Starting with a Firm Foundation: Building a Sustainability Program for the Australian Research Data Commons

Claire Stirm
San Diego Supercomputer Center
University of California, San Diego
La Jolla, California, USA
0000-0003-4188-137X
cstirm@ucsd.edu

Kerry Levett
Australian Research Data Commons
Adelaide, South Australia, AUS
kerry.levett@ardc.edu.au

Nancy Maron
BlueSky to BluePrint Consulting
New York, New York, USA
nancymaronconsulting@gmail.com

Sandra Gesing
Discovery Partner Institute
University of Illinois Chicago
Chicago, Illinois, USA
sgesing@uillinois.edu

Andrew Treloar
Australian Research Data Commons
Melbourne, Victoria, AUS
0000-0002-8911-3081
andrew.treloar@ardc.edu.au

Siobhann McCafferty
Australian Research Data Commons
Brisbane, Queensland, AUS
siobhann.mccafferty@ardc.edu.au

Juliana Casavan
Casavan Consulting
West Lafayette, Indiana, USA
julianabcasavan@gmail.com

Abstract—**The United States Science Gateways Community Institute (SGCI) and the Australian Research Data Commons (ARDC) have offered their communities support and training around project sustainability. The support given has varied from in-person sustainability training, such as SGCI's Focus Week, direct advice giving, planning consultation, and personnel support to make sure the project continues to succeed within the traditional academic environment. All of these methods have helped research projects in the United States and in Australia.**

In 2020, SGCI and ARDC initiated the planning of a Sustainability Program. The Sustainability Program was based on curriculum developed by the SGCI instructors, while being customized to cover core topics with follow-on cohort feedback sessions with the ARDC program staff. The Sustainability Program gave projects in-depth training on core strategies to "think like a business" while operating in an academic environment; technology best practices for science gateway user-interfaces; and long-term sustainability strategies to receive continued support.

Index Terms—**sustainability, training, science gateways, virtual research environments**

I. INTRODUCTION

Sustainability continues to be a widely discussed topic within academic research, especially in relationships with research software projects. Over 90% of researchers answering surveys say that they use software for their research and over 65% express that they even could not do their research without software [1]. With more and more projects dependent on research software and science gateways, sustainability of such solutions is crucial to allow researchers to focus on their work instead of setting up or re-developing frameworks when availability of existing research software and science gateways disappear when financial or community support goes away [2]. Sustainability of science gateways has many facets: from technical aspects such as good practices in software engineering to usability of science gateways to community building practices to attract a large community. Widely used science gateways have better chances to be further funded or supported, one further aspect of sustainability. Diversifying funding contributes to ongoing operations of science gateways and being able to develop additional features beneficial for the community.

Organizations such as the United States Science Gateways Community Institute (SGCI) [3]–[5] and the Australian Research Data Commons (ARDC) [6] offer support and training to projects navigating sustainability challenges and potential paths towards a more secure future. In 2020, these two organizations formed a partnership to deliver a unified, customized sustainability program to projects supported by ARDC. In this paper we discuss the preparation and delivery of the ARDC Sustainability Program. Additionally, we present the outcomes of the training.

II. Background

Researchers who build science gateways are often seeking to address, understand, and solve a specific problem. These challenges could pertain to urban development planning, monitoring fish populations, or certifying modeling tools for researchers. Curious researchers find inspiration in the challenge and seek to find a solution that positively impacts broader audiences, such as their fellow research peers, students, policy makers, or the general public. Once a solution has been conceptualized or produced in the form of a product or service, follow-on questions emerge.

- How can the project team define the value of their work in a way that inspires opportunities for continued support?
- How can the project reach intended audiences or even create change in a societal system?
- How can the project further develop within the means provided in an academic community?

These questions stem from a need for sustainability in order to continue the efforts of the project.

Cyberinfrastructure projects combining access to data with software tools and underlying compute resources are commonly referred to as virtual research environments, science gateways, or research platforms. These projects straddle the line between computer science and hard sciences, as they bring together these two worlds to provide digital access to cutting-edge research models, educational softwares, or a central interactive data repository. Due to their blended nature, they need a team that has computer science and engineering, system administration, and research skills. They also have a continued need for software and hardware to enable consistent access for their audiences, while incorporating additional resources as needed when growth occurs. All of these needs add up. Often projects receive a lump sum of funding at the initiation of an idea yet struggle to receive a consistent form of funding without additional buy-in from new audiences or new support groups. Sometimes the sustainability challenge is not always monetarily driven but driven by a need for expertise in specific development practices or community engagement.

These discussions and more have emerged from peer workshop series such as WSSSPE (Workshop on Sustainable Software for Science: Practice and Experiences) [1], on-campus conversations with researchers [7], journals such as JORS (Journal of Open Research Software) [8], and funding agencies such as the United States National Science Foundation (NSF) [9]. Other professional organizations carry forward these thought pieces such as the UK Software Sustainability Institute (SSI) [10], the US Research Software Engineers Association (US-RSE) [11], the consortium of Advanced CyberInfrastructure - Research and Education Facilitators (ACI-REF) [12], and the United States Science Gateways Community Institute (SGCI) [13].

In August 2016, the Science Gateways Community Institute (SGCI) was created to provide subsidized services and resources to the developers and users of science gateways. During the first year of SGCI, founder Nancy Wilkins-Diehr, presented at the University of Queensland with an overview of SGCI's offerings and interest in collaborations including the International Coalition on Science Gateways and the International Workshop on Science Gateways [14]. This foundation led to continued following and conversations between SGCI and ARDC.

SGCI developed a science gateway sustainability training model, Focus Week, to train science gateway owners in all lifecycle stages of the core sustainability tools [15]. Three main areas are addressed in the training:

1) Core business strategy skills as they apply to leading an online digital presence, such as understanding stakeholder and user needs; business, operations, finance, and resource planning; marketing and project management.
2) Technology best practices, including the principles of usability and user-centered design for science gateway environments.
3) Long-term sustainability strategies, such as alternative funding models; case studies of successful gateway efforts; licensing choices and their impact on sustainability.

The training is set up for five days as an interactive workshop with a maximum number of ten teams to be accepted for each workshop. The teams consist ideally of project members in diverse roles such as the science gateway's Principal Investigator, project managers, lead research scientists, community managers, or developers. The first Focus Week, which was previously called Bootcamp, was held in April 2017. Ten teams were in attendance with one observer from Nectar Australia, the research cloud founded in 2013 and that has been applied by ARDC from 2018. This observation of the SGCI provided training had the initial goal to initiate similar conversations and training in Australia. This interaction led to a partnership forming between Nectar and SGCI. The feedback of the participants showed that the concept was found beneficial, thought-provoking and entertaining with some room for improvement.

Past the initial delivery of Focus Week, SGCI has continued to provide the United States and international communities with sustainability training sessions, largely held as in-person workshops. Twenty-one training sessions have been delivered by SGCI Instructors to over 670 participants that work on science gateways or research projects. In one example, a shortened sustainability course was delivered to attendees at the 2018 International Workshop for Science Gateways [16]. In December 2020, SGCI delivered the Focus Week sustainability training as a two-week virtual course to 57 attendees. While the majority were project teams from the United States, there were attendees joining internationally including staff members from the Australian Data Research Commons (ARDC).

The ARDC is funded by the Australian Government's National Collaborative Research Infrastructure Strategy and is designed to accelerate Australian research and innovation by driving excellence in the creation, analysis and retention of high-quality data assets [6]. The ARDC partners with the research community and industry to build leading-edge digital

research infrastructure and runs a series of programs including the ARDC Platforms Program.

The ARDC Platforms Program seeks to enable transformative research across all disciplines using advanced software and platforms and supports 26 projects that are building platform infrastructure (i.e. virtual research environments) between 2019-2023. After participating in the December 2020 Focus Week, the ARDC team partnered with the SGCI Focus Week instructors to design and deliver a sustainability training program that would meet the needs of project teams supported by the ARDC.

III. SGCI FOCUS WEEK CURRICULUM

The SGCI Focus Week is an intensive workshop designed for innovative research teams to work together on producing a strong sustainability plan [17]–[19]. Participating projects leave Focus Week with a clearer definition of their project's value, its audience, and its positioning in the competitive landscape. Below are the core Focus Week exercises that teams complete.

- Napkin Drawing: Learn how to effectively communicate your project.
- Understanding Your Audiences & Key Stakeholders: Identify who cares about your project, and determine why they care; explore potential new user groups.
- Environment-Mapping the Landscape: Spend time researching and mapping out your known and new competition, as well as open opportunities where your project is applicable.
- Marketing Tactics Tools: Learn how to communicate your project's value to your audiences with selective marketing.
- Goal Setting: Think about the big picture ideas. Learn how you can set the right goals for the right reasons, and learn how you can measure your success.
- Value Proposition: Build a concise value statement that articulates the unique value your project delivers to its users.
- Budgeting: Discover how you can forecast a budget that will help you plan for life beyond the grant.
- Market Development: Explore the possible customer groups and subgroups that will find value in your project other than the original audience your project is intended to serve.
- User-Centered Design: Learn from a usability expert on best practices when designing cyberinfrastructure user interfaces.

The last day of the workshop is a "Pitch Day" with each team presenting their sustainability plans. This is a very rewarding experience for not only the presenting teams but their fellow cohort members as they can see how each other have grown upon each exercise and the outcome of putting all the exercises into one conclusive presentation. Many of SGCI's past attendees have shared their experiences from attending sustainability training sessions in the SGCI storybook [20]. Here are a few notable quotes from past participants:

"ESIP Lab has funded over 20 projects since we attended Focus Week, and I'm able to take the tools that I learned and just use that language and pass it on to those projects. It gave me a mindset to continually evaluate and re-evaluate the Lab, too—Is our value proposition the same? What's our niche?" — Annie Burgess, ESIP Lab

"We're so grateful for Focus Week because it is exactly the right thing for people who want to expand, broaden, and capitalize on their gateways. You can't do any of that without the training and resources provided by Focus Week. It was such an eye-opening experience for us and it remains, behind the scenes, what keeps us from going over a cliff." — Jason Fleming, Coastal Emergency Risks Assessment Tool

IV. RE-IMAGINING FOCUS WEEK FOR ARDC

Through attending the virtual Focus Week in 2020 hosted by SGCI, ARDC staff started to explore a collaboration: a jointly-run sustainability program for ARDC projects. Taking the Focus Week curriculum as a starting point, the SGCI team was asked to modify and customize the program to suit ARDC project leads. Rather than a one-week, in-person intensive, could Focus Week be presented as a series of virtual, interactive events, delivered over time? Working together, the SGCI team met with the ARDC team to develop a plan that would provide as much content as possible, with teams many time zones apart.

To address the time-zone differences, the teams determined that evening sessions in the US would work as morning sessions for Australian participants, and to address Zoom-fatigue, it was suggested (and eagerly adopted) that no sessions run longer than 90 minutes. A major benefit of in-person delivery of the Focus Week program - both to instructors and to participants - has been the ability to interact with project leaders after each topic has been discussed, and the ability for participants to take the time to work through the new material, and apply what they are learning to their own projects. To build in feedback while using a virtual delivery approach, the ARDC program was re-structured:

- First, SGCI instructors would present a core topic, in a virtual session, encouraging interactivity as much as possible.
- Each day's presentation was then immediately followed by an SGCI ARDC debriefing session, permitting the SGCI instructors to meet with ARDC program leads to review the material covered, and prepare for the work sessions to come
- A formal "work session" for participants was then led by ARDC staff to provide feedback to project teams and answer questions.

The curriculum was customized to cover core topics with follow-on cohort feedback sessions with the Platforms program staff after each Tuesday and Wednesday session - see Table 1 for details.

TABLE I
ARDC SUSTAINABILITY PROGRAM TIMELINE

	Product/Service - Definition of Value	
	CORE TOPIC (90 minutes)	*COHORT FEEDBACK (60 minutes)*
21 Feb	Introduction to Sustainability. Drawing: Effectively communicate the value of your project through verbal and visual communication.	*None*
22 Feb	Audience: Explore and assess your audiences and stakeholders to better engage them.	Team Breakouts, Feedback to teams from ARDC Platforms team
23 Feb	Landscape: Define who your competitors/potential collaborators are and how you differentiate your product from theirs	Team Breakouts, Feedback to teams from ARDC Platforms team
24 Feb	Value Proposition: Identify the primary value that your project brings to its users and community.	*None*
	Defining Your Funding Model	
	CORE TOPIC (90 minutes)	*COHORT FEEDBACK (60 minutes)*
3 Mar	Goal Setting & Budgeting (2h): Use impact-driven goals to begin developing a budget and financial forecast	*None*
9 Mar	Revenue Models: Exploration into different revenue types that can work within the academic world	Team Breakouts, Feedback to teams from ARDC Platforms team
17 Mar	Sales/Marketing: Learn principles of sales and marketing, and how to develop a plan to connect with your target audience and promote your gateway.	Team Breakouts, Feedback to teams from ARDC Platforms team
23 Mar	Sustainability Strategy "Pitch": Each team presents the exercises they completed during the workshop in the form of a "pitch"	*None*

An important element that was incorporated into the ARDC Sustainability Program was time for feedback after core training sessions. These feedback sessions were scheduled to allow projects to have time with the ARDC staff members to share their concepts from the core topic exercise and discuss. This feedback allowed the projects to share ideas with their support team and it provided ARDC staff the opportunity to help advise on workable approaches for next steps.

V. ARDC SUSTAINABILITY PROGRAM DELIVERY

The ARDC Sustainability Program was held over the course of two months. The first part of the program was delivered over the course of one week, Monday, February 21, 2022 through Thursday, February 24, 2022. The sessions delivered are described in Table 1: Part 1. The goal of delivering all of these sessions in one concurrent week was to lay the foundation for sustainability practices and build enthusiasm for the concepts. The second part of the program was delivered as individual sessions once a week, described in Table 1: Part 2. The second part of the program was built on top of the foundation week with time in between sessions to allow for research and deeper team conversations to take place. The ending session was "Pitch Day". A total of 74 attendees participated in the ARDC Sustainability Program. Additionally, 16 projects delivered their completed "Pitch Day" presentations at the end of the program.

VI. FEEDBACK FROM ARDC SUSTAINABILITY PROGRAM PARTICIPANTS

At the end of the program, participants expressed to ARDC staff that they found the program highly valuable. With the goal of capturing some of this feedback, an exit survey was sent to the teams that attended the entire program.

An open text question at the end of the survey asked respondents to "share any additional comments, questions, or concerns." Three survey participants shared:

- "Our platform made a lot of valuable progress in these sessions and spent a lot of time outside of the sessions to work on the activities. We ended up with a better understanding of our platform. I feel like there should be more focused workshops like these for platforms in the future."
- "It was much more useful than I expected it to be. I think what our team has learnt will go a long way towards helping create a sustainable platform. Thanks very much!"
- "Great initiative!"

One survey participant also pointed out the value in being able to meet with the ARDC team during program feedback sessions.

"I think the afternoon sessions where teams were able to meet and work on the exercises were really valuable. We had a breakthrough in one session when we locked in our value propositions and another when we identified the two axes of our market

landscape. We definitely needed the time in these sessions to have those discussions."

Participants were asked in the final survey, "How important do you think the following components are to the success of your project?". There were ten curriculum items they were asked to rate with the options of extremely important, very important, moderately important, slightly important, and not at all important. The following response was provided to this question:

- Basics of sustainability strategy
  - 100% of survey participants rated this as "Extremely important"
- Understanding your audience
  - 100% of survey participants rated this as "Extremely important"
- Competitive landscape
  - 50% of survey participants rated this as "Extremely important"
  - 50% of survey participants rated this as "Very important"
- Value Proposition
  - 75% of survey participants rated this as "Extremely important"
  - 25% of survey participants rated this as "Slightly important"
- Market development
  - 75% of survey participants rated this as "Extremely important"
  - 25% of survey participants rated this as "Slightly important"
- Goal setting
  - 25% of survey participants rated this as "Extremely important"
  - 75% of survey participants rated this as "Very important"
- Budgeting
  - 75% of survey participants rated this as "Extremely important"
  - 25% of survey participants rated this as "Very important"
- Revenue Models compatible with Open Education
  - 25% of survey participants rated this as "Extremely important"
  - 75% of survey participants rated this as "Very important"
- Sales Marketing
  - 50% of survey participants rated this as "Extremely important"
  - 50% of survey participants rated this as "Very important"
- Delivering a pitch
  - 50% of survey participants rated this as "Extremely important"
  - 25% of survey participants rated this as "Very important"
  - 25% of survey participants rated this as "Slightly important"

Participants were also asked, "How would you rate the length of the series" in which the choices were way too long, a bit long, just right, a bit too short, and way too short. To this question, 75% of participants shared that it was the right length and 25% of participants found the series to be a bit too short. Finally, participants were asked, "Overall, how well did the series meet your expectations" in which the choices were extremely well, very well, moderately well, slightly well, and not well at all. Survey participants responded with 50% saying extremely well, 25% very well, and 25% slightly well.

The incorporation of feedback went past the initial program delivery. Based on the SGCI's practice of holding follow-on sessions to connect with teams after the sustainability program, the ARDC team scheduled a three-month follow-on session to meet with projects and hear updates on goals teams had set for themselves. Platform teams met the majority of their three-month goals, particularly those that had set goals specific to investigating or implementing sustainability activities such as engaging effectively with their audiences, or developing operational budgets.

Both the SGCI and ARDC teams have viewed this collaboration as a success. ARDC's project Steering Committee noted that this work permitted project teams to begin looking beyond the time-bound project funding and duration, and develop the knowledge required for projects to become ongoing operational infrastructure. Sustainability is now considered a standing agenda item for them. ARDC staff stated that it was beneficial for them to participate as well as they heard the same language as the participating projects and could reflect on ARDC practices through this sustainability lens.

VII. CONCLUSION

The ARDC Sustainability Program provided an opportunity to deliver sustainability training to projects supported by ARDC which benefited teams gearing up to initiate sustainable paths to continue supporting their project's goals. Additionally, this opportunity provided ARDC and SGCI the space to partner on a larger shared vision, to provide education on how projects can continue past initial support, and to have a shared space to discuss what it means for a project to be sustainable with project teams.

As the ARDC develops its Thematic Research Data Commons (RDCs), it is taking key learnings from the course into the design of new programs. Future projects supported by the ARDC will be supported to plan for sustainability from the beginning of the project by, for example, creating activity-based budgets, and engaging more deeply with their user communities.

As SGCI funding from NSF has begun to conclude, a new Center of Excellence has been awarded by NSF to continue providing sustainability training and other services to the science gateways community. Called the Center of

Excellence to Extend Access, Expand the Community, and Exemplify Good Practices for CI through Science Gateways (SGX3), the science gateways community will continue to be offered sustainability training through in-person Focus Week workshops and virtual Jumpstart short courses [21].

Acknowledgment

References

[1] Katz, D. S., Niemeyer, K. E., Gesing, S., Hwang, L., Bangerth, W., Hettrick, S., Idaszak, R., Salac, J., Chue Hong, N., Corrales, S. N., Allen, A., Geiger, R. S., Miller, J., Chen, E., Dubey, A., Lago, P. (2017). Report on the Fourth Workshop on Sustainable Software for Science: Practice and Experiences (WSSSPE4), arXiv:1602.02296 [cs.SE]

[2] Nangia, U, and Katz, D. (2017). Survey of National Postdoctoral Association - Dataset [Data set]. Zenodo. https://doi.org/10.5281/zenodo.843607

[3] Wilkins-Diehr, N., and Crawford, T. (2018). NSF's Inaugural Software Institutes: The Science Gateways Community Institute and the Molecular Sciences Software Institute, in Computing in Science Engineering, vol. 20, no. 5, pp. 26-38, Sep./Oct. 2018, doi: 10.1109/MCSE.2018.05329813.

[4] Wilkins-Diehr, N., Zentner, M., Pierce, M., Maytal Dahan, Lawrence, K., Hayden, L., and Mullinix, N. (2018). The Science Gateways Community Institute at Two Years. In Proceedings of the Practice and Experience on Advanced Research Computing (PEARC '18). Association for Computing Machinery, New York, NY, USA, Article 53, 1–8. https://doi.org/10.1145/3219104.3219142

[5] Gesing, S., et al. (2017). 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, http://hdl.handle.net/10125/41919

[6] Australian Research Data Commons - ARDC. (2023). Retrieved April 4, 2023, from https://ardc.edu.au/

[7] Gesing, S., Lawrence, K., Dahan, M., Pierce, M. E., Wilkins-Diehr, N., Zentner, M. (2019). Science gateways: Sustainability via on-campus teams, Future Generation Computer Systems, Volume 94, Pages 97-102, ISSN 0167-739X, https://doi.org/10.1016/j.future.2018.09.067.

[8] Journal of Open Research Software. (2023). Retrieved April 4, 2023, from https://openresearchsoftware.metajnl.com/

[9] National Science Foundation - NSF. (2023). Retrieved April 4, 2023, from https://www.nsf.gov/

[10] The Software Sustainability Institute. (2023). Retrieved April 4, 2023, from https://www.software.ac.uk/

[11] The United States Research Software Engineer Association. (2023). Retrieved April 4, 2023, from https://us-rse.org/

[12] Aci-Ref. (2023). Ref. ACI. Retrieved April 4, 2023, from https://aci-ref.github.io/

[13] Gesing, S., Dahan, M., Zentner, M., Wilkins-Diehr, N., and Lawrence, K. (2019). The Science Gateways Community Institute: Collaborations and efforts on international scale, Future Generation Computer Systems, Volume 101, Pages 951-958, ISSN 0167-739X, https://doi.org/10.1016/j.future.2019.07.024.

[14] Wilkins-Diehr, N. (2016). The Science Gateways Community Institute. University of Queensland Research Computing Centre Seminar Series. Retrieved April 4, 2023, from https://rcc.uq.edu.au/event/737/science-gateways-community-institute

[15] Gesing, S., et al. (2017). Science Gateways Incubator: Software Sustainability Meets Community Needs, 2017 IEEE 13th International Conference on e-Science (e-Science), Auckland, New Zealand, 2017, pp. 477-485, doi: 10.1109/eScience.2017.77.

[16] IWSG 2018. (2023). IWSG 2018 Program. Retrieved April 4, 2023, from https://sites.google.com/a/nd.edu/iwsg2018/program

[17] Gesing, S., Zentner, M., Cleveland, S., Casavan, J., Craddock, R.C., Stirm, C., and Hoebelheinrich, N. (2018). Science Gateways Community Institute Incubator Pitch Deck: Success Stories from the 2nd & 3rd Bootcamp.

[18] Gesing, S., Zentner, M., Casavan, J., Heiland, R., Hillery, B., Marru, S., Pierce, M., Maron, N., Mullinix, N. and Vorvoreanu, M., (2017). Science Gateways Bootcamp: Strategies for Developing, Operating and Sustaining Science Gateways.

[19] Maron, N. (n.d.). Publications. BlueSky to BluePrint. Retrieved April 4, 2023, from https://www.blueskytoblueprint.com/publications

[20] Science Gateways Community Institute. (2022). SGCI storybook Connecting people, creating solutions, accelerating discovery. sciencegateways. Retrieved April 4, 2023, from https://sciencegateways.org/about/storybook

[21] SGX3. (2023). Center of Excellence to Extend Access, Expand the Community, and Exemplify Good Practices for CI through Science Gateways. sciencegateways. Retrieved April 4, 2023, from https://sciencegateways.org/