%</td>
</tr>
<tr>
<td>Computational resources</td>
<td>59%</td>
</tr>
<tr>
<td>Mobile technology</td>
<td>59%</td>
</tr>
<tr>
<td>Database structure, optimization, and query expertise</td>
<td>59%</td>
</tr>
<tr>
<td>Data mining and analysis</td>
<td>58%</td>
</tr>
<tr>
<td>Cybersecurity consultation</td>
<td>57%</td>
</tr>
<tr>
<td>Website construction</td>
<td>57%</td>
</tr>
<tr>
<td>Software engineering process consultation</td>
<td>57%</td>
</tr>
<tr>
<td>Source code review and/or audit</td>
<td>51%</td>
</tr>
<tr>
<td>High-bandwidth networks</td>
<td>45%</td>
</tr>
<tr>
<td>Scientific instruments or data streams</td>
<td>44%</td>
</tr>
<tr>
<td>Management aspects of a project</td>
<td>38%</td>
</tr>
</tbody>
</table>
Science Gateways Survey 2014

Well-designed gateways require a variety of expertise

- Usability Consultant: Wished we had this - 34%, Yes, we had this - 16%
- Graphic Designer: Wished we had this - 36%, Yes, we had this - 30%
- Community Liaison/Evangelist: Wished we had this - 20%, Yes, we had this - 18%
- Project Manager: Wished we had this - 45%, Yes, we had this - 17%
- Professional Software Developer: Wished we had this - 31%, Yes, we had this - 44%
- Security Expert: Wished we had this - 26%, Yes, we had this - 14%
- Quality Assurance and Testing Expert: Wished we had this - 42%, Yes, we had this - 15%
Science Gateways Community Institute

- 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

### Network / Cohort Formation

#### Common Experiences
- Training sessions
- Group interactions

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

### An Ongoing Dispassionate Ear

---

10
Technologies

- Widely used complete frameworks (Galaxy, HubZero, Open Science Framework, Globus Data Portal etc.)
- RESTful APIs and support of multiple programming languages in widely used frameworks (Apache Airavata, the Agave platform, etc.)
- Reused interface implementations such as the one of CIPRES with its RESTful API (CIPRES has served more than 20,000 users to date)
- Science gateways as a service with provision of hardware in the background such as SciGap (Science Gateway Platform as a Service)

Lessons learned: approaches should be technology agnostic, using APIs and standard web technologies OR deliver a complete solution

Community Engagement is key

HubZero instances world wide
Usability

“After all, usability really just means that making sure that something works well: that a person … can use the thing - whether it's a Web site, a fighter jet, or a revolving door - for its intended purpose without getting hopelessly frustrated.”

(Steve Krug in “Don't make me think!: A Common Sense Approach to Web Usability”, 2005)
User-Centered Design

Goal?

- Ease of use
- Ease of learning
- Satisfaction

user-friendly  usability
User-Centered Design

What?

What are the types of artifacts we design?

interfaces    entire systems
User-Centered Design

Who?

Who are the users?
What are their goals? needs? mental models? expectations?
User-Centered Design

Goal: enable users to accomplish their goals with:

- Ease of use
- Ease of learning
- Satisfaction
User-Centered Design

Why?

- Ease of use
- Ease of learning
- Satisfaction
Unified Theory of Acceptance and Use of Technology (UTAUT)  
Venkatesh et al, 2003

- Will it help my job performance?
- Is it USEFUL?
- Is it better than the old way?

Direct determinants of intention

Moderators

Facilitating Conditions
Effort Expectancy
Social Influence
Performance Expectancy
Behavioral Intention
Use Behavior

Direct determinants for use behavior
Unified Theory of Acceptance and Use of Technology (UTAUT)
Venkatesh et al, 2003

- Easy to use?
- Easy to learn?
- Time-consuming?
- Efficient?
Unified Theory of Acceptance and Use of Technology (UTAUT)  
Venkatesh et al, 2003

- Can I do this? Do I have knowledge? Support? Resources?
- Does it fit in with my work style?
Diffusion of Innovations - Rogers, 1962

Perceived characteristics of the innovation:

- relative advantage
- compatibility
- complexity
- trialability
- observability

persuasion → decision

- adopt
- reject
User-Centered Design

Why?

Because usability and user experience are predictors of technology adoption.
User-Centered Design

What can I do?

- research
- test
- iterate
Usability Test

Spend one hour a week on a usability test and the increase of usability is overproportional!
Usability Test

Go to http://sciencegateways.org/

Find:
• How can you partner with us?
• How can you apply for developer services?
• How can you sign up for the email list?
Bootcamp at a Glance

- 5 full days
- Knowledge dissemination
- Interactivity
- Community formation
- Putting away the normal daily routine
- Homework
The lightbulb effect, who I can unify community request for support to create a funding request to campus leadership. I'm planning to organize the gateway community on our campus, including humanities, libraries, engineering and science and create a strategic budget request to our campus leadership.

For me, it was the realization of our team after the second day that we weren't in a black box of "where do we go from here", but could see a path to making our gateway sustainable in the next 5 years (50% nsf funding) and 10 years (0% nsf funding).

I was really impressed by the common challenges faced by such diverse projects. They were diverse in their maturity, goals, and disciplinary contexts and yet we all could learn from each other by addressing the shared challenges.
Extended Developer Support

**Focus**
- Front-end development
- Gateways using all types of CI
- Both sides give 2-month to 1-year commitment
- Well-defined engagements with work plans
- Technology agnostic

**Mission**
- Bring new gateways into existence
- Adapt existing gateways to new resources and technologies
- Provide “burst” support to help gateways with smaller issues

**Benefits**
- Reinforce Incubator lessons
- Develop deep understanding of community needs that feed into other Institute areas
- Capture and document support efforts for scalability
- Hands-on opportunities for student participants

---

**Data Instruments Analysis Tools**

**Education Computation**

**Collaboration Workflows**

**Sensor**

---

Airavata HUBzero AGAVE Jupyter Galaxy And more…
Scientific Software Collaborative

End-to-End Solutions
- Serve a diverse set of scientific domains
- Out-of-the-box gateway solution that can be customized
- Based on Docker – executable images that are the skeleton for a secure and functioning gateway
- Portable and reproducible
- Community-contributed

Software Marketplace for Science Gateways
“Use-what-you-need”
- API integration
- Variety of services
  - Information
  - Security
  - Execution
  - Data
  - Event
  - Accounting
  - Hosting opportunity

Gateway Discovery
- Open registry
- Promotes use of existing science gateways
- Community-contributed
- Admin approval
- Automated cleanup

Engage Other Areas of Institute
- Support projects leverage Collaborative components
- Framework evolves as a result of gateway engagements
- Community outreach

Software Integration & Community Contribution
- Docking mechanisms for community-contributed software, including NSF SI2
- Incorporate community standards
Scientific Software Collaborative

Diagram:

- **Gateway Users**
- **Gateway Developers**
- **Software Providers**

**Science Gateways**
- Gateway Registry
- **Hosted Environments**
- Platform APIs
  - Security
  - Event
  - Account
  - Etc.
- **Software Services**
  - SSI/ SSE/ Other Providers

**Developer Toolkits**
- Docker Containers
  - Institute & Community Contributed

**Education & Outreach**
- Training Materials
  - Videos, Workshops, Recipes
- Documentation
  - Tutorials, Quickstart Guides

**Platforms**
- Cloud Systems
- High-performance Computers
- Databases
- Storage
Community Engagement and Exchange

**Website Activities**
- Discussion forums
- Gateway showcase with case studies
- Symposium series
- News: media coverage, related happenings, academic publications, job openings, events calendar
- Curated blog with guest authors,
- Professional development: synchronous and asynchronous training
- Capture client/user feedback on web and through other areas

**Annual Conference**
- Tutorials and workshops
- Paper presentations
- Invited keynotes and panels
- Interactive elements: Open Space, poster session
- Travel support for students and campus IT staff

*Builds on 10 years of experience with GCE and IWSG series*

**Campus Gateway Groups**
- Task force builds campus-based expertise
- Channel for scaling institute services

**Outreach to Complementary NSF Initiatives**
- NSF SI2 projects
- Large NSF projects
- Science and Technology Centers
- Engineering Research Centers
- MolSSI software institute collaboration
Conferences and Workshops

• 2006: GCE (Gateway Computing Environment) workshop series started – successful 10th anniversary in 2015

• 2009: IWSG (International Workshop on Science Gateways) workshop series started – successful 9th anniversary in 2017

• since 2013: GCE partners with IWSG on yearly special issue
  9th IWSG took place 19-21 June 2017 in Poznań, Poland – next year in Edinburgh, Scotland

• since 2015: GCE and IWSG partnering with Australian IWSG-A (International Workshop on Science Gateways – Australia) on yearly special issue
  3rd IWSG-A will take place 16-17 October in Brisbane, Australia

• 2016: GCE extended to Gateways conference with 120 participants at the first event
  Gateways 2017 will take place 23–25 October 2017 in Ann Arbor, Michigan, USA
Workforce Development

4 Focal Areas

Promoting Gateway-Related Career Paths
- Student-related conference programs
- Campus opportunities
- Job boards

Providing Financial Support
- Enabling students learning gateway skills
- Including internship experiences

Establishing Center for Training and Education at ECSU
- Vigorous schedule of on-site and virtual training
- Development of training and course curricula about science gateways technologies

Integrating Gateways into Course Content
- Providing broader access to high-end resources

Partners

Google Summer of Code (GSoC)
National Organization for the Professional Advancement of Black Chemists and Chemical Engineers (NOBCChE)

SGCI Institute Areas
Association of Computer/Information Sciences and Engineering Departments at Minority Institutions (ADMI)
Molecular Science Software Institute
Opportunities

• Get advice (sustainability, usability, cybersecurity, other) on your existing gateway
• Work with SGCI to build a gateway for you
• Find a working gateway or gateway development software in our catalog
• Learn how to set up a gateway group on your campus
• Keep up to date on gateway developments
  • webinar series, gateways in the news, google scholar feed, case studies, annual conference, blog posts
• Enroll a student in an internship program to learn gateway development
• Partner with SGCI on your own projects, technologies and events
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
  • HPC usage…

• A breadth of interesting roles
  • HPC programmers, 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
Your Campus

Is your campus seeing an increasing number of research projects that include web-based applications using HPC? Does each group have to hire developers independently?

This can be time consuming and inefficient, but there is a solution.
YOU ARE NOT ALONE!

Addressing Software Sustainability on Your Campus

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. Individuals find it challenging to recruit and sustain teams that offer the diversity of expertise necessary for developing gateways.

Even ants wish they had an extra pair of hands when developing science gateways!
Sustainability via On-Campus Teams

Potential salary cost distribution

- Typical Research Team
- Extended Research Team
- Shared On-Campus Team

- Q&A Expert
- Security Expert
- Usability Expert
- Graphic Designer
- Professional Software Developer
- Postdoc
- PhD Student
Sustainability via On-Campus Teams

Potential salary cost distribution

- Typical Research Team
- Extended Research Team
- Shared On-Campus Team

Increased possibility of research team to focus on research

- Q&A Expert
- Security Expert
- Usability Expert
- Graphic Designer
- Professional Software Developer
- Postdoc
- PhD Student
Sustainability via On-Campus Teams

Potential salary cost distribution

Increase of quality of application and efficiency of software development

- Q&A Expert
- Security Expert
- Usability Expert
- Graphic Designer
- Professional Software Developer
- Postdoc
- PhD Student
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

Some 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 Mission

CRC engages in computational science, fosters multidisciplinary research and provides advanced computational tools and services. The CRC works to facilitate discoveries across science, engineering, the arts and humanities, social sciences, business and other disciplines.
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
Lessons Learned by the Four Teams

• You need an evangelist to build/sustain teams
• You are going to have staff turnover
• Plug into your talent reservoir
• 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
• Use internal and external resources for creating/maintaining science gateways
Funding Mechanisms – External Resources

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

- Extended Developer Support
- Software & visibility for gateways
- Diverse expertise on demand
- Incubator
- Scientific Software Collaborative
- Workforce Development
- Community Engagement & Exchange
- Sharing experiences & knowledge as a community
- Longer-term, hands-on support
- Student opportunities & educator resources
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 facilitators/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 …
Identify key contacts and evangelists on your campus

- Deans?
- Department chairs?
- HPC colleagues?
- Web developers?
- Gateway group?

“We like to bring together people from radically different fields and wait for the friction to produce heat, light and magic. Sometimes it takes a while.”
Sharing resources

• Share human resources (librarians, data scientists, web developers, HPC specialists,…)?
• Share free resources such as office space?
• Which departments?
• Possibility between universities?
• Presentations at events (new faculty events, meetings within faculty, meetings between departments, …)?
Role Play

Devil’s advocate

Ask for free resources!
Role Play

It will be painful...

but beneficial to put yourself in this role
Scenario

NSF offers funding for three years for a “cyberpractitioner”: first year 80%, second year 50%, third year 20%. You would like to use the funding for a usability specialist.

Ask the dean (Henry)!
Your own roadmap
YOU ARE NOT ALONE!

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