

The Science Gateways Community Institute: Going Beyond Borders

Sandra Gesing
University of Notre Dame
Notre Dame, USA
sandra.gesing@nd.edu

Maytal Dahan
The University of Texas at
Austin, USA
maytal@tacc.utexas.edu

Michael Zentner
Purdue University
West Lafayette, USA
mzentner@purdue.edu

Nancy Wilkins-Diehr
University of California
San Diego, USA
wilkinsn@sdsc.edu

Katherine Lawrence
University of Michigan
Ann Arbor, USA
kathla@umich.edu

Abstract—The Science Gateways Community Institute (SGCI), opened in August 2016, provides free resources, services, experts, and ideas for creating and sustaining science gateways. It offers five areas of services to the science gateway developer and user communities: the Incubator, Extended Developer Support, the Scientific Software Collaborative, Community Engagement and Exchange, and Workflow Development. While all these services are available to US-based communities, the Incubator, the Scientific Software Collaborative and the Community Engagement and Exchange serve also the international communities. We aim at reaching out and supporting beyond borders on international scale with diverse measures and our intent is to form and deepen collaborations with partner organizations and coalitions beneficial and/or related to the science gateways community. Research topics are independent of national borders and researchers spread worldwide can benefit from each other's research results, software, data and from lessons learned — via online materials and publications or at international events. The gateway community has long benefitted from this type of exchange. This paper will present related work describing the benefits of international collaborations generally, and specifically as they relate to science gateways. We go into detail regarding SGCI's ongoing work on an international scale and its work planned in the near future.

Keywords—*science gateways community institute; international collaboration; community engagement and exchange; scientific software collaborative; incubator*

I. INTRODUCTION: THE ORIGINS OF THE SCIENCE GATEWAYS COMMUNITY INSTITUTE

The Science Gateways Community Institute (SGCI) was established as one of the US National Science Foundation's first Software Institutes in August 2016 [1]. This award was the result of over 10 years' work studying the community and its scientific contributions and the benefits that could be gained through a central organization providing increased visibility to these efforts [2]. Seven months in, early interest has been strong. Many are recognizing that work they have done for many years can be classified as a science gateway. Many more are recognizing what an increased contribution their work might make if it were exposed as a science gateway.

Early work in the National Science Foundation's TeraGrid program [3] involved the creation of policies and a program to provide access to NSF funded supercomputers for "community

leaders who are building discipline-specific cyberinfrastructure capabilities". These capabilities were by and large web-based, and the science gateway concept was born. These gateways came from many disciplines and had many different funding models. Over the eight years of the program we saw a variety of gateways large and small come and sometimes go, sometimes despite strong user interest [4]. Frequently failing gateways could undermine confidence in and potential impact of these powerful interfaces. Two subsequent studies and a 5000-respondent survey [5] highlight characteristics, environments and support systems that contributed to successful science gateways. These studies have directly informed the design of the Science Gateways Community Institute.

SGCI's current conference series builds on significant work by others. The Open Grid Computing Environments program, that advanced the use of web portals, began its workshop series in 2005 [6]. In 2009, the International Workshop on Science Gateways (IWSG) series began and in 2015 an Australian series, IWSG-A [7]. These three series have a history of cross attendance as well as keynote speakers from the other series.

II. INTERNATIONAL COLLABORATION IN RESEARCH

The benefits of international collaborations are well supported. The Technopolis group and the Manchester Institute of Innovation Research analyzed in 2009 the drivers of international collaboration in research [8]. Their examination is based on international collaborations in the EU context but also beyond it including the US, Australia, Brazil, Canada, China, India, Japan, Mexico, Russia, and South Africa. They found that the main four drivers of international collaborations are

- Improving national competitiveness
- Supporting less developed countries by developing science, technology and innovation capabilities
- Tackling global societal challenges
- Creating good and stable diplomatic relationships (and indirectly ensuring international security)

While globalization is not a new trend, it becomes more and more evident in industrial research and the worldwide mobility of researchers. The success of LIGO [9] and the detection of gravitational waves is one of the recent

achievements of international collaboration that was attained by more than 1,000 scientists from 15 different countries. The 2009 report elucidates that the increased political debate and urgency of global challenges such as climate change, health issues and sustainable energy resources has led to more international collaborations. Further findings include that smaller countries tend to collaborate internationally more often than bigger ones, and the authors refer to literature reviews that show that publications with international partners are more cited than publications with one author or a group of authors from one country. This trend has many reasons and does not refer to the quality of publications by single authors but is evident in citation counts.

For SGCI, international research has—besides the big picture of such collaborations—also very practical aspects. Software solutions already developed in countries outside of the US can be also applied in the US if the source code is openly accessible and if access technologies and policies for its usage are aligned. Software frameworks and technologies are in general not bound to a location; they can be bound to hardware requirements or require high-speed networks for efficient and effective use.

Despite all the possibilities and opportunities of international collaborations, there are also challenges. The established system sciences conference series HICSS (Hawaii International Conference on System Sciences) [10], for example, for more than 10 years has had a minitrack analyzing and discussing the challenges for global virtual teams—not only in academia but also in industry and governmental projects [11]. They investigate challenges including temporal, cultural, and language differences; diversity in management and work styles; and conflict management. Research recommendations suggest alleviating challenges by increasing the awareness of these challenges and by training people. While SGCI offers training and webinars in the context of science gateways with consideration of these aspects, it will not offer training on the topic of international collaboration itself.

The last decade has led to diverse, mature, science-gateway frameworks such as HUBzero [12], the Agave Platform [13], Apache Airavata [14], and Galaxy [15]. Their worldwide usage is an example for the potential of sharing science gateways



Figure 1: Global users of HUBzero-based science gateways

beyond borders (Fig. 1 illustrates the usage of HUBzero-based science gateways around the globe) and are success stories for international usage of science gateway technologies.

The remainder of this paper is organized as follows. Section 3 covers related international organizations and programs. SGCI is described in section 4 with a particular focus on the international aspects of each service area. Section 5 goes into detail about ongoing work between SGCI and its international counterparts and future work.

III. RELATED WORK

The UK Software Sustainability Institute (UK SSI) [16] was funded in 2010 and serves the UK's research software community as well as partners on an international level. Its goal is to cultivate better, more sustainable, research software to enable world-class research. Its investigations have elucidated that 70% of UK researchers rely on software in their research [17]. The institute has worked with over 50 projects to directly improve their codes, written over 80 best practices and organized training events to teach over 1,000 learners the basics of software engineering. They have built a platform to provide researchers and developers access to the contacts, information and training necessary to develop reliable and reproducible research software. Besides UK-based and online activities, the UK SSI organizes a yearly Research Software Engineers Conference open for international participation and is majorly involved in the organization of further events such as the workshop series Working towards Sustainable Software for Science: Practice and Experience (WSSSPE) [18]. Similar to the goals for international collaborations of SGCI, they have established partnerships in Canada, the US, and European countries. They achieved to be the de facto point of contact for research software good practice. Science gateways form a subgroup of research software and SGCI aims at building a long-term partnership with the UK SSI. The institutes can complement each other with services in the international research software and science gateways landscape.

Science gateways are often called Virtual Research Environments (VREs) in Europe and the Horizon 2020 program regularly announces solicitations calling for projects in the context of VREs in general or dedicated to a certain research domain. For example, VRE4EIC (A Europe-wide interoperable Virtual Research Environment to Empower multidisciplinary research communities and accelerate Innovation and Collaboration) [19] targets key data and software challenges in supporting multidisciplinary data-driven sciences. Its specific objectives include among others:

- Increase the VRE usability for multi-disciplinary research;
- Increase the deployment of VRE on different research infrastructures by abstracting and reusing building blocks and workflows;
- Improve the contextual awareness and interoperability of metadata.

These objectives are well aligned with objectives of SGCI that are concerned with supporting the development of user-friendly science gateways and reusing existing mature science gateway frameworks. A goal is to partner with the project on suitable best practices.

ENVRIplus [20] is a Horizon 2020 project “bringing together Environmental and Earth System Research Infrastructures, projects and networks together with technical specialist partners to create a more coherent, interdisciplinary and interoperable cluster of Environmental Research Infrastructures across Europe”. The project published a report with a vision for a worldwide environmental research infrastructure. The recommendations include topics independent of the research domain such as improving interdisciplinary collaboration, sharing computationally expensive results, data-intensive federation support and software sustainability, which overlap with goals of science gateways. The project organizes and teams up on events for best practices to create environmental research infrastructures and is also a very suitable potential partner for SGCI to share experiences and best practices on software solutions for interdisciplinary research.

Science gateways are an increasingly common component of funded activities by many agencies in the US. The National Science Foundation (NSF) and National Institutes of Health (NIH) directly mention science gateways in solicitations and roadmaps, [21]; moreover, NIH and Intelligence Advanced Research Projects Activity (IARPA) also allow for applications requesting funding for international partners. SGCI can offer all services of the institute to such international collaborations when there is support from a US-based partner.

IV. SGCI’S CONCEPT FOR INTERNATIONAL COLLABORATIONS

SGCI consists of five service areas, which closely interact and complement each other to serve the communities tailored to their needs and increase the sustainability and reusability of science gateways.

- The Community Engagement and Exchange (CEE) brings the community together through interaction and professional development.
- The Scientific Software Collaborative (SSC) leverages and promotes existing investments in gateway technologies.
- The Incubator team provides expertise and guidance for the science gateway lifecycle.
- The Extended Developer Support (EDS) enhances gateways through technical help from experienced gateway developers.
- The Workforce Development nurtures the next generation of science gateways users and developers.

CEE, SSC and the Incubator present near term targets for engaging and supporting international collaborations on science gateways via SGCI. These are the areas that can be offered to the international community, while the EDS and Workforce Development are actively supporting various US-

based projects and students and are available for US-based communities but not on an international level. The restriction on national scale results inherently from characteristics of the funding. Thus, we describe here in detail the three areas that are available to the international community.

A. *Community Engagement and Exchange*

CEE provides a wide variety of measures to interact with the SGCI and within the community surrounding science gateways. The yearly Gateways conference series succeeds the successful Gateways Computing Environment (GCE) workshop series and attracted more than 120 participants from the US, Poland, Italy, Mexico and Australia in 2016. It is a major event for the community to share experiences, present their work, identify new issues and shape future directions of research in the context of science gateways. It takes place in the US and partners with the European workshop series International Workshop on Science Gateways (IWSG) and International Workshop on Science Gateways–Australia (IWSG-A) on a yearly journal special issue as well as on invitations of keynote speakers.

Besides these science-gateway focused events, SGCI staff also present and publish at international technology-focused or domain-focused events, which attract communities on different scale, e.g., at the American Geophysical Union (AGU) Fall meeting with approximately 24,000 participants in 2016, at the Hawaiian International Conference on System Sciences (HICSS) with about 1000 participants yearly, or the 2nd Personalized Medicine Conference with about 40 participants.

In addition to face-to-face meetings, SGCI features diverse online options for interaction with different levels of involvement. Regular newsletters inform not only about news regarding the activities and products of the SGCI but about related events, publications and achievements of science gateway projects. The monthly webinar provides the opportunity to present use cases, projects, technologies and/or concepts. Researchers all over the world are also invited to write blog posts, offer their project as a case study, and advertise related events. SGCI continuously works on best practices for various topics related to science gateways, especially in the Incubator and Extended Developer Support service areas. We aim for international contributions that consider the distinct science gateway landscapes in different countries and on different continents with regard to the variation of technologies, policies and funding opportunities. Additionally, researchers can also request webinars on diverse topics or meetings led by SGCI staff to be presented at their home institutions. For example, they can ask for a webinar on software development sustainability via on-campus groups.

The partner program has been launched to build long-term partnerships between organizations, projects and the SGCI to benefit the community in multiple ways:

- Leverage the expertise of complementary organizations to support and advise the gateway community in the best possible way.

- Exchange existing services to provide the best array of services to gateway clients.
- Cross-promote resources and services offered by SGCI and partners.
- Provide a streamlined process to access resources and services provided by a partner.
- As appropriate, provide student exchange or internship opportunities.

Interested organizations collaboratively develop a Letter of Commitment that includes background information on the organization and a description of how the organization's services/resources are augmented by involvement with SGCI as well as how it is anticipated that SGCI will benefit from the partnership. While this program has started only recently, we plan to immediately reach out to our international collaborators of many years to form partnerships.

With this mixture of the outreach and interaction measures we aim at reaching out to and growing the existing science gateways community. Researchers might work with or develop science gateways without being aware of that their web-based solution is a science gateway and that there are more solutions and a vibrant community that could be beneficial for their own research. One of the reasons is that various terms are used to label science gateways, such as virtual research environments in Europe or virtual labs in Australia. While we do not try to unify the various labels, we strive for increasing the visibility of research and developments in the context of science gateways and bring the communities together irrespective of

the preferred expression or their locations.

B. Scientific Software Collaborative

The Scientific Software Collaborative (SSC) area of the Institute has the goal of engaging a variety of gateway developers—from researchers looking for end-to-end solutions to developers looking to contribute their software services—by promoting gateways, software and services available. The SSC will provide researchers the information they need to build their gateway using existing software or by leveraging existing gateways, enabling researchers to focus their efforts on their domain-specific challenges instead of the technology needed to get there.

The SSC will deliver software solutions that facilitate the following:

- End-to-end solutions to help developers from various scientific domains who want an “out-of-the-box” hosted gateway for their application, even if they have minimal IT support.
- A “use-what-you-need” approach to help developers discover available software and services to create their own customized gateways.
- Software for developers who want to actively extend open-source gateway software to meet their needs
- Discovery of gateways for those looking to find existing gateways they might want to use or collaborate with.

One of the initial milestones of the SSC will be to deliver a community Gateway Catalog. The Gateway Catalog is a

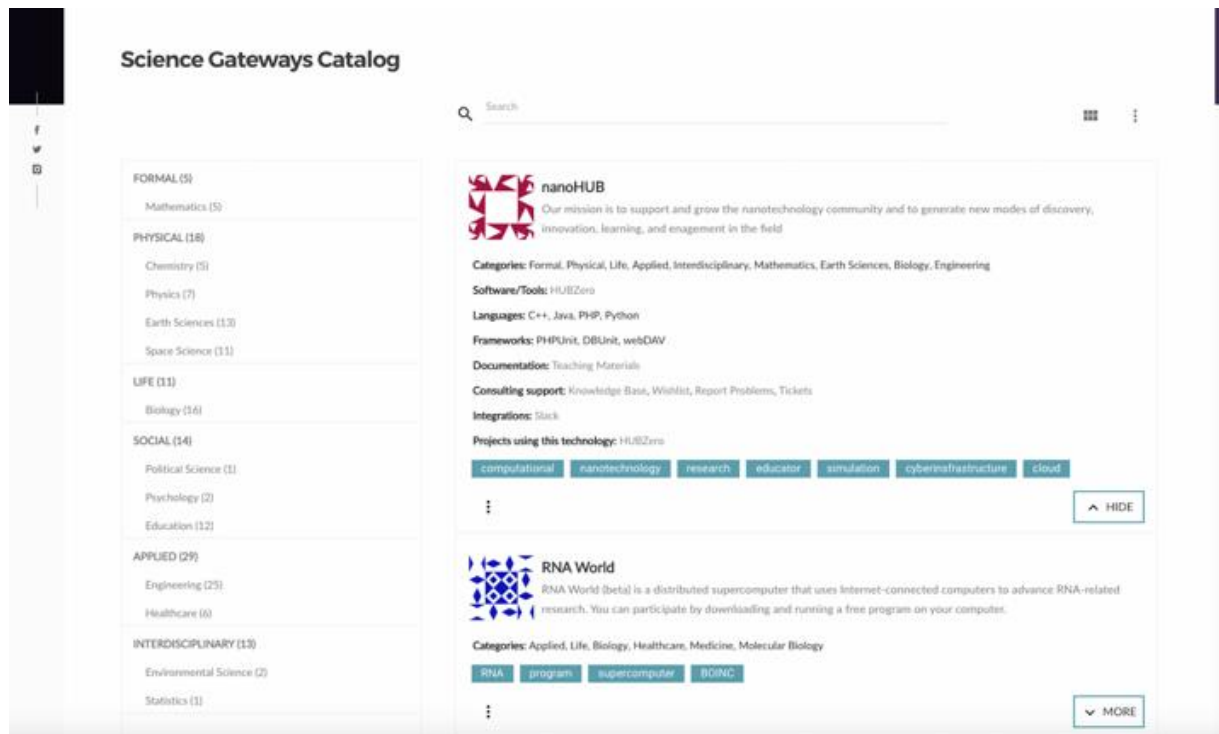


Figure 2: Prototype of the Gateway Catalog

publicly available, easy-to-use listing of available science gateways. The initial catalog will be seeded with entries the Institute is familiar with and will continue to grow with community contribution. The community, including international gateways, will have the ability to create accounts and add their gateway to the listing. The Gateway Catalog includes basic information such as name, description, URL, but also more in depth metadata such as language used to build the gateway, frameworks, documentation and more. This catalog is openly accessible and available for international contributions. Future iterations of the catalog will also have an API endpoint for automatically sharing and consuming gateway data. The catalog team will also look at working with national and international collaborators to share data in the catalog and consume data with other similar projects. With the changing landscape and the ever-growing science gateway environment researchers aren't limited to using technologies within their borders. Many technology products are open-source and available for use. By leveraging SGCI to offer an inclusive environment for promoting gateway projects across borders it opens up opportunities for usage from a diverse set of participants. An example prototype of the catalog is shown in Fig. 2. When the catalog is in production we will engage gateway developers from international arenas such as Europe and Australia to contribute their gateways to the listing.

As the SSC area continues to expand, we will also be promoting gateway software. We intend to list gateway frameworks as well as gateway components providing authentication, information services, workflows, etc. This offers additional opportunities for international collaborators to share software solutions via SSC. The international community has an opportunity to collaborate with the SSC area of SGCI in various forms as stated above, and the SSC is open to other opportunities to collaborate in other ways that advance the software infrastructure of developing science gateways. In addition to being contributors, international opportunities are available in the form of the affiliates program, which entails a commitment to support users of their software platforms or components. By having international affiliates become a part of the program we can leverage the expertise of complementary organizations to support and advise the gateway community in the best possible way. Furthermore we can cross-promote resources and services to provide the best support to gateway clients.

C. Incubator

The Incubator is focused on helping people who are at various stages of their gateway lifecycle—ranging from planning through operations—to step back from the effort and view it anew with a different set of eyes that are more focused on the business of gateways than on the technological or scientific aspects. We accomplish this through two primary activities. The first is a week-long in-person cohort training session called a “gateway bootcamp.” The goals of this session are to equip participants with that second set of eyes, to establish interpersonal relationships with others in the cohort engaged in creating gateways, and to familiarize participants with the in-depth consultations they might seek from the SGCI. The second is an in-depth consultation to focus on specific

identified areas where a gateway effort needs help. For gateways in the planning stage, this may take the form of helping to create a credible case for the value of the proposed gateway to its likely stakeholders. For gateways that are already operational, it may take the form of examining alternative ways they may measure impact to demonstrate to funders and other stakeholders the value that they bring to their community. For gateways in the mid-to-late funding stage, it may involve helping develop a sustainability strategy.

While SGCI cannot provide direct support to specific gateway projects based outside the US, our sponsor views international engagement opportunities quite positively and encourages ways for us to convey information to and learn from groups involved in gateways regardless of the country of origin. For example, an international group is joining our first “gateway bootcamp” as observers. Our goal with this observation is to work with other gateway support organizations as a mutual learning opportunity rather than to help a specific international gateway effort.

Our expectation therefore is to have a follow-up debrief to learn about the gateway-development climate in the observer's region, how it differs from our own, and how to adapt our message for international audiences. Given that user bases of gateways span national boundaries, this knowledge is important in helping gateways understand the implications of some of their sustainability strategies. Depending on the feedback in our first session, we may seek to have interested international participants participate as lecturers in future sessions.

The Incubator also can offer paid engagements with international organizations to offer training cohorts in their regions and with specific gateway efforts who may benefit from the advisory services we offer. Additionally, since the Incubator team includes some core SGCI personnel but also independent consultants, those consultants are able to offer their services without restriction from our funding agency. Finally, there are also granting agencies in the US and abroad that have specific initiatives to foster international collaboration. Provided that such efforts have a partially US-based team, we are able to offer our services to these efforts.

V. FIRST STEPS AND FUTURE WORK

There has been much interest in advancing the progress of science gateways (virtual laboratories, virtual research environments) internationally. The International Coalition of Science Gateways [22] has been proposed as an organization that would provide a framework for organizations, forums and individuals with an interest in international, regional, national, cross-institutional, domain-specific, and interdisciplinary initiatives in science gateways to exchange information. Several initial groups have expressed interest in participating. These included SGCI (USA), NeCTAR (Australia), NESI (New Zealand), Sci-GaIA (Africa), Academia Sinica Grid Computing Center (Taiwan), Software Sustainability Institute (UK), VRE4E1C (Europe), IWSG (Europe), CANARIE (Canada), Research Data Canada (Canada), the IEEE Technical Area on Science Gateways (International). First steps included drafting a survey paper that would provide an overview of

programs internationally, including successes and best practices as well as lessons learned.

Almost simultaneously in the Research Data Alliance, the Virtual Research Environments Interest Group (<https://www.rd-alliance.org/groups/vre-ig.html>) was proposed to “bring together those initiatives actively developing VREs, virtual laboratories and science gateways, along with the representatives of the common infrastructure services and the researchers seeking to make use of these technologies”. This effort is geared toward the definition of reference architectures, standard components, interfaces and best practices in the construction of science gateways. At current writing, this interest group consists of 68 members from all over the world. Meetings have been held at several RDA conferences and a case statement was developed for the group to review in March, 2017. The goal here is to move VRE-IG from an interest group to a working group. Working groups (WGs) have 18-month deliverables, milestones and work plans and their activities are a core focus of RDA. In contrast, interest groups are long-lived and strategic and are expected to 'spawn' WGs for specific 18-month tasks producing outputs. They promote communication and coordination and can produce surveys, recommendations and reports.

VI. SUMMARY

SGCI, still in its infancy, looks forward to engaging with the important international community. We are building on over 8 years of collaborative work with international organizations to organize gateway workshops and special journal issues to enhance communication among the community and recognize the contributions of the science gateway developer community. As we in SGCI build out our own services, we do it in the context of these ongoing international engagements and with an eye toward how we can most valuably share developments in and offerings of our Institute with the wider community. We hope to engage with other organizations building their own services with an eye toward international collaboration as well.

Via the combination of face-to-face meetings, publications, online resources, the partnership program, the SSC and the Incubator service area, the SGCI aims at not only reaching out to a wide and diverse community but also serving as an international focal point for science gateways. SGCI closely collaborates with researchers and creators of science gateways despite national or continental borders or political bans. We are aware of the diversity of challenges in international collaborations and would like to start a discussion to extend the existing measures and ideas about science gateways on an international scale. The SGCI has the opportunity to expand beyond currently existing partnerships, and even though it is located in the US, many measures such as the webinar series are independent of its location. As we do this, we would like to continually engage with that community to solicit feedback on current directions and adapt for the future.

ACKNOWLEDGMENTS

This research is supported by NSF Award Id ACI-1547611.

REFERENCES

- [1] <http://sciencegateways.org/>
- [2] Gesing, S., Wilkins-Diehr, N., Dahan, M., Lawrence, K., Zentner, M., Pierce, M., Hayden, L.B., and Marru, S. “Science Gateways: The Long Road to the Birth of an Institute”. Proc. of HICSS-50 (50th Hawaii International Conference on System Sciences), 4-7 January 2017, Hilton Waikoloa, HI, USA, 2017. <http://hdl.handle.net/10125/41919>
- [3] Wilkins-Diehr, N., Gannon D., Klimeck, G., Oster, S., Pamidighantam, S, “TeraGrid Science Gateways and Their Impact on Science”, IEEE Computer, volume 41, number 11, November 2008.. <https://doi.org/10.1109/MC.2008.470>.
- [4] Lawrence, K.A., Wilkins-Diehr, N., 2012. "Roadmaps, Not Blueprints: Paving the Way to Science Gateway Success." XSEDE12 Conference, July 16-20, 2012. ACM Digital Library, DOI=<http://doi.acm.org/10.1145/2335755.2335837>.
- [5] Lawrence, K. A., Zentner, M., Wilkins-Diehr, N., Wernert, J. A., Pierce, M., Marru, S., and Michael, S. (2015) Science gateways today and tomorrow: positive perspectives of nearly 5000 members of the research community. *Concurrency Computat.: Pract. Exper.*, 27:4252-4268. <http://dx.doi.org/10.1002/cpe.3526>.
- [6] <http://www.collab-ogce.org/ogce>
- [7] <http://iwsg.info/>
- [8] Patries Boekholt, Jakob Edler, Paul Cunningham, Kieron Flanagan. "Drivers of International collaboration in research". Final Report. 2009. http://www.eurosfair.prd.fr/7pc/doc/1266832886_drivers_of_international_cooperation_in_research.pdf
- [9] B. P. Abbott et al. (LIGO Scientific Collaboration and Virgo Collaboration) “Observation of Gravitational Waves from a Binary Black Hole Merger”. *Phys. Rev. Lett.* 116, 061102 – Published 11 February 2016
- [10] <http://hicc.org/>
- [11] <http://scholarspace.manoa.hawaii.edu/handle/10125/42288>
- [12] M. McLennan, R. Kennell, "HUBzero: A Platform for Dissemination and Collaboration in Computational Science and Engineering," *Computing in Science and Engineering*, 12(2), pp. 48-52, March/April, 2010.
- [13] Dooley, Rion et al. "Software-as-a-service: the iPlant foundation API." 5th IEEE Workshop on Many-Task Computing on Grids and Supercomputers (MTAGS) Nov. 2012.
- [14] Marru, Suresh, Lahiru Gunathilake, Chathura Herath, Patanachai Tangchaisin, Marlon Pierce, Chris Mattmann, Raminder Singh et al. "Apache Airavata: a framework for distributed applications and computational workflows." In *Proceedings of the 2011 ACM workshop on Gateway computing environments*, pp. 21-28. ACM, 2011.
- [15] Goecks, J, Nekrutenko, A, Taylor, J and The Galaxy Team. “Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences”. *Genome Biol.* 2010 Aug 25;11(8):R86
- [16] Stephen Crouch, Neil Chue Hong, Simon Hettrick, Mike Jackson, Aleksandra Pawlik, Shoaib Sufi, Les Carr, David De Roure, Carole Goble, and Mark Parsons. The software sustainability institute: Changing research software attitudes and practice. *Computing in Science & Engineering*, 15(6):74–80, 2013.
- [17] <https://www.software.ac.uk/>
- [18] <http://wssspe.researchcomputing.org.uk/>
- [19] <https://www.vre4eic.eu/>
- [20] <http://www.envriplus.eu/>
- [21] <https://dibbs17.org/report/Presentations/KeynoteQualters.pdf>
- [22] <http://www.icsciencegateways.org>