The Science Gateways Community Institute at Two Years

Nancy Wilkins-Diehr
San Diego Supercomputer Center
University of California, San Diego
La Jolla, CA 92093-0505
wilkinsn@sdsc.edu

Michael Zentner
HUBzero Platform
Purdue University
West Lafayette, IN 47907
mzentner@purdue.edu

Marlon Pierce
Science Gateways Research Center,
Pervasive Technology Institute
Indiana University
Bloomington, IN 47408
marpierc@iu.edu

Maytal Dahan
Texas Advanced Computing Center
The University of Texas at Austin
Austin, Texas 78758
maytal@tacc.utexas.edu

Katherine Lawrence
School of Information
University of Michigan
Ann Arbor, MI 48109
kathla@umich.edu

Linda Hayden
Center of Excellence in Remote Sensing
Education and Research
Elizabeth City State University
Elizabeth City, NC 27909
LBHayden@ecsu.edu

Nayiri Mullinix
School of Information
University of Michigan
Ann Arbor, MI 48109

ABSTRACT*

The Science Gateways Community Institute was one of the first two software institutes funded by the National Science Foundation's Office of Advanced Cyberinfrastructure in August, 2016. The structure of and services offered by the institute were developed as a result of seven years of planning grants that funded focus groups, a 5000-person survey and the development of a strategic plan. Now two years in, we provide an overview of the institute's service offerings and their usage, reflect on the experiences of some early clients, review our approaches to metrics and evaluation, and describe some lessons learned. We also describe the lightweight, adaptive management approach employed by the institute.

SGCI is organized into five service areas: Incubator, Extended Developer Support, Scientific Software Collaborative, Community Engagement and Exchange, and Workforce Development. This paper will highlight early successes in all five areas, from client achievements to conference experiences to our impact on students. We highlight areas where the institute has evolved — based on community feedback — from what was originally envisioned. We describe our use of the Entrepreneurial Operating System as a lightweight management approach for a highly adaptive organization. Finally, we include early plans for the execution phase of the institute.

CCS CONCEPTS

• **Information Systems → World Wide Web**; • **Human-centered Computing → Interaction Design**

KEYWORDS

Science gateways, software institutes, software sustainability

1 INTRODUCTION

Science gateways are known by many names and have many definitions. Portals, virtual research environments, research platforms, and virtual laboratories are just some of the additional terms used. In 2014, a 5000-person survey of US principal investigators and leaders in the academic community showed that science gateways are used across all domains of science and for a variety of purposes. Leading uses include gateway interfaces to educational tools, computational tools, data analysis tools and data collections. Fifty-seven percent of survey respondents were involved in gateway development in some way. They found they

PEARC '18, July 22–26, 2018, Pittsburgh, PA, USA

ACM 978-1-4503-6446-1/18/07…$15.00
https://doi.org/10.1145/3219104.3219142

needed — yet often could not afford — a variety of expertise including usability experts, cybersecurity professionals, quality assurance testers, community evangelists, software engineers, and project managers. They also wished for advice on technology selection, licensing, and business planning. A full analysis of this survey is provided in the 2014 publication "Science gateways today and tomorrow: Positive perspectives of nearly 5000 members of the research community." [1].

But beyond the survey response, it is clear that gateways play a pivotal role as a connector of resources and services in an increasingly complex research environment. Instruments of all types generate data that needs analysis, sometimes by massive supercomputers. Communities unfamiliar with cyberinfrastructure now have digitized data available to them making new modes of research suddenly feasible. But well-designed gateways require careful thought and planning.

Despite widespread use and a role as connectors of services, gateways have sometimes followed an unproductive cycle in which they were developed for research projects then used by early adopters, only to falter as they reached greater traction with the community because the initial research grant had ended. Some didn't have access to or couldn't afford the varied types of expertise needed. Working in isolation, developers often duplicated solutions already created by others, reducing their ability to focus on higher level functionality.

The Science Gateways Community Institute (SGCI) was created in 2016 and designed in response to the studies that identified the challenges gateway developers faced and the solutions that would help them. "Science Gateways: The Long Road to the Birth of an Institute" [2] describes this path and the resulting institute.

2 EARLY EXPERIENCES AND KEY LEARNINGS

Now two years in, we describe our experiences operationalizing this institute. While the structure and offerings that make up SGCI were community-driven, the institute is unique in its assembly of these services into a collective whole.

The National Science Foundation's Office of Advanced Cyberinfrastructure funds its flagship software institutes in two (but really three) phases. There are separate awards for conceptualization and implementation phases. The 5-year implementation awards consist of design (18-month) and execution phases. During the conceptualization phase, institutes work to determine community interest and compile a well-developed strategic plan. In our award, the same team involved at the conceptualization phase was selected for implementation. This allowed participants to develop strong working relationships and a shared understanding of the institute's components. We were able to immediately set up the structure, processes and personnel necessary to begin offering services. How this took shape in each functional area is described below.

2.1 Incubator

Evidence from our large-scale survey indicated a need for assistance with aspects of running a gateway other than just the technical tasks. Several members of the Institute management team have run their own gateway creation and operation groups at various scales within their own institutions and realize the importance of these aspects such as sustainability planning, budgeting and revenue models, project management, technology selection, cybersecurity, and identifying and reaching the desired audience. The team members also know from experience what a luxury it is to be able to afford such resources on a full-time basis — a luxury simply not available to most single-focus projects. Yet, through the Institute, an opportunity existed to deliver this expertise levered across the many client projects anticipated in the Institute. Through the Institute leadership's networks, experts in each of these areas were known. It became the challenge to organize a delivery vehicle for this expertise. The survey identified the need for such experts but not how their expertise should be delivered.

Based on previous experiences with accelerator programs in an entrepreneurial setting, the team decided that this pattern had several positives that could equally serve the gateway community: imparting of knowledge that helps clients help themselves when in need, the potential for extended client consulting engagements when greater assistance is needed, and importantly the formation of a community with members aware of each other's identities and projects, and with a desire to help each other even without the intervention of the Institute. The Institute therefore established an intense, in-person Bootcamp as a gateway into potentially more extended Incubator consulting services.

The Bootcamps are held twice a year for 5-day periods — all day, each day — to help attendees form bonds that they may carry outside of the intense week. Care is taken to make each session interactive, to the point where 50% of the time during the week involves some sort of interaction and group report-out. The focus is both on providing knowledge and working practical exercises on how to apply that knowledge. The week progresses with ways to communicate the attendees' gateways, identifying the audience they desire to reach, the value they provide, how they execute the delivery of that value, and how they may seek to sustain it past the grant period. Outside of formal class hours, we provide an experience where members can socialize around lunches and dinners that are out of the ordinary conference fare. At the writing of this paper, two Bootcamps have been held. The response has significantly surpassed the expectations of the Institute team. Anecdotally, attendees freely share highly positive assessments of their experiences and learnings including the subjects discussed, the formation of the network of peers, the meals, and even the length of the program. In post event surveys, attendees quantify the growth of their confidence in the topics discussed during the week. While the feedback is nearly all positive, suggestions have included providing "office hours" during the week to work through specific issues with the instructors, expanding some of the sections with more examples, and looking at some additional case studies.

The Incubator has also engaged a total of 18 projects for consulting purposes. These engagements have included usability studies, financial modeling with sustainability analysis, technology selection, and a variety of short-term, question-and-answer activities. While it was anticipated that the consulting load would be intense engagements of one or two months in duration, the reality more closely resembles that of startup companies. Many consulting engagements need brief periods of intense help (hours to a few days) followed by extended periods of inactivity while clients work through the issues discussed during the short period. Fewer engagements involve continuous periods of extended assistance beyond a week. Notably, usability consultations follow this pattern. As a result, the Incubator consulting capacity is higher than anticipated in the number of clients that may be simultaneously engaged, with overall less intense continuous engagements.

2.2 Extended Developer Support

SGCI's Extended Developer Support (EDS) is modeled after the highly successful Extended Collaborative Support Services [3] effort of the NSF-funded XSEDE project [4]. EDS is designed to help new gateways come into existence and existing gateways add major new capabilities that are needed for growth and sustainability; Incubator Bootcamp provides the complementary services for helping gateways make these strategic choices.

EDS support is allocated for an initial six months, with an additional six-month extension for projects that are progressing well but need more effort to complete their proposed work. A gateway client is assigned an SGCI consultant who works with the client gateway at 25% effort. The first stage of the consultation is the development of a work plan, which identifies key deliverables, monthly milestones, and external constraints on work (such as external deadlines). The intention of the work plan is to provide a coherent and feasible outline of the work to be accomplished; the contents of the work plan may be adjusted as the work proceeds.

We have adopted a light weight application process. Development of a coherent work plan to initiate the project is used as a gating exercise; projects that are not yet in the stage to formulate a work plan are directed to other SGCI consulting services, including the Incubator to further crystallize their gateway ideas. We assign consultants for an initial six months with an informal review before continuing the second six months. We typically cap total support for twelve months. Clients who have used their twelve months of support can request additional support, but new clients are given priority.

While the XSEDE science gateways program focuses primarily on integrating existing gateways with XSEDE resources, SGCI's EDS support is more flexible. Development of frontend user environments, integration of non-XSEDE resources such as campus clusters and commercial clouds, and development of non-computational services such as services to support scientific collaboration, citizen science projects and gateways that manage cyberinfrastructure are all suitable for EDS support.

Funding levels for EDS enable us to support twelve projects at a time. At the time of writing, we have supported twenty clients; see https://sciencegateways.org/consulting/clients for a full listing. Client gateways have been funded by the Department of Energy, the National Institute of Health, and the Department of Homeland Security in addition to the NSF. Support has included the creation of a gateway to serve new campus computing clusters (USD Campus Gateway), a gateway for the client scientist to manage contributed data (ENIGMA), gateways that can manage virtual infrastructure (CloudLaunch, VC3), a gateway to enable a Research Coordination Network for the ecological sciences (ESA), and a gateway to support chemistry education (Chem Compute). This brief sampling is intended to show the breadth of gateway collaborations we have supported, in addition to gateways that simplify High-Performance Computing (HPC) and High-Throughput Computing (HTC) job submission. We do that, too, as exemplified by the nSides, SimCCS, IPT, COSMIC^2, and LSU Systems Biology consultations.

The institute's experience is still rather limited, but we have seen gateways be most beneficial in communities where data generation is exploding (cryo-EM), the need for powerful computation is great, and the experience of the user community with large-scale computation is limited. Gateways can also help in fields with access to newly digitized data, for example text analytics of newspaper archives as exemplified in a criminal justice project. Gateways have been less successful when data formats dictated by a federal agency changed without the knowledge of the developer. The key here is to be closely tied in to changing data standards if a gateway relies on external data.

We have worked with teams creating general interfaces to campus computing systems where they plan to create a gateway to address the needs of those running the most popular codes, thereby reducing the majority of their support load. These interfaces can be rather simple in nature, but still provide a lot of value to the end users and to local support staff. Because gateways using XSEDE come with the additional requirement of an allocation and the requisite monitoring of renewal of that allocation, we see clients move in this direction only if the computing needs require it. But if they do, XSEDE provides an outstanding source of world- class computing and full support of the science gateway usage model.

A key challenge for the next phase of the institute is to scale out EDS support. Our goal is to gain maximum impact for each consultation beyond the direct support for clients. Most successful consultations are described in a corresponding blog entry (developed with Community Engagement and Exchange staff) at sciencegateways.org, with an emphasis on the technical challenges that were overcome. Our intention here is to provide gateways with overviews of technical solutions to common problems (such as authentication) so that they can evaluate and adopt solutions on their own without needing full EDS support. EDS also works closely with the Scientific Software Collaborative to get client gateways listed in the software catalog and to help gateway frameworks improve their engineering processes so that they can be made easily available through gateway hosting services. Workforce Development interns and outreach programs will also help us increase gateway-building expertise at a broader number of sites. Interns are paired with EDS staff each summer to work on client projects. EDS staff will also be working with the

Incubator team to develop a tech-focused Bootcamp spinoff that will help new gateway providers with technology choices.

2.3 Scientific Software Collaborative

Years ago, a focus group attendee commented, "wouldn't it be great to have an app store for gateways". SGCI's Scientific Software Collaborative (SSC) is our response to that, but delivers much more. SSC's offerings help those needing assistance in their gateway building efforts, including choosing technologies and integrating new features and capabilities (e.g., visualization or computational tools, education support resources, etc.). The initial vision for SSC is based on two guiding principles: Gateways should not be built as a series of one-off efforts but, at the same time, a single software solution will not fit all problems. The goal of the SSC is to promote science gateways, simplify development, and expand the capabilities of science gateways. The SSC provides solutions that facilitate the following:

- Discovery of gateways for those looking to use production gateways
- Discovery of gateway-related software and assistance for those looking to build gateways
- Hosting environments for gateway developers who need to develop, test, and create alpha releases of their gateways before migrating to their own platforms

During early operations, the SSC team provided infrastructure support for the institute as a whole due to the team's experience in this field of work. SSC set up the entire JIRA infrastructure for ticket and issue support for helpdesk services, workflow configuration for EDS, and SSC project management. SSC set up Trello and Slack, used by the entire team for management and communication (described in Section 2.6). Work continues as our functionality and needs expand. For example, Google forms must communicate with Jira and our Customer Relationship Management (CRM) system in Hubspot. Connecting these infrastructure components is not always straightforward. The SSC team has also worked with CEE to configure two Virtual Machines with Liferay 6.2 Community Edition for the new website.

Moving past infrastructure for the project, the SSC team developed and released the new SGCI Science Gateway Catalog and the SGCI Hosting Environment. Users can browse the catalog, search for specific metadata, and filter by categories or tags. Users also have the ability to create an account, login, and contribute their own science gateway and software to the listing. The addition then goes into "pending" mode and must be approved by an administrator before being publicly viewable. On the hosting front, the SSC team has worked on configuring OpenStack using TACC's Rodeo system as a sandbox hosting environment. This is now open for applications and already has a few early users. Both of these efforts will be actively promoted to the community and updated as appropriately based on user feedback.

The SSC will continue to evolve over time and engage the community early on to measure the impact and success. The components of the Scientific Software Collaborative will be leveraged by the Incubator and Extended Developer Support to help researchers build and improve their science gateways. We work closely with Community Engagement and Exchange to offer comprehensive documentation and outreach services specific to the framework and with Workforce Development to engage students in learning about the software components and improving their programming skills. We will also continue to expand our services to create a community around SSC offerings by launching the Affiliates program that will allow software collaborators to offer their services to SGCI customers.

2.4 Community Engagement and Exchange

Community Engagement and Exchange (CEE) supports the heart of the gateways community, as a way for people to connect with, support, and learn from each other. The area achieves this by organizing events and resources targeted toward the unique needs and interests of the gateway community. CEE also supports the institute as a whole, engaging new participants and clients through promotion, questionnaires, written materials, and social media.

The two primary events organized by CEE include an annual conference and a Webinar series. The Gateways conference is an expansion of the previously successful Gateway Computing Environments (GCE) workshop series (since 2005) that had been hosted primarily at larger computing conferences, introducing the scheduling limitations associated with being co-located. The Webinar is a monthly, one-hour seminar that rotates through a selection of topics including tools, integrations, best practices, and any other useful resources that could help developers improve their gateways. CEE also, in collaboration with the Incubator, provides expertise about creating campus-based gateway development groups.

The inaugural Gateways 2016 conference was scheduled for early November 2016, anticipating an earlier start date for SGCI's funding. With the official "opening" on August 1, the only conference activity had been the Call for Participation and reservation of the venue. In three months, we launched the entire expanded, 2-day conference while also hiring the CEE Coordinator (responsible for conference logistics) who joined at the start of October. We had hoped to increase attendance from the usual 30–40 participants to 50 or 60. Instead, we had tremendous interest with 118 attendees and 16 remotely joining the live stream of presentations. Responding to attendee feedback, the Gateways 2017 conference added a full day for tutorials and encouraged SGCI clients to share their work and experiences.

The Webinar series has also grown and developed in response to observed interest and participant feedback. The initial model was to identify a rotating set of topics, but we found that fewer people attended our Gateway Showcase (featuring two gateways in the hour). Discussion and observation prompted us to reshape the focus of the topics to deliver more usable information that attendees could apply to their own projects. Additionally, Webinars had required registration for us to better know who would be attending and what questions they might have for the presenters, but we have discontinued registration and provide the

same Webinar "meeting room" for every event. Both changes have improved attendance.

CEE's other substantial efforts focus on the development and maintenance of a website, including significant content creation, to support the sharing of resources, news, information, and SGCI's services and events. With the formation of the institute, our first goal was to redesign the website to better support anticipated needs and expanded content. We began by providing information about SGCI's consulting services (provided by EDS and the Incubator), resources (publications, learning materials, etc.), events (Incubator's Bootcamp, Workforce Development's programs for students and educators, and CEE's programs), and community information (news, calendar, newsletter, job listings). In addition, our blog has expanded to foster our community connections, including a monthly "People of SGCI" to introduce our varied and distributed staff. As time has passed and our services have matured, we have added "success stories" featuring "graduated" clients, guest blog posts, and "tech blogs" explaining how to work with gateway technologies (in collaboration with EDS).

CEE also persistently promotes the opportunities afforded by the Institute and its broader community. Our newsletter mailing list of approximately 1400 people was seeded by participants in the Institute's formative focus groups and survey, but we have actively extended our reach by identifying and connecting with pertinent communities through Twitter, LinkedIn, and online forums. As a more direct approach, SGCI staff have promoted SGCI's services and resources through presentations and invited talks at conferences, workshops, webinars, and meetings of professional groups or domain-centered associations; the staff has averaged more than 4 presentations per month since inception.

CEE's program to promote and support campus-based gateway development groups has the goal of helping campuses build a more cost-effective, sustainable, and efficient way of supporting the growing number of gateway projects. Such groups also support career paths for research-focused developers. We have reached out to specific university campuses as well as to professional groups through this effort. These groups, such as XSEDE Campus Champions and ACI-REF, have individuals who are particularly well positioned to identify the need for and facilitate the creation of campus-based development groups. This effort has included face-to-face and web-based presentations as well as visits to campuses, meetings, and conferences. The topic is also part of the Incubator's Science Gateways Bootcamp.

Going forward, the CEE team will continue to look for opportunities to shape our services to meet the needs of our community as well as look for new opportunities to expand our community's boundaries.

2.5 Workforce Development

Transforming the frontiers of science requires users and developers who are trained and motivated to tackle the difficult challenges of working in uncharted territory. SGCI works to build the nation's capacity to generate the workforce needed to meet these challenges. The primary SGCI approach to addressing this performance goal is through engagement and the integration of research and education. First, we identify talented young people, through national searches and through our partnerships with the National Organization for the Professional Advancement of Black Chemists and Chemical Engineers (NOBCChE) and the Association of Computer and Information Science/Engineering Departments at Minority Institutions (ADMI), then connect them to the science gateway community through directed experiences. Recruitment efforts now also take place during PEARC, Grace Hopper, SXSW and other conferences.

The move into the execution phase will be based on lessons learned and guided by a skills development report, which has identified non-technical, basic-technical and advanced technical topic areas. Non-technical skills include communication skills, ethics, and internship practices. Basic technical skills include basic software engineering and system administration while advanced skills include distributed systems, cluster/cloud computing, messaging systems, cybersecurity and others. Internships will continue to build relationships with the community while providing students options to apply and enhance their skills.

The service delivery model for non-technical and basic technical skills will include face-to-face training during the academic year, a summertime 4-week coding institute, and conference workshops, and hackathons will continue with the content closely tied to non-technical and basic technical skills topics. Elizabeth City State University (ECSU) will deliver both non-technical skills and basic-technical skills. The service delivery model for advanced technical skills will be led by Indiana University (IU). IU currently offers a CSCI-B649, Topics in Systems course on Science Gateway Architectures. IU also conducts advanced technical workshops and tutorials during PEARC, SGCI's Gateways conferences, and other appropriate conferences.

Using these models of service delivery, Workforce Development will educate the next generation of users and developers who will create the software that cyberinfrastructure requires to advance gateways and will engage underrepresented minorities who are less likely to self-identify as contributors to or users of Gateway services.

2.6 Management

As described in sections 2.1 through 2.5, SGCI consists of many different activities. Coordinating such diverse offerings could mean a higher management load. But as a project, we wanted the budget to be spent as much as possible on community services. In addition, as a community-designed organization we had to be ready to adapt to the changing needs of the community. So we needed a system that was lightweight and flexible. We settled on the Entrepreneurial Operating System™ (EOS) approach to planning and execution and a suite of low-cost or free tools to efficiently collaborate and communicate.

2.6.1 Entrepreneurial Operating System

The EOS approach to management is based on the book Traction by Gino Wickman [5]. EOS breaks down the

components of any business into 6 pieces: **vision, people, data, issues, process and traction**. While a full description of this process is the subject of the book and beyond the scope of this article, we summarize the key concepts here as they relate to SGCI.

EOS starts with a vision. The SGCI team had two intensive in-person planning sessions with a professional EOS Implementer at the project's outset. Our **vision** includes a core focus, core values and a brief marketing strategy, as well as a 1-year plan, 3-year picture and 10-year target. The **people** component consists of assembling staff who understand the mission, want to contribute, and are capable of contributing. In SGCI, where anyone on the team could be representing the institute at any point in time, this was very important. **Data** involves setting up a Scorecard and selecting data to track weekly and quarterly. **Issues** are addressed both on weekly calls and at more in depth Quarterly Pulse meetings. This creates a regular forum for discussing items, brought by anyone in the leadership, that require decisions. We identify the business problem that needs to be solved, discuss the issue, and reach an effective solution. Any organization needs well-defined, repeatable systems, so **processes** are simply documentation of SGCI's operational actions, understood by all.

Finally, and perhaps most important is **traction**. Traction is realized through our regular structured meetings (weekly and quarterly) when we select quarterly "Rocks" that break a long-range vision into attainable steps. Rocks are contributed individually, then prioritized and selected by the group. This unique process achieves two goals. Individual generation means all voices are heard. In the group prioritization and selection, the leadership team acts as a board of directors for the project, thinking about the project as a whole and not just the areas we lead.

2.6.2 Organizational Tools

All projects need a technology infrastructure in order to function. Many components of our infrastructure are quite common, but we feel the listing of them and the uses for each may help others in similar projects ramp up more quickly. Cost was also an issue as we strive to keep management costs low.

Google docs. These are used for many things: event planning, service request forms, work plans, slide templates, logos, flyers, staffing lists and presentations made, paper and abstract development, templates for communications, website content and success stories, internship planning, and so on. This project produced detailed monthly reports for the first 16 months of the project, so we needed a very streamlined process. Applying the same reporting structure from one month to the next and updating text makes report completion less time-consuming. The ability to work in parallel and communicate directly with others via comments in documents is also crucial. We also use Google docs for EOS materials, such as our Scorecard.

Trello. We use Trello to run the weekly L10 meetings. We find it ideal for tracking progress as well as assigning new "to-do's" as a result of our weekly discussion. We have a Trello board for the Leadership Team as well as boards for each area within SGCI. Cards can be transferred easily across boards as needed. Trello is also used for our Project Schedule. This makes it easy to review during Quarterly Pulse meetings. **Gantiffy** displays that Trello board in a Gantt chart view.

Slack. Like many organizations, we found Slack a useful communication tool for immediate topics and a great help in reducing email.

Jira. We use Jira primarily for tasks within our Extended Developer Support projects, though we initially used it for our helpdesk system as well.

Hubspot. Early on we used spreadsheets and Trello boards to log customer interactions, but this quickly became unmanageable. Very soon, we knew we needed a customer relationship management (CRM) system to better keep track of our interactions with clients. We chose Hubspot based on its ability to automate record-keeping through email, extensive documentation, and low cost. This has been ideal to log multiple interactions from staff members across the project and also review new clients at weekly meetings.

3 EARLY SUCCESSES

3.1 QUBES

The Quantitative Undergraduate Biology Education and Synthesis (QUBES) science gateway's mission is to foster a community of math and biology educators who share resources and methods for preparing students to use quantitative approaches to tackle real, complex, biological problems. It operates as a virtual hub that consists of online learning communities that connect teachers with quantitative biology projects, resources, and expertise while also supporting implementation in the classroom.

Since its founding in 2014, the QUBES platform has evolved and grown, and so has its need for support. When the SGCI launched in 2016, they were among the first to submit a request for services and have, since their initial request, engaged with each service area offered by the SGCI.

- Extended Developer Support: custom development support
- Incubator: Science Gateways Bootcamp and usability consulting
- Workforce Development: hosted an intern through the summer internship program
- Community Engagement and Exchange: attended and presented at Gateways 2017
- Scientific Software Collaborative: added QUBES to the Gateway Catalog

Having access to these resources had a significant impact on QUBES and has helped the team achieve their goals. In reflecting on the experience of engaging with SGCI, the PI for the project, Sam Donovan, said, "As education experts, we have some strong opinions about the types of faculty activities we want to support—but realizing those ideas in a gateway context was beyond our capacity. Working directly with programmers and development project managers through the EDS program was personalized, professional and extremely productive." Co-PI M. Drew LaMar added, "SGCI's role in providing UX support has been a huge

benefit to us. It is important to mention, though, that it is more than just the initial UX study—SGCI is a community of like-minded colleagues interested in the same things we are, which is maximizing the impact and minimizing the pains of utilizing cyberinfrastructure in faculty's research and teaching."

The QUBES team is not quite done evolving, and they anticipate requesting further services from SGCI to take their platform to the next level.

3.2 COSMIC²

The COSMIC² science gateway offers an easy, web-based platform for structural biologists to determine macromolecular structures using cryo-electron microscopy (cryo-EM). Over the last five years, cryo-EM has undergone advancements that allow for the determination of atomic structures which, in turn, has led to a revolution in the field of structural biology. So much so, that the Nobel Prize for Chemistry 2017 was awarded to three scientists for cryo-EM. Researchers around the world are now collecting and analyzing unprecedented amounts of data and, per experiment, an individual scientist could easily process up to 10-30 terabytes of data from a single dataset collected over the course of a week. Datasets of this size require high-performance computing (HPC) resources and, as a result, new users to cryo-EM face a number of obstacles when it comes to handling large datasets and dealing with managing and submitting jobs to HPC resources. The COSMIC² science gateway was developed to lower these barriers.

The PI of the COSMIC² science gateway, Michael Cianfrocco, requested services from SGCI's Extended Developer Support in order to implement technologies that would help make things run more smoothly for their users. Cianfrocco worked with two SGCI consultants to integrate two Globus technologies. The first was Globus Auth, a federated identity management service that manages user identities and access to resources by allowing users to utilize their existing organization login to access the gateway. The second was Globus Transfer, which offers the ability to transfer large datasets in a secure and reliable manner, giving COSMIC² users the ability to transfer their terabyte-sized datasets to the gateway. Once the data is transferred, the user can set up an analysis task and submit it as a job to the XSEDE HPC resource Comet at the San Diego Supercomputer Center. When the analysis job has been completed, the user can view and download their results from the gateway.

About the work that SGCI completed on COSMIC², Cianfrocco had this to say: "We knew that the need was there for a cryo-EM gateway, and having the experts at SGCI available made the gateway development process much more efficient since they have the starting technologies to work with and they know how to solve the unique problems facing gateway developers."

3.3 Conference and journal

As described in section 2.4, despite a short planning timeline, attendance at the inaugural Gateways 2016 tripled when compared to previous gateway workshops. Broad support from the program committee included international reviewers for submissions and some international attendees and presenters. Gateways 2017 continued this success while attracting over 60% new attendees. For Gateways 2018, we have increased community participation, soliciting volunteers to help shape the conference program and oversee the review process.

Following each conference, we have collaborated with the International Workshops on Science Gateways (IWSG, both in Europe and Australia) to publish a joint special journal issue from extended versions of papers presented at the conference or workshops. All presenters were invited to submit, and we received plenty of submissions. This journal provides a unique venue for publishing practice and experience papers on science gateways.

3.4 Internships and Workforce Development

Early results from the Workforce Development internship program show the benefits of building confidence and skills in the interns. Interns also deepen their engagement with the community through the assignment of a mentor for the duration of the internship. For the first summer, internship placements were made to SGCI sites (TACC, Purdue, IU) as well as a placement at William and Mary University part-funded by an SGCI client.

As we move forward with these successful internships, workforce development will implement procedures that build mentoring skills and early interactions between students and mentors. The practices will encourage interns and mentors to interview or to jointly work on challenge problems prior to final placement. For summer 2018, we are developing a mentoring resource center and internship webinars will be conducted for the summer 2018 programs. Intern/mentor relationships have developed from several sources, including conference pairings and client interactions.

Building non-technical and basic technical skills in future interns will be facilitated by a NIH grant to locate a CIPRES developer at ECSU working on one of the most successful science gateways. In addition to writing CIPRES software, that developer will also provide instruction for onsite and virtual skills development events. This level of co-operation with clients speaks well for the future of the internship program as well as the growing number of clients who submit request to host interns at their site.

4 FUTURE WORK

After getting functional processes set up and attracting a steady stream of interested clients during the design phase, we envision several next steps. First, methodically reaching out to researchers, educators, and scholars who can benefit from our services, particularly from underrepresented demographic groups and underrepresented disciplines. While Workforce Development has been very successful in these areas, we would now like to expand that success to the other service offerings.

The technology landscape continues to evolve. Jupyter Notebooks have offered a radical departure in gateway development techniques. There will be other such departures. Our staff members need to continue to sharpen their skills in order to be most effective. The definition of gateway will also continue to

widen. Future directions might include ties to publications, data, and the ability to reproduce results through a gateway. We already see gateways being incorporated into online textbooks so that students can immediately perform calculations they are studying.

Finally, and perhaps most importantly, we need to work with the community to develop expertise independent of SGCI offerings. For example, developing a program to instantiate "mini-institutes" on campuses, organizing all the expertise that SGCI offers, but at a campus level, will allow us to replicate our success in a more scalable way.

SGCI will need to continue to react to community feedback and external developments as we go forward. One way we can accomplish this is by establishing a science advisory panel (in addition to our existing Steering Committee) to engage representative members of our community in shaping the opportunities and impacts that we pursue. On a strategic level, we will continue to engage new partners who can extend the services and expertise that we offer and the communities that we serve. We also plan to continue our methodical approach to goal setting, moving forward toward our long-term goal to support science gateway developers as they radically change the way science is conducted.

ACKNOWLEDGMENTS

This work was supported National Science Foundation grant ACI-1547611.

REFERENCES

[1] Lawrence, K. A., Zentner, M., Wilkins-Diehr, N., Wernert, J. A., Pierce, M., Marru, S., & Michael, S. (2015). Science gateways today and tomorrow: positive perspectives of nearly 5000 members of the research community. Concurrency and Computation: Practice and Experience, 27(16), 4252-4268.

[2] S. Gesing, N. Wilkins-Diehr, M. Dahan, K. Lawrence, M. Zentner, M. Pierce L. Hayden and S. Marru. 2017. Science Gateways: The Long Road to the Birth of an Institute. Proceedings of the 50th Hawaii International Conference on System Sciences. http://hdl.handle.net/10125/41919.

[3] Wilkins-Diehr, N., Sanielevici, S., Alameda, J., Cazes, J., Crosby, L., Pierce, M. and Roskies, R., 2015, March. An Overview of the XSEDE Extended Collaborative Support Program. In International Conference on Supercomputing (pp. 3-13). Springer, Cham.

[4] Towns, J., Cockerill, T., Dahan, M., Foster, I., Gaither, K., Grimshaw, A., Hazlewood, V., Lathrop, S., Lifka, D., Peterson, G.D. and Roskies, R., 2014. XSEDE: accelerating scientific discovery. Computing in Science & Engineering, 16(5), pp.62-74.

[5] Gino Wickman. 2012. Traction. BenBella Books. ISBN-13: 978-1936661831.