How to Position Your Gateway for Success: Ten Good Practices

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

Kerk F. Kee
College of Media and Communication
Texas Tech University
Lubbock, TX, USA
kerk.kee@ttu.edu

Sean B. Cleveland
University of Hawaii-System
Honolulu, HI, USA
seanbc@hawaii.edu

Annelie Rugg
University of California, Los Angeles
Los Angeles, CA, USA
anneliergug@ucla.edu

Steven R. Brandt
CCT
Louisiana State University
Baton Rouge, LA, USA
sbrandt@cct.lsu.edu

***Abstract*—Science gateways have gained a lot of traction in the last twenty years, as evidenced by projects such as the Science Gateways Community Institute (SGCI) and the Center of Excellence on Science Gateways (SGX3) in the US, The Australian Research Data Commons (ARDC) and its platforms in Australia, and the projects around Virtual Research Environments in Europe. A few mature frameworks have evolved with their different strengths and foci and have been taken up by a larger community such as Hubzero, Tapis, and Galaxy. However, even when gateways are built on successful frameworks, they continue to face the challenges of ongoing maintenance costs and how to meet the ever-expanding needs of the community they serve with enhanced features. It is not uncommon that gateways with compelling use cases are nonetheless unable to get past the prototype phase and become a full production service, or if they do, they don't survive more than a couple of years. While there is no guaranteed pathway to success, it seems likely that for any gateway there is a need for a strong community and/or solid funding streams to create and sustain its success. With over twenty years of examples to draw from, we present in this paper ten factors common to successful and enduring gateways that effectively serve as good practices for any new or developing gateway.**

***Index Terms*—science gateways, good practices, sustainability**

I. Introduction

Science gateways have proven to be invaluable environments for community collaborations, especially around data and computational tool sharing. They encompass not just those environments known as 'gateways,' but also those referred to as 'platforms', 'virtual research environments,' or 'dashboards.' The extensive uptake of the concept is evident not just in the instances of the gateways themselves, but also in related events and publications. However, creating and sustaining a successful gateway is not a trivial task, and for all the successes there are at least as many that did not succeed [1]. Sustaining a gateway requires careful planning, design, development, deployment, and maintenance, as well as attention to the needs and expectations of the target user community [2]. Early success is often dependent on initial funding that facilitates adoption by a sizeable community. Early success and growth in use is also often the result of taking an open source approach that enables the community to further develop and create their own instances. Even with these initial successes, fledgling gateway communities tend to underestimate the maintenance resources required, while in other cases, users who are not familiar with the development processes needed to extend features for their individual instances are stalled without access to further funding. With so much gateways history behind us, we are in a position to help newer gateways to not repeat earlier mistakes. Instead, in this paper, we suggest ten good practices for characteristics of gateways that are more likely to position gateway developers and providers for success. These characteristics fall into two groups: technical and socio-economic. All are equally important to a gateway's success and adoption and complement each other.

II. Background

In the last two decades quite a few mature frameworks have evolved with different strengths and foci and also quite a few that failed to stay sustainable and are not available in the ecosystem any more. For this paper we picked three successful examples that we are experts in. An exhaustive list of all successful frameworks is not within the scope of this paper. Hubzero [3], Tapis [4] and Galaxy [5] were selected because these are the ones we have the most experience with and we know from years working with these frameworks and the involved teams which factors contributed to the success. Hubzero and Galaxy are complete solutions that enable users to install a standard version with basic features. Tapis follows the concept that it exists of standard-based services and a developer can spin up efficiently a gateway tailored to a specific community. All three solutions are fostered by an evangelist team and take into consideration some of the main challenges for success. Main challenges in developing science gateways include gateway sustainability, security, usability,

scalability, and interoperability. Despite these challenges, science gateways have many benefits such as providing access to advanced resources for science and engineering researchers, educators, and students. They can also streamline online user-friendly interfaces that combine a variety of cyberinfrastructure components in support of a community-specific set of tools, applications, and data collections.

III. Good Practices

We developed two groups of good practices. We would like to emphasize these are good practices to contribute to the sustainability and we see them as the minimum but not as the only measures to guarantee success. Broadly speaking, the first five practices (i.e., adaptability, extensibility, & scalability; usability & accessibility; reproducibility; FAIR-compliance; standard-based) are practices for technical design of gateways, while the last five practices (i.e., evangelism; community & advocacy for funding/stakeholder buy-in; visibility & awareness; use cases & success stories; documentation and onboarding resources) are practices for social organization of gateways. We see the success of gateways as the outcome of socio-technical strategies, and one without the other would be insufficient and/or incomplete efforts.

1) Adaptability, Extensibility, and Scalability
   A gateway should be able to adapt to changing user requirements, technologies, and cyberinfrastructure capabilities over time. It should also be extensible to support new features, functionalities, and integrations with other systems and services. One of the key design principles for a gateway is flexibility. A gateway should not be a rigid or static system, but rather a dynamic and adaptable one. A gateway should be able to accommodate the evolving needs and preferences of its users, as well as the changing landscape of technologies and cyberinfrastructure capabilities that it relies on or interacts with. A gateway should also be extensible, meaning that it should allow for the addition of new features, functionalities, and integrations with other systems and services, without compromising its performance, security, or usability. A flexible and extensible gateway can enhance the user experience, increase the gateway's value and impact, and foster innovation and collaboration. Hubzero, for example, integrates with successful environments such as Jupyter Notebook and RStudio. Tapis can easily scale up via the RESTful services.
2) Usability and Accessibility
   A good gateway should have a user-friendly design that makes it easy to use and understand for its intended users, regardless of their level of expertise, background, or preferences. A good gateway should also be accessible to users with different abilities, devices, and network conditions. This means that the gateway should follow the principles of universal design and web accessibility standards, such as providing clear labels, instructions, feedback, and alternative formats for different types of content. A good gateway should also be responsive and adaptable to different screen sizes, resolutions, orientations, and input methods. A good gateway should also be reliable and secure, ensuring that the data exchange is fast, accurate, and protected from unauthorized access or modification. All three of our examples are designed with these factors in mind.
3) Reproducibility
   A gateway should enable users to reproduce their own or others' research results using the same data, methods, parameters, and cyberinfrastructure resources. It should also provide provenance information and documentation for transparency and accountability. A gateway is a tool that facilitates scientific discovery and collaboration by allowing users to access, share, and reuse data, methods, parameters, and cyberinfrastructure resources. One of the key features of a gateway is that it enables reproducibility of research results, both by the original researchers and by others who want to verify or extend their work. A gateway should also provide provenance information and documentation for each step of the research process, such as the data sources, the methods used, the parameters chosen, and the results obtained. This way, a gateway can ensure transparency and accountability of the research conducted through it. All three gateway frameworks enable to share data and simulations in a project and/or community.
4) FAIR-compliance
   A gateway should follow the FAIR principles of making its data and services Findable, Accessible, Interoperable, and Reusable by humans and machines [4]. It should also adhere to relevant standards, protocols, formats, and metadata schemas for data and service interoperability. Since a gateway is a platform that provides access to data and services for a specific scientific domain or community. To ensure that the data and services are of high quality and can be easily discovered, integrated, and reused by both humans and machines, a gateway should follow the FAIR principles. These principles are:

   - Findable: Data and services should have unique and persistent identifiers, rich metadata, and be registered or indexed in searchable resources.
   - Accessible: Data and services should be retrievable by their identifiers using standard protocols that are open, free, and universally implementable. Metadata should be accessible even when the data are no longer available.
   - Interoperable: Data and services should use a formal, shared, and broadly applicable language for knowledge representation, and use vocabularies that follow FAIR principles. Data and services should also include qualified references to other data and services.
   - Reusable: Data and services should be well-described with accurate and relevant attributes, clear

and accessible data usage licenses, detailed provenance, and meet domain-relevant community standards.

By following the FAIR principles [6], a gateway can enhance the value and impact of its data and services for the scientific community and society at large. All three of our chosen examples have an open-source version available.

5) Standards-based, Priorities and/or Maintainability
A gateway should use established standards and best practices for its architecture, design, development, testing, deployment, operation, security, and sustainability. It should also leverage existing tools, frameworks, libraries, platforms, and services that comply with these standards. Since gateways are often crucial components of a distributed system that connects different networks, protocols, and applications. It must act as an intermediary between the clients and the servers, providing various functionalities such as routing, load balancing, authentication, authorization, encryption, caching, logging, monitoring, and more. To ensure the quality, reliability, performance, and security of a gateway, it is essential to follow established standards and best practices throughout its lifecycle. These standards and best practices cover various aspects of the gateway's architecture, design, development, testing, deployment, operation, security, and sustainability. They also help to avoid common pitfalls and challenges that may arise in complex and dynamic environments. Moreover, a gateway should leverage existing tools, frameworks, libraries, platforms, and services that comply with these standards and best practices. These can help to reduce development time and cost, improve the maintainability and scalability, and enhance the interoperability and compatibility of the gateway with other systems. In the realm of security, Gateways present unique challenges. The current best practices for dealing with this issue are enunciated in [7] and embodied in the principle of providing the minimum access needed by the application [8]. For all the standards and best practices listed here, it is important to remember that they are moving targets, that new and better ways of maintaining Gateways are being developed. Tapis' architecture, for example, is based on RESTful services, which are a standard in web development.
6) Evangelism
A gateway should have a dedicated person or team who can act as an evangelist for the gateway's vision, mission, goals, and benefits. They should also be able to advocate for the gateway's needs, challenges, and opportunities among different audiences and forums. One of the key roles in a gateway project is that of a gateway champion. This is someone who can communicate effectively the vision, mission, goals, and benefits of the gateway to various stakeholders and potential users. A gateway champion can also represent the gateway's interests and challenges in different venues and platforms, and seek out opportunities for collaboration, funding, and outreach. A gateway champion can be an individual or a team, depending on the size and scope of the gateway. Having a dedicated gateway champion can help increase the visibility, impact, and sustainability of the gateway [9].
7) Community and Advocacy Funding/Stakeholder (Institutions, Communities) Buy-in
A gateway should secure adequate funding and resources for its development and operation from various sources such as grants, sponsors, partners, users, or revenue streams. It should also obtain buy-in and support from its key stakeholders such as domain experts, cyberinfrastructure providers, researchers, educators, students, and policymakers. One of the main challenges for a gateway project is to ensure its sustainability and growth over time. This requires a strategic approach to securing and managing the financial and human resources needed to build and maintain the gateway. A gateway should diversify its funding sources and explore various options such as applying for grants, seeking sponsors, establishing partnerships, charging fees, or generating revenue from other activities. A gateway should also engage and communicate with its key stakeholders and demonstrate the value and impact of the gateway for its target community. This includes domain experts, cyberinfrastructure providers, researchers, educators, students, and policy makers. By obtaining their buy-in and support, a gateway can increase its visibility, credibility, and usability.
8) Visibility and Awareness
A gateway should raise awareness about its project, person, community, network, and impact among potential and existing users, stakeholders, funders, and collaborators. It should also provide mechanisms for feedback, communication, outreach, and dissemination. A gateway is a platform that enables collaborations of users, stakeholders, and funders on a specific project, community, or network. A gateway has a dual role of showcasing and engaging. It showcases the value, achievements, and impact of its project, person, community, or network to potential and existing audiences. It not only enables research and democratizes scientific discovery, but also engages its users in meaningful ways through feedback, communication, outreach, and dissemination channels. A gateway should aim to raise awareness about its purpose, vision, goals, and activities among its target groups and beyond. It should also seek to foster a sense of belonging, participation, and collaboration among its users, stakeholders, funders, and collaborators. A gateway should be designed and maintained with these objectives in mind.
9) Use cases and Success Stories
A gateway should provide examples of successful use

cases, scenarios, and stories that demonstrate the value and impact of the gateway for its users and stakeholders. For instance, a gateway could showcase how it helped researchers to solve a complex problem, publish a high-impact paper, or create a new curriculum. A gateway should also showcase examples of successful gateway frameworks that can serve as models or inspirations for other gateway developers and providers. Some of these frameworks are Tapis, Hubzero, and Clowder [10]. These frameworks provide reusable components and best practices for building and maintaining gateways that are scalable, secure, and user-friendly.

10) Documentation and Onboarding Resources
A gateway should not only provide access to resources, but also guide its users on how to use them effectively and efficiently. Therefore, a gateway should offer tips, guidelines, recommendations, and best practices for its users to get the most out of the gateway's features, functionalities, and services. These tips can include how to navigate the gateway, how to search and filter the available resources, how to customize the user interface and preferences, how to collaborate and communicate with other users and experts, how to troubleshoot and report issues, and how to provide feedback and suggestions for improvement.

A gateway should also support its developers and providers who are responsible for creating, maintaining, and enhancing the gateway. It should provide guidelines for its developers and providers to maintain and improve the quality, performance, and sustainability of the gateway. These guidelines can include how to design and develop user-friendly and secure features, how to test and debug the gateway functionality and performance, how to monitor and analyze the gateway usage and impact, how to update and upgrade the gateway software and hardware, how to document and disseminate the gateway information and achievements, and how to solicit and incorporate user feedback and suggestions for improvement.

IV. Conclusion

In conclusion, this paper documents ten factors common to successful science gateways as a way to help today's gateway developers and providers go beyond initial launch to successful ongoing use and adoption by a strong and growing community. By following these best practices, gateway developers and providers are more likely to create a successful gateway that enables access to advanced cyberinfrastructure resources and services for scientific research and education. Still, these best practices alone are not sufficient to guarantee success for any one gateway. Other important aspects come into play that are unique to the use case or the context, such as the right time for a solution in a given ecosystem, or the uptake by a crucial group in a given research community. What is presented here constitutes a set of practices that any gateway community can directly implement or plan for, and which should position that community for longer-term success and sustainability. We will continue the work on the ten practices by adding per practice the challenges and limitations. Furthermore, we plan to extend the practices by ten practices what to avoid.

Acknowledgment

We would like to acknowledge NSF OAC 1547611 (SGCI) and NSF OAC 2231406 (SGX3).

References

[1] Calyam, P., Wilkins-Diehr, N., Miller, M. Brookes, E., Arora, R., Chourasia, A., Jennewein, D., Nandigam, V., Lamar, M. D., Cleveland, S., Newman, G., Wang, S., Zaslavsky, I., Cianfrocco, M., Ellett, K., Tarboton, D., Jeffery, K., Zhao, Z., González-Aranda, J., Perri, M., Tucker, G., Candela, L., Kiss, T. and Gesing, S. Measuring Success for a Future Vision: Defining Impact in Science Gateways/Virtual Research Environments Concurrency and Computation, 2020;e6099. doi. https://doi.org/10.1002/cpe.6099

[2] Gesing, S., Zentner, M., Casavan, J., Hillery, B., Vorvoreanu, M., Heiland, R., Marru, S., Pierce, M., Mullinix, N., Maron, N. Science Gateways Incubator: Software Sustainability Meets Community Needs. 2017 IEEE 13th International Conference on e-Science (e-Science), Auckland, 2017, pp. 477-485. doi:10.1109/eScience.2017.77

[3] Gesing, S., Stirm, C., Klimeck, G., Zentner, L., Wang, S., Villegas Martinez, B.M., Diaz Eaton, C., Donovan, S., Zhao, L. , Song, C., Kim, I.L., Strachan, A., Zentner, M. and Kalyanam, R. Open Science via HUBzero: Exploring Five Science Gateways Supporting and Growing their Open Science Communities Proc. of HICSS-55 (55th Hawaii International Conference on System Sciences), Open Science Practices in Information Systems Research, January 2022, http://hdl.handle.net/10125/79420

[4] Joe Stubbs, Richard Cardone, Mike Packard, Anagha Jamthe, Smruti Padhy, Steve Terry, Julia Looney, Joseph Meiring, Steve Black, Maytal Dahan, Sean Cleveland, Gwen Jacobs. Tapis: An API Platform for Reproducible, Distributed Computational Research. In: Arai K. (eds) Advances in Information and Communication. FICC 2021. Advances in Intelligent Systems and Computing, vol 1363. Springer, Cham. https://doi.org/10.1007/978-3-030-73100-7_61

[5] The Galaxy Community. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update, Nucleic Acids Research, Volume 50, Issue W1, 5 July 2022, Pages W345–W351, doi:10.1093/nar/gkac247

[6] Wilkinson, M. D. et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018 (2016).

[7] Basney, Jim, and Von Welch. "Science gateway security recommendations." 2013 IEEE International Conference on Cluster Computing (CLUSTER). IEEE, 2013.

[8] Saltzer, Jerome H., and Michael D. Schroeder. "The protection of information in computer systems." Proceedings of the IEEE 63.9 (1975): 1278-1308.

[9] A. Olshansky, C. Hayes and K. F. Kee, The Characteristics of Influencers and Opinion Leaders of Science Gateways and Cyberinfrastructure for Innovation Diffusion, in Computing in Science Engineering, doi: 10.1109/MCSE.2023.3276905.

[10] Luigi Marini, Indira Gutierrez-Polo, Rob Kooper, Sandeep Puthanveetil Satheesan, Maxwell Burnette, Jong Lee, Todd Nicholson, Yan Zhao, and Kenton McHenry. 2018. Clowder: Open Source Data Management for Long Tail Data. In Proceedings of the Practice and Experience on Advanced Research Computing (PEARC '18). Association for Computing Machinery, New York, NY, USA, Article 40, 1–8. https://doi.org/10.1145/3219104.3219159