Received: 22 May 2020 | Revised: 25 October 2020 | Accepted: 2 November 2020

DOI: 10.1002/cpe.6192

RESEARCH ARTICLE

WILEY

If you build it, promote it, and they trust you, then they will come: Diffusion strategies for science gateways and cyberinfrastructure adoption to harness big data in the science, technology, engineering, and mathematics (STEM) community

Kerk F. Kee[1] | Bethanie Le[2] | Kulsawasd Jitkajornwanich[3]

[1]College of Media and Communication, Texas Tech University, Lubbock, Texas, USA

[2]Office of Strategic Marketing and Communications, Chapman University, Orange, California, USA

[3]Data Science and Computational Intelligence (DSCI) Laboratory, Department of Computer Science, Faculty of Science, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand

Correspondence
Kulsawasd Jitkajornwanich, Data Science and Computational Intelligence (DSCI) Laboratory, Department of Computer Science, Faculty of Science, King Mongkut's Institute of Technology Ladkrabang, Bangkok 10520, Thailand.
Email: kulsawasd.ji@kmitl.ac.th

Funding information
US National Science Foundation, Grant/Award Numbers: 1939067, 1453864

Abstract

In the big data era, for science gateways (SG) and cyberinfrastructure (CI) projects to have the greatest impacts, they need to be widely adopted in the scientific community. However, diffusion activities, or activities aimed to spread SG/CI in the science, technology, engineering, and mathematics community, are often an afterthought in projects. We warn against the fallacy of *"If You Build It, They Will Come."* Projects could be intentional in promoting tool adoption. Based on an analysis of 83 interviews with 66 administrators, developers, scientists/users, and outreach educators of SG/CI, we identified seven external communication practices—raising awareness, personalizing demonstrations, providing online and offline training, networking with the community, building relationships with trust, stimulating word-of-mouth persuasion, and keeping reliable documentation. With these strategies, we revised the pop culture line to *"If You Build It, Promote It, and They Trust You, Then They Will Come."* We also observed the beliefs that external communication is mainly necessary when seeking continuous funding, and it belongs to the skillset of nontechnical staff. These two beliefs may explain why external communication is underemphasized in many SG/CI projects. The article serves as evidence to justify a bigger budget in funding proposals for diffusion strategies to increase adoption and broader impacts.

KEYWORDS

cyberinfrastructure, diffusion of innovations, external communication, science gateways, technology acceptance, technology adoption

1 | INTRODUCTION

Big data requires a complex system to extract its hidden insights and values. Cyberinfrastructure (CI), also referred to as e-science or e-research in the European Union,[1] is a complex system that consists of large-scale datasets, data storage systems, remote instruments, advanced computing technologies, interdependent software, and so forth bundled together as an infrastructure to help dispersed teams of experts across diverse institutions to harness big data for research and education in STEM (science, technology, engineering, and mathematics) and other disciplines.[2,3] In other words, CI is a complex sociotechnical system that allows scientists to innovate and make discoveries with big data, in a way otherwise not possible.[4] However, the necessity for mastering advanced computer programming has limited the widespread user adoption of CI.

To address this challenge, science gateways (SG) were developed to reduce the need to master advanced programming skills to use CI. SG refers to a "community-specific set of tools, applications, and data collections that are integrated together via a portal or a suite of applications, providing access to grid-integrated resources".[5] Through a user-interface, SG allows scientists and educators to access shared software, data, computing (cloud) services, instruments, educational materials and other discipline-specific resources as a wide community of users.[6]

For SG/CI projects to have the greatest impact on scientific discoveries, the tools need to be widely adopted in the STEM community. However, the tools will not always spread by themselves. The developer teams behind the tools need to systematically and strategically engage in external communication to promote the tools. Otherwise, many tools may fall by the wayside, and eventually die out after the initial funding cycles conclude.

The lack of intentional promotion of SG/CI tools is often due to what we term the fallacy of *"If You Build It, They Will Come."* A line in pop culture inspired by the 1989 movie entitled "Field of Dreams," it creates the hopeful mindset that if a person strongly believes in a vision, the vision will eventually attract a large group of followers. In the movie, the main character (Ray) heard a voice that gave him the vision to build a baseball field in a cornfield in Iowa. Despites the initial struggles, the movie ends with hundreds of cars approaching the baseball field. While the message in the movie is inspiring, the reality of many SG/CI projects do not often match the same outcome as the movie.

The goal of this article is to explore the importance of external communication to the diffusion of SG/CI. By diffusion, we refer to the systemic spread of SG/CI as a collective innovation in the STEM community.[7] Given this goal, we set out to investigate what SG developer teams specifically and CI professionals broadly can do to help attract users to their tools. We organize the article in the following way. First, we give a brief overview on the diffusion of innovations theory (hereafter referred to as diffusion theory) and the technology acceptance model (TAM), and their applications to workplace technologies in the public marketplace broadly and the studies of SG/CI in the STEM community specifically. Second, we discuss the data collection and analysis in the methodology section. Third, we present the research findings, including interview excerpts that illustrate the practices we identified. Fourth, we conclude with limitations and implications.

2 | SOCIAL SCIENCE RESEARCH ON INNOVATION DIFFUSION AND TECHNOLOGY ACCEPTANCE

2.1 | Diffusion of innovations theory and TAM

The diffusion of innovations theory[8] is a communication theory that explains what triggers the adoption of innovations in a social system. It is a general theory developed to describe and predict the uptake of innovations. An innovation can be any technology, behavior, meme, ideology, information, and so forth, perceived as new to a group, organization, community, and society. The theory includes innovation attributes, user categories, and social networks as key predictors of adoption across various social, technical, cultural, and political contexts. The classic innovation attributes are relative advantage, perceived compatibility, perceived complexity, trialability, and observability.

A related theory is the TAM.[9] TAM presents a model that predicts the likelihood of users adopting a new system of technology. It proposes that the perceived ease of use and perceived usefulness of the technology are determinants of the user's attitudes and behavioral intent towards actual system use. Perceived ease of use is contextualized as the extent to which an individual believes that using a particular system would be effortless. Perceived usefulness is defined as the extent to which an individual believes that using a particular system will enhance one's job performance.

The TAM was originally developed to understand the adoption of information systems and workplace technologies. Davis proposed that the user's behavioral intention to use email and a data editing technology depends on its perceived usefulness and perceived ease of use.[9] TAM has been applied in a variety of emerging technologies such as, but not limited to, smartphone mobile applications,[10-12] wearable health devices,[13] social media,[14] and online learning.[15] For example, Melzner et al.[16] used the TAM for an integrated framework to assess the suitability of mobile health applications as a measure of workplace health promotion. In their conclusion, the researchers recommended companies to start an information campaign that incorporates the TAM variables in an effort to achieve the adoption of the mobile health applications. Lane and Stagg[17] also used the TAM in their research about the adoption of iPads among higher education staff. Their data show that university staff consider iPads to be easy to use, useful and compatible with their job tasks, and the researchers suggest for workplaces to consider user-centered technology support for staff members who found iPads less easy to use. Furthermore, the TAM was utilized in a study on green behavior in the workplace,[18] where the researchers found that education and training contribute to employees' adoption of green workplace behavior.

The introduction of new workplace technologies can help people, businesses and organizations perform their daily tasks more quickly and efficiently, optimizing productivity. Similarly, CI and SG are a job-related technology that can provide breakthroughs in the research workplace. However, active efforts by the technology organization will need to accompany the technologies to ensure successful implementation of the technology. We note that mobile apps and iPads in the cited studies were designed and created for consumer use and the general public, thus having a low learning curve and a design that is convenient and familiar for most users to adopt. CI and SG, on the other hand, usually have a high learning curve, and are designed for specialized and technologically sophisticated users. The case of SG/CI presents a unique opportunity to investigate the adoption, acceptance, and diffusion of complex innovations.

2.2 | Multidisciplinarity, and the sustainability and diffusion of SG and CI

Past research about social and organizational practices of SG and CI indicate that having a team of multidisciplinary experts as collaborators contribute to the project's chances of success. Multidisciplinary expertise is needed because the grand challenges and problems investigated in SG/CI projects are often not within a single discipline. However, the multidisciplinary nature of these projects can also be one of the obstacles in the adoption of SG/CI. Gesing et al.[19] showed that a challenge in achieving greater sustainability of SG through on-campus teams is the concern to maintain the diversity of expertise in the multidisciplinary group so that the project can benefit. Furthermore, according to Katz and Proctor,[20] the SG/CI infrastructure, itself, must address and support the collaborative, multidisciplinary team, which includes the response to larger and more complex data while promoting interoperability between systems and policies. Gesing et al.[21] expanded on this in their study about using SG for bridging the differences between research infrastructures. The researchers stated that SG offers sharing possibilities within a community but that is dependent on the willingness of researchers to share methodologies and data, and the technical dependencies of methodologies and data across disciplines.

Despite the challenges brought about by the necessary multidisciplinarity of the SG/CI community, some tools have been successfully adopted and sustained over the years. In a study on CI software sustainability, Stewart et al.[22] reported that projects successful in sustaining long-term operation usually employ open-source software licenses, and they also tend to utilize best practices known in the software engineering community. Furthermore, these projects tend to be led by leaders who are committed to the vision, and the projects hold regular meetings (e.g., user forums or conferences) where the developers and the users can engage with each other in person, at least on an annual basis. Finally, sustainable projects also tend to explore different financial models for funding at different points of the longitudinal project. This last point suggests that projects at different stages of development and maturity may benefit from different types and approaches of financial support.

Furthermore, Katz and Proctor[20] defined CI sustainability in terms of five key elements. First, dependent infrastructure maintains the same level of functionality over time, despite changes in other parts of the infrastructure. Second, collaborative infrastructure is combinatorial in nature with other infrastructural elements in order to meet the needs of users. Third, new users understand the CI functionality and usability, and are able to learn and ask relevant questions. Fourth, existing users are provided with infrastructural elements that are modular and can be adapted to meet their needs. Finally, the infrastructural elements continuously incorporate and implement new progresses in science and theoretical advances in research. While sustainability focuses on the macro processes in the CI/SG ecosystem, acceptance and adoption focuses on the micro perception of the technologies. We will discuss the perceptions of SG/CI attributes next.

At the level of innovation attributes to promote acceptance and adoption, Kee et al. described 10 attributes of SG/CI that attract new users.[23] They expanded on diffusion theory's classic five attributes and TAM's two key predictors to include attributes such as the SG/CI meeting users' needs, having organized access, trialability, observability, relative advantage/usefulness, simplicity/ease of use, compatibility, adaptability, and being community-driven, and well-documented. Furthermore, Kee and Schrock argued that social and organizational practices engaged by SG/CI developer teams increase the teams' capability to promote their tools' adoption and diffusion.[24] The 12 practices they identified are having multidisciplinary expertise, setting shared goals, using common language, having bridging liaisons, establishing productive routines, meeting face-to-face, demonstrating altruistic leadership, having clear roles, engaging user feedback, raising sustainable funding, growing organizational capacity, and maintaining personnel continuity.

The practices identified by Kee and Schrock[24] are inward-facing. In other words, they described what the teams can do internally among members of a project. In the present article, we set out to identify communication practices that are outward-facing. We aim to identify external strategies that teams can engage to diffuse SG/CI as an innovation to the scientific community as the target social system. Therefore, we pose the research questions (RQs), RQ1: *"What external communication strategies can promote the adoption and diffusion of science gateways and cyberinfrastructure?"* and RQ2: *"How is external communication perceived by the developer teams behind science gateways and cyberinfrastructure?"*At the heart of these RQs is the notion that building an innovation also requires strategic external communication and a communication savvy team to attract a large user-base. *"If You Build It, They Will Come"* is a fallacy.

3 | METHODOLOGY

The data for this article is a subset ($n = 83$) of a dataset collected for a larger study, which focuses on the social and organizational dimensions of big data projects during SG/CI diffusion. The larger dataset included 192 semistructured interviews between March 2016 and February 2019. Interviews were transcribed verbatim using a contract transcription service.

The interview protocol consisted of a variety of questions related to CI/SG definitions, adoption/implementation, influencers/opinion leaders, organizational capacity, and so forth. Key questions most germane to this article are: "When did you first hear about CI and/or SG? Were there particular people that influenced your initial involvement?," "Our research team is interested in discovering ways to accelerate CI and/or SG adoption and diffusion. What do you think your organization can do or is currently doing to trigger adoption sooner?," "If you were to create a checklist of the key factors for organizational capacity to accelerate adoption and diffusion of CI and/or SG, what would you include?." Furthermore, the interviews were semistructured, so the responses were based on other organic follow-up questions during the interviews, with the goal of understanding

adoption and diffusion as the ultimate outcome. It is worth noting that the nature of semistructure qualitative interviews is that informants have the flexibility to share their thoughts in response to the questions asked or focus on what they care about, what was salient in their minds, and what they would like to tell the interviewers. The variations enriched the study with details, improvisations, and informants' lived experiences in a way that cannot be captured by a structured survey questionnaire.

Informants in this study were recruited from the publicly available online programs of XSEDE/PEARC conferences in 2015, 2016, 2017, and 2018. Listed presenters and authors were first searched online via Google, and then an invitation email was sent to invite them for an in-person interview at the conference (if the recruitment timing was before the actual conference) or a telephone interview (if the conference had passed). XSEDE/PEARC conference programs were utilized for recruitment because listed presenters and authors are active contributors and stakeholders of the CI and SG communities. At the end of each interview, some referrals were made by the informants and pursued by the research team. We next describe the sampling process that selected the subset from the complete dataset for the present article.

3.1 | Systematic sampling

The larger dataset was mined for this article during three phases. During Phase 1 in January 2019, we sorted 123 transcripts that were available at that time into a roster based on chronological order of the interviews. Using the roster of the larger dataset, systematic sampling was performed with an interval of 6 (i.e., every 6th informant from the start of the roster) on the chronological list of informants of 123, yielding a subset of 20 randomly selected transcripts for analysis. A conference paper was written using the first subset of 20 interviews for the Gateways 2019 conference and presented in September 2019 in San Diego, California, USA.[25]

Then Phase 2 of the analysis took place during February to April 2020. We then had the complete set of 192 transcripts, and resorted them in alphabetical order of informants' last names. Once again, we performed another round of systematic sampling using the same interval of 6, and yielded 31 interviews. Based on close examination, five transcripts were sampled in both Phases 1 and 2, thus the analysis in Phase 2 consisted of 46 interviews. The first draft of this article was written based on these 46 interviews.

According to Baxter and Babbie,[26] one of the most rigorous sampling techniques in social science research is systematic sampling. Systematic sampling first sorts data into a particular order (e.g., alphabetical order of last names of individuals are clustered based on ethnicity given their last names; chronological order of interviews so currency of input is sequenced based on a timeline), and then select a random interval (i.e., an interval of 6 in this case) to select a subset for analysis. The interval of 6 was arbitrary, and an interval of 5, 8, or another number could have been selected. The goal of systematic sampling is to increase randomness (hence diversity and/or representativeness) in sample, by sorting of data into a set order, and then apply a consistent interval in selecting cases to be analyzed. This approach is to minimize the influence of biases due to demographic or temporal overlaps in informants or interviews.

A goal of systematic sampling is to increase representativeness of the data, thus broader transferability of the findings across projects and groups in the STEM community. Biases and limitations may stem from time periods (e.g., 2016 versus 2018) and social groups (e.g., cultural groups that share ethnic origins as signaled by last names). Given the nature of the special issue calling for submissions from participants in the Gateways and IWSG conferences, we recognized two opportunities to randomize the sample for a more representative analysis in this article. The two phases of analyses also served to confirm the findings, and expanded two particular themes (out of five) in the first Gateways conference paper into two additional related but distinct themes in the present article, which made the final framework with seven strategies more robust. Furthermore, the opportunity to engage in Phase 2 also allowed us to add RQ2 to the current article, as we expanded the scope beyond the first conference paper in the first submission to the special issue.

In September 2020, we performed Phase 3 of sampling in response to reviewers' comments. In order to broaden stakeholder perspectives by further increasing the diversity of professional roles, we performed a round of purposive sampling. Purposive sampling is a technique used when looking for data specific to a target population or a target segment of the population.[26] Purposive sampling yielded 32 additional transcripts for analysis. Results did not change the findings, but further confirmed the themes identified in earlier analyses. Furthermore, many of the informants interviewed played multiple roles across different projects. Moreover, even if they play the role of an administrator and/or outreach educator today, they may have been a technology developer or a coproducing user before. Therefore, their insights often reflect a mix of experiences with SG/CI. Given the three phases of sampling, the analysis reported in this article submitted to the journal special issue on SG is based on a total of 83 semistructured interviews with 66 informants (some informants participated in follow-up interviews).

While not all quoted in the article, informants whose transcripts are included in the analysis came from 22 US states (AL, AK, AZ, CA, CO, FL, HI, IA, IL, IN, MA, MN, MT, NE, OK, PA, SC, TX, UT, VA, WI, WY), Washington D.C., Canada, and Germany. The sample analyzed in this article include 26% female ($n = 17$) and 74% male ($n = 49$) who represented a diverse range of stakeholders, including center administrators, technology developers, scientists/users, and outreach educators of SG/CI. While the dataset is predominantly male (which is characteristic of the field), the variety of job roles and geographical locations gave us a holistic perspective for this article.

The notion of scientists/users requires some clarifications. In the sample, since scientists are the users of SG/CI for their scientific work, there are three types of users.[27] First, a scientist-developer is a scientist who developed his/her own SG/CI tool in order to do his/her own science, because

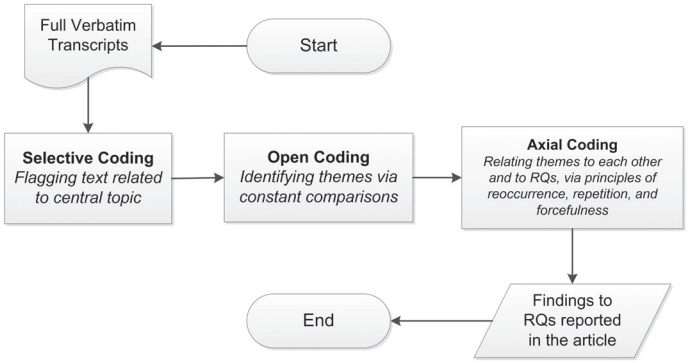

FIGURE 1 A visual diagram of the analysis process guided by grounded theory

there is no readily available tool out there that fits the research. Second, a coproducing user is a user who collaborated with a developer to design an SG/CI tool for his/her own research. Tools developed by the first two groups of users are then adopted by the third group of pure users, who simply adopt existing tools. Regardless of which group of users, some adaptations during implementation are common.

3.2 | Systematic analysis

We took a qualitative social science approach, in that data collection and analysis followed the principles and strategies of grounded theory.[28] The purpose of a grounded analysis is to investigate a phenomenon in its early stages, before a quantitative analysis is feasible. As a strategy for exploratory research, grounded theory analysis may appear anecdotal on the surface. We next explain how a grounded analysis was performed systematically in the present article.

The grounded theory approach includes the use of bottom-up coding methods to identify themes and concepts. We began by *selective coding*, which involved flagging a large portion of interview text related to communication practices that are outward-facing. Then the analysis continued with *open coding*, which organically allowed themes and concepts to emerge during careful interpretations. Through constant comparisons, concepts/themes were either further differentiated or combined, until distinct categories were identified. Finally, *axial coding* related the concepts/themes via specific relationships. We finalized the themes based on Owen's criteria of repetitions (i.e., the same practice being mentioned by multiple informants in the same way), reoccurrence (i.e., the same strategy being discussed by informants in different ways and from different perspectives), and forcefulness (i.e., a practice/strategy that is mentioned by few informants, but the idea can have significant and transformative impacts) to select themes.[29] In order to further illustrate the data analysis process, a diagram of grounded theory analysis is shown in Figure 1. Findings based on grounded theory analysis described in the following section emerged organically to consolidate and postclassify the responses into the appropriate categories. The interview questions were not designed based on the strategies apriori and deductively to retrofit the responses into preset categories.

4 | RESEARCH FINDINGS

We begin this section with reporting on the findings to RQ 1, "What external communication strategies can promote the adoption and diffusion of science gateways and cyberinfrastructure?" Based on a systematic analysis, seven strategies emerged. They are summarized in Table 1. These strategies were inductively derived from synthesizing informants' insights in an organic fashion, without imposing any existing framework. Furthermore, these strategies represent the collective insights of CI/SG stakeholders across different roles, as many of the informants were not strictly a user or a developer. Therefore, we list their multiple roles in the excerpts presented. For example, a scientist-developer[27] is a scientist (or a user) who also develops his or her own SG/CI tool in order to carry out his or her own research. A scientist-developer may indicate that she is also an administrator and/or an outreach educator across different projects. In that case, we list all three roles when introducing the representative excerpts. We elaborate on the seven strategies with informants' excerpts in the following section. Table 2 collects all the excerpts discussed for RQ1 in a fashion that allows readers to directly examine the raw data.

4.1 | External communication strategies for SG/CI diffusion

4.1.1 | Raising awareness with success stories of reputable peers

For SG/CI to become widely adopted, potential users first have to become aware of the tools. Raising awareness of the tools involves message dissemination, which is comprised of the public relations, marketing, and advertising efforts by project teams. A scientist-developer and outreach

TABLE 1 Seven diffusion strategies to promote SG/CI in the STEM community

Diffusion strategies	Definitions
Aimed at short-term and active promotion	
1. Raising awareness with success stories of reputable peers	Getting the words out, making the tool known to potential users through face-to-face meetings, newsletters, and/or social media, with success stories of reputable peers to induce confidence. Additional efforts can also support building a website, and being listed in an organized repository.
2. Personalizing demonstrations	Learning about potential users' work and work environment, in order to personalize demonstrations for them to see a strong fit with SG/CI and their own research.
3. Providing online and offline training	Making training available through organized materials, tutorials, short courses, classes, webinars, and bootcamp summer institutes. These can be offered either online and/or offline, in small groups and/or large workshops. Additional efforts can also provide user-support, address issues on community forums, and use chatbots to trouble-shoot.
4. Networking with the community	Promoting the tool at community activities (e.g., conferences), and getting to know what other tools exist, and what other peers are doing in the community. Community networking can trigger peer-to-peer support among domain scientists also.
Aimed at long-term and organic diffusion	
5. Building relationships with trust	Cultivating relationships with users, such that the users maintain adoption, and even share credits in joint publications, as the users believe that the developers help the users succeed and provide support during implementation.
6. Stimulating word-of-mouth persuasion	Generating word-of-mouth promotion and diffusion as users tell peers in their domain about the tools, because users come to trust the developers and the tools. Additional efforts can also identify current events (e.g., a crisis) that can utilize the tools, and giving diffusion an extra boost.
7. Keeping reliable documentation	Maintaining good documentation to help users actively learn, keep up with changes/bugs/updates, and follow good examples. Reliable documentation helps users develop the belief that the tool is sustainable, hence promoting adoption.

Abbreviations: CI, cyberinfrastructure; SG, science gateways; STEM, science, technology, engineering, and mathematics.

educator from Utah described this as a common problem, "The thing that still everybody struggles with is how to make everybody aware that we exist and we can help" (12 April 2016). An administrator and scientist-developer from Washington who specializes in computing and software systems concurred, "There is a sense of advertising that needs to happen in order for you to communicate to those who might use your software that you've created this and that it exists" (14 June 2016). Both informants spoke on how the aim of their outreach efforts is to let others know their tools' existence. Furthermore, the first informant suggested that raising awareness is a common challenge in the field. Engaging in outreach is key to raising awareness. An administrator and technology developer from California with a background in physics summarized this goal with the reflection question: "How do you reach out to your potential customers and make sure they know you exist?" (3 May 2016). An administrator and outreach educator from Oklahoma offered an answer to this question, "We do a lot of outreach to tell people that we're here … So it's … the communication, PR [public relations], and marketing part of [SG/CI]" (25 May 2016). Outreach efforts get the word out to have potential users recognize the SG/CI projects, and the tools that they offer.

How exactly can raising awareness be accomplished? Informants offered some in-person communication strategies. A technology developer from Indiana shared, "[W]e get a list of every incoming faculty every single semester, and then we go personally meet with every single one of them that will return our e-mail" (11 July 2017). Another informant who is an administrator and technology developer from Colorado added, "[W]e tried last year to really visit every department on campus and total faculty meeting and talk about what service we provide. It's still a large … group of people on campus that don't know about us" (21 March 2016). In these two quotes, informants talked about going from door-to-door in person as a strategy to help target potential users to become aware of SG/CI on their home campuses.

Furthermore, raising awareness can be accomplished with asynchronous communication strategies. An informant from Oklahoma talked about using newsletters. She explained that featuring stories in a newsletter is particularly effective to get researchers to think, "Oh, I do research similar to that. I didn't know I could use the HPC [high-performance computing] center for that" (25 May 2016). Using social media is another asynchronous strategy. An administrator and technology developer in Pennsylvania suggested that building a "social media presence, … [to send] announcements, updates, [and] news about what's going on [within the SG/CI project]" (24 October 2016) are steps to raise general awareness in message dissemination. A technology developer and administrator from Colorado involved in atmospheric research shared that his center's outreach also include the utilization of social media: "[W]e try to increase our social media activity, so that faculty or students become more aware of our service" (21 March 2016). A technology developer from Alabama offered this rationale, "So if your users really like a particular social platform [e.g., Facebook, Twitter, and so forth], then you may want to have a support handle on that social plat[form], [so] you can talk to your users where they like to talk" (23 March 2016). The strategy here is to meet the potential users where they are, and then bring them onboard from there.

TABLE 2 Raw excerpts illustrating the seven strategies

Diffusion strategies	Original excerpts
Aimed at short-term and active promotion	
1. Raising awareness with success stories of reputable peers	"The thing that still everybody struggles with is how to make everybody aware that we exist and we can help" (12 April 2016). – Scientist-Developer & Outreach Educator, UT
	"There is a sense of advertising that needs to happen in order for you to communicate to those who might use your software that you've created this and that it exists" (14 June 2016). – Scientist-Developer & Administrator, WA
	"How do you reach out to your potential customers and make sure they know you exist?" (3 May 2016). – Administrator & Technology Developer, CA
	"We do a lot of outreach to tell people that we're here … So it's … the communication, PR [public relations], and marketing part of [SG/CI]" (25 May 2016). – Administrator & Outreach Educator, OK
	"[W]e get a list of every incoming faculty every single semester, and then we go personally meet with every single one of them that will return our e-mail" (11 July 2017). – Technology Developer, IN
	"[W]e tried last year to really visit every department on campus and total faculty meeting and talk about what service we provide. It's still a large … group of people on campus that don't know about us" (21 March 2016). – Administrator & Technology Developer, CO
	"[W]e try to increase our social media activity, so that faculty or students become more aware of our service" (21 March 2016). – Administrator & Technology Developer, CO
	" … social media presence, … [to send] announcements, updates, [and] news about what's going on [within the SG/CI project]" (24 October 2016). – Administrator & Technology Developer, PA
	"So if your users really like a particular social platform [e.g., Facebook, Twitter, and so forth], then you may want to have a support handle on that social plat[form], [so] you can talk to your users where they like to talk" (23 March 2016). – Technology Developer, AL
	"You need success stories … to convince others to jump on … So that'll convince them that it's not just a fad if serious players are using it, serious people have used it and have had success with it" (23 March 2016). – Technology Developer, AL
	"[S]uccess stories make a big difference … [O]nce somebody in the department has a big publication, and it resulted in at least in part or maybe even in large part due to utilizing a supercomputer … [then] everybody wants in on that" (6 May 2016). – Scientist-Developer & Administrator, NE
	"[T]here are some people who have … built a lot of fame and credibility over producing and maintaining software artifacts that are such high quality that they get used, you know, a lot and people get very good reputations for that" (2 June 2017). – Technology Developer & Administrator, UT
2. Personalizing demonstrations	" … [M]y husband is faculty here. He says, 'I keep getting those emails about training in Hadoop. I don't know what Hadoop is. I don't know what it can do for me. You guys keep sending me these emails about training in Hadoop. I wish you would send me an email that tells me what it does so I know if it's good for me or not'" (6 April 2016). – User & Outreach Educator, SC
	"You're actually going out … [asking], 'What is your research?' They'll tell you, and then you say, 'Okay. So, cyberinfrastructure could improve things this way, this way, and that way. This is how you get access to cyberinfrastructure'" (24 July 2017). – Outreach Educator, WY
	"[G]iving them a kind of a tour of the tools and how they're used can spark their interest in using them or. . . thinking of ways those tools can apply to their work. So kind of like the Software Carpentry … It's called carpentry because it's like … taking somebody into your workshop and saying—Look, here's a vice … and here's the kinds of things you can do with a vice. And here's a hammer and here's a saw—and showing people that these are the kinds of tools that people can use to build things. And you do that the same way with cyberinfrastructure and it just kind of plants the seed" (25 May 2016). – Administrator & Outreach Educator, OK
	"The technology developer spent a day in the clinic and saw what it was like … So, we have shadowed them, not in the exam room, but just from the center and see what their technology looks like, see the tools that they have, see the decision they're trying to make. That was super enlightening for us. Because we could immediately see—Gosh, this is taking way longer than it should. Why is it taking so long?—It's for a variety of reason[s], but it helped us understand what the workflow was [like] … " (25 July 2018). – Administrator & Technology Developer, TX
	"[Y]ou might want to work with us … [for] a month to a year or longer. … [T]here's a significant amount of work that has to be done in many cases to get the domain scientist's problem solved properly with the cyberinfrastructure … " (2 June 2016). – Administrator, Coproducing User, Pure User, & Outreach Educator, IL
	"[H]istory builds trust" (16 November 2018). – Coproducing User, Scientist-Developer, Administrator, CA

(Continues)

TABLE 2 (Continued)

Diffusion strategies	Original excerpts
3. Providing online and offline training	"We provide a lot of training. So helping train students or faculty or researchers to use, and train in new technologies they are not used to" (21 March 2016). – Administrator & Technology Developer, CO
	Just like a university has freshman orientation, we offer similar classes and tutorials for beginning users to understand how to make effective use of the systems, how to get their work into the format and into the job submission service that the large systems use. Even just to familiarize themselves with the different systems and which system might be most appropriate for [their] science" (11 May 2016). – Administrator, TX
	"It's mainly through tutorials, presentations, conferences, or classes that we are doing online or at [a major HPC center in Texas] … " (18 May 2016). – Scientist-Developer, Technology Developer & Outreach Educator
	" … short courses, online materials, asynchronous materials, webinars, the whole host-bootcamp summer institutes" (25 July 2018). – Administrator, TX
	"Having one workshop and then never using it again is not very helpful. But having a lot of webinars are really helpful because that's really low investment on time, travel, and so forth. It's usually online and revisitable" (14 June 2017). – Coproducing User & Outreach Educator, IN
4. Networking with the community	"A lot of [knowledge comes] from the interactions and … relationships … [at] conferences where the new technologies are discussed and introduced, and having a network that goes beyond just our center [and] being involved in the national landscape of cyberinfrastructure" (6 April 2016). –Administrator & Outreach Educator, UT
	"[Attending conferences] really exposed me to the community, and I was able to meet people … That really made a big difference for me because it made me realize what people were doing in the field" (14 August 2017). – Outreach Educator, NC
	"That has a really positive impact on my own research … There was a scientist [from Pittsburgh] that was working with me and [he] could tell me, 'Okay, this is a good way to go. This is not a good way to go.' Even though I'm an expert in this field myself … [B]ecause he worked on … four projects similar to mine, he could offer some insight to it" (3 July 2017). – Coproducing User & Scientist-Developer, TX
Aimed at long-term and organic diffusion	
5. Building relationships with trust	"Being able to make the relationships with [potential users] to … coax them into doing something new and interesting … that can maybe help them improve their work" (11 July 2017). – Technology Developer, IN
	"It is about working hand in hand and coming alongside to create excellence. That builds trust because they know you're there for them—to help them out. And that goes a long way … " "And then also making sure that you position yourself so that you can execute. Sometimes you can have the trust of a person, but you can lose their trust [during] your execution" (July 2018). – Technology Developer, UT
	"[C]ollaboration involves quite a lot of trusting in the people you're working with that, when you come to publishing, you're both going to be appropriately listed as authors for reward and credit for your contribution" (7 April 2017). – Coproducing User & Scientist-Developer, UT
	"[A]t the end of the day, I think that those types of things are largely a kind of a trust between PI and like some of the lead developers … There's a discussion about why you think this is a great thing to do. And then if it's kind of a little bit informal … Yep, sounds great, let's do it" (20 November 2018). – Coproducing User & Scientist-Developer, VA
	"[I]dentifying the key individuals in lab that are the ones that are the early adopters … Because not only is it not contextual, but there is a trust factor … [T]hat's paramount. … So if you do not have a trust factor, if you don't have skin in the game, and I say—Oh, yeah, you should go try this, do that—And then it fails, or it doesn't work … That's the part where you start losing users, and they feel that you're not vested in their growth … " (29 April 2016). – Scientist-Developer, Technology Developer, & Administrator, AZ
	"A lot of times … whoever leads the research leads from what they know best … And either the faculty has to trust … the postdoc or any incoming grad students can do it in some new method or methodology … [or] they just stay doing the way they've done it before … That has nothing to do with what tools and what capacity of computation or storage that you're providing for somebody" (5 November 2018). - Administrator, MA
	"I would try to create a situation where people really trusted us to help them and took our advice … I think we could build a much nicer synergy between the people who run the cyberinfrastructure and the people who use it" (14 July 2016). – Pure User & Administrator, CA
	"[F]or a lot of researcher, what really matters is will the transfer happen and not how fast it will be. It should not take 500 years, but if they want to do something, they want it to be dependable, right? And they want to trust that if something goes wrong, somebody will be there to help them" (26 April 2016). – Pure User, Technology Developer, & Administrator, WI

(Continues)

TABLE 2 (Continued)

Diffusion strategies	Original excerpts
6. Stimulating word-of-mouth persuasion	"[W]e try to build the relationship with the faculty long-term … we hope [through word-of]-mouth the faculty can tell their colleague[s] how the CI can help them [i.e., the colleagues] in terms of research … We have some faculty … they heard [that] their colleague[s] … can do something [with SG/CI], and they start to get interested" (13 March 2017). – Pure User, Outreach Educator, & Administrator, IN
	"[S]imilar to like how your friends in the dorm give you recommendations about classes, there [are] also user communities that develop, either within particular research groups or even across particular disciplines that give folks ideas on how to use system[s] most effectively, kind of tricks of the trade in terms of configuring a simulation to a particular system" (11 May 2016). – Administrator, TX
	"[A]s far as advertisements—a lot of these scientific applications are really domain specific, so there isn't as much advertising as much as colleagues providing the information to other colleagues they know would be utilizing it" (5 November 2018). – Technology Developer, MA
	"Faculty and researchers don't have a lot of time to just experiment with whatever the latest/greatest thing is. So, having somebody in the community that's local to where they are, somebody they've built up a trust around, that they can go to and say - Is it worth my time to invest in maybe experimenting with this?" (5 November 2018). – Technology Developer & Outreach Educator, CA.
	"[B]ring the domain scientists closer to the computational resource with tools that are trusted because they are developed by peers in the field, who are the innovators and who are funded and are motivated to develop those things … [Y]ou need the domain scientist that speaks the domain science stuff among their peers" (12 July 2017). – Technology Developer, PA
7. Keeping reliable documentation	"We do have a wiki online, but … it's not very user friendly … [I]t's more sort of a kind of improvised way of doing some internal documentation" (19 January 2019). – Pure User & Scientist-Developer, AZ
	" … write better documentation for end users to get some training" (12 July 2017). – Administrator & Technology Developer, Canada
	"[O]nce we work on certain projects, we write tutorials and wikis … So documentation and wiki is important" (17 July 2017). – Pure User, IA
	" … getting real world use cases well-documented and disseminated would be a powerful sales pitch" (3 May 2016). – Administrator & Technology Developer, CA
	"[P]oor documentation is a barrier to entry … [D]ocumentation is the unsung hero of most cyberinfrastructure and also the sort of unrecognized weakness that a lot of cyberinfrastructure has … [H]aving complete documentation that [is] written hopefully with the input of someone who's a user and not just the maintainer [and] the developer of the package, is super important … " (11 July 2017) – Coproducing User, CA

Whether it is a newsletter or a social media post, the informants suggested using success stories of reputable peers. A technology developer in Alabama explained, "You need success stories … to convince others to jump on … So that'll convince them that it's not just a fad if serious players are using it, serious people have used it and have had success with it" (23 March 2016). This strategy of using success stories is echoed by a scientist-developer in computational chemistry who is also an administrator from Nebraska. He stated, "[S]uccess stories make a big difference … [O]nce somebody in the department has a big publication … resulted in at least in part or maybe even in large part due to utilizing a supercomputer … [then] everybody wants in on that" (6 May 2016). These informants talked about success stories of SG/CI users in their domain research. However, another aspect of success stories speaks to the success of the tools and/or the developers. A technology developer and an administrator from Utah suggested, "[T]here are some people who have … built a lot of fame and credibility over producing and maintaining software artifacts that are such high quality that they get used, you know, a lot and people get very good reputations for that" (2 June 2017). Collectively, the informants shared that success stories, especially stories of high-profile users, showcase to potential adopters that the SG/CI tool works and it has been vetted by credible users. It also helps when the developers have a strong reputation themselves.

4.1.2 | Personalizing demonstrations

Personalizing demonstrations is to have potential adopters and beginning users see a strong fit between SG/CI and their own research work. While this strategy makes a lot of sense on surface, it is not always employed in practice. Sometimes the developer team promotes a technology without considering the need to personalize messages for the target audience. A user and facilitator in South Carolina recounted, "[M]y husband is faculty here. He says, 'I keep getting those emails about training in Hadoop. I don't know what Hadoop is. I don't know what it can do for me"' (6 April 2016). She described how developer teams tend to send a blanket and general message about a technology, which can leave the target users puzzled. She then continued relaying her husband's complaint with, "You guys keep sending me these emails about training in Hadoop. I wish you would send me an email that tells me what it does so I know if it's good for me or not."

The missing link suggested by the first excerpt is personalization. In order to put personalization into practice, an outreach educator from Wyoming told us that outreach work should not simply be sending out mass messages and waiting to answer questions. Instead, "you're actually going out … [asking], 'What is your research?' They'll tell you, and then you say, 'Okay. So, cyberinfrastructure could improve things this way, this way, and that way. This is how you get access to cyberinfrastructure"' (24 July 2017). This informant believed that not every potential user can immediately recognize SG/CI's usefulness. Therefore, it is important to find out what potential users are doing and then personalize the tools to users' specific situations to help them recognize the advantages of implementing SG/CI in their own research.

Similarly, an administrator and facilitator from Oklahoma with a mathematics background emphasized personal demonstrations, as she explained,

> [G]iving them a kind of a tour of the tools and how they're used can spark their interest in using them or … thinking of ways those tools can apply to their work. So kind of like the Software Carpentry … It's called carpentry because it's like … taking somebody into your workshop and saying—Look, here's a vice … and here's the kinds of things you can do with a vice. And here's a hammer and here's a saw—and showing people that these are the kinds of tools that people can use to build things. And you do that the same way with cyberinfrastructure and it just kind of plants the seed. (25 May 2016)

Both excerpts underscore the notion that promoting adoption is a two-way street, where the developer team also find out users' work, and help them explore how SG/CI can fit into their work, in addition to one-way public relations and marketing efforts.

Personalization requires not simply understanding what the research users do, but also the workflow and physical environment of users. An administrator and technology developer from Texas talked about shadowing as a technique to achieve this understanding, and provided the following example in the medical context,

> The technology developer spent a day in the clinic and saw what it was like … So, we have shadowed them, not in the exam room, but just from the center and see what their technology looks like, see the tools that they have, see the decision they're trying to make. That was super enlightening for us. Because we could immediately see—Gosh, this is taking way longer than it should. Why is it taking so long?—It's for a variety of reason[s], but it helped us understand what the workflow was [like] … (25 July 2018)

The excerpt above describes how shadowing can help developers better understand how users work, therefore, able to design SG/CI that better fit their daily work environment. No matter how good an SG/CI tool is, if it does not fit the users' work environment, diffusion will be impeded.

Furthermore, personalization involves spending time working together on making the adoption work. A administrator, coproducing users, pure user, and outreach educator in Illinois told us that it is important to facilitate deep engagement between the developer and the user—"[Y]ou might

want to work with us … [for] a month to a year or longer." He further explained, "[T]here's a significant amount of work that has to be done in many cases to get the domain scientist's problem solved properly with the cyberinfrastructure … " (2 June 2016). Spending time is necessary not only to ensure that the technology works for solving the problem, it also builds trust. A coproducing user, scientist-developer, and administrator in California stated, "[H]istory builds trust" (16 November 2018). With trust as a foundational condition, once a user is interested in using SG/CI, the next step is to invite them to participate in training.

4.1.3 | Providing online and offline training

Given the high learning curve of SG/CI, training is essential for turning short-term curiosity into adoption, and/or ensuring long-term continuation and successful implementation. Otherwise, a potential user or a new user may stop exploring at this point. An administrator and technology developer from Colorado explained, "We provide a lot of training. So helping train students or faculty or researchers to use, and train in new technologies they are not used to" (21 March 2016). An administrator from Texas used a metaphor to illustrate his organization's approach:

> Just like a university has freshman orientation, we offer similar classes and tutorials for beginning users to understand how to make effective use of the systems, how to get their work into the format and into the job submission service that the large systems use. Even just to familiarize themselves with the different systems and which system might be most appropriate for [their] science. (11 May 2016)

The freshman orientation metaphor refers to educational efforts as a vehicle for potential users to imagine and understand where their work, jobs, and needs fit within SG/CI, with deeper technical understanding of how the tools work. Training takes place in various formats. We examined the interviews to identify the common formats. For example, a scientist-developer, technology developer, and outreach educator shared, "It's mainly through tutorials, presentations, conferences, or classes that we are doing online or at [a major HPC center in Texas] … " (18 May 2016). Another informant from Texas who identifies as an administrator and technology developer added, "short courses, online materials, asynchronous materials, webinars, the whole host-bootcamp summer institutes" (25 July 2018). Collectively, we synthesized the formats into organized materials, tutorials, short courses, classes, webinars, and bootcamp summer institutes, and these can be offered either online and/or offline, in small groups and/or large workshops.

Given that training is key, it is important to make training accessible and convenient for users throughout the entire adoption and implementation process. It is necessary for users to revisit training materials over time. A coproducing user and outreach educator from Indiana shared her insight by stating, "Having one workshop and then never using it again is not very helpful. But having a lot of webinars are really helpful because that's really low investment on time, travel, and so forth. It's usually online and revisitable" (14 June 2017). This excerpt emphasizes the needs to make training available online and available for users on an ongoing basis. Furthermore, while most of the training activities take place at a university campus where there is a research computing center, informants also mentioned conferences as the venues where presentations and training can occur. The next theme examines further the networking aspect of conferences as part of external communication.

4.1.4 | Networking with the community

Networking with the STEM community is key to becoming effective at promoting SG/CI to new users. Informants talked about the need for developers to attend conferences to promote SG/CI through giving presentations and tutorials about their tools. This argument is clearly articulated by the interview excerpts and discussion in the previous theme. We will not reiterate them here. Instead, we would like to emphasize the observation that developers shared with us about how networking at conferences in the community can further help developers become more effective promoters of their own tool when they are familiar with what other SG/CI tools are available in the community.

An administrator and outreach educator from Utah told us that her organization is aware of how much they can benefit by attending conferences, stating, "A lot of [knowledge comes] from the interactions and … relationships … [at] conferences where the new technologies are discussed and introduced, and having a network that goes beyond just our center [and] being involved in the national landscape of cyberinfrastructure" (6 April 2016). Moreover, an outreach educator with a background in parallel computing from North Carolina added, "[Attending conferences] really exposed me to the community, and I was able to meet people … That really made a big difference for me because it made me realize what people were doing in the field" (14 August 2017). By networking with the community at conferences, developers not only give presentations and tutorials to promote their own tools, they also become aware of what other peers are doing and the latest knowledge related to building SG/CI tools.

Similarly, users reported that networking with their domain communities help them adopt and implement SG/CI. A coproducing user and scientist-developer from Texas shared her experience meeting a peer scientist at a conference, who later helped her with implementing SG/CI in

her research—"There was a scientist [from Pittsburgh] that was working with me and [he] could tell me, 'Okay, this is a good way to go. This is not a good way to go"'(3 July 2017). She further explained, "That has a really positive impact on my own research … Even though I'm an expert in this field myself … [B]ecause he worked on … four projects similar to mine, he could offer some insight to [SG/CI]." Networking with the community helps both the developers and users with SG/CI diffusion.

4.1.5 | Building relationships with trust

A key part of successful SG/CI projects includes relationship building and establishing trust between developer teams and potential users. To demonstrate why relationships with users matters, a technology developer from Indiana told us, "Being able to make the relationships with [potential users] to … coax them into doing something new and interesting … that can maybe help them improve their work" (11 July 2017). As previously discussed, potential users may not immediately recognize how SG/CI can help them with their work, and/or they may need the training to begin using the tool meaningfully. Having a relationship with a potential user can help them get over the hump and adopt. This informant also added that every potential user is unique, and that building a relationship requires starting from scratch each time. Furthermore, diffusing a tool requires the potential users to have confidence that the developer teams are committed to the new users' success in using the tool. A technology developer from Utah noted, "It is about working hand in hand and coming alongside to create excellence. That builds trust because they know you're there for them—to help them out. And that goes a long way … " (July 2018). This excerpt shows how promoting SG/CI is not simply about the point of adoption, but also working with users during implementation until the work is done. To the point of implementation, this informant from Utah continued, "And then also making sure that you position yourself so that you can execute. Sometimes you can have the trust of a person, but you can lose their trust [during] your execution." His excerpt also shows that trust building is a continuous effort. One oversight can undo the good work previously established.

The strategy of building relationship with trust is also repeated by users. For example, a coproducing user and scientist developer from Virginia stated, "[A]t the end of the day, I think that those types of things are largely a kind of a trust between PI and … the lead developers" (20 November 2018). He further explained, "There's a discussion about why you think this is a great thing to do. And then if it's kind of a little bit informal … [Then the user agrees]—"Yep, sounds great, let's do it." His excerpt shows that having trust between the PI and the lead developers is often what informally convinces users to adopt, not simply having a discussion about why SG/CI adoption is a good idea.

Another trust-related strategy is to help someone on the PI's team to be able to adopt SG/CI for the PI. A scientist-developer, technology developer, and administrator from Arizona shared, "[I]dentifying the key individuals in lab that are the ones that are the early adopters … Because not only is it not contextual, but there is a trust factor … [T]hat's paramount" (29 April 2016). He further explained why trust is important in SG/CI adoption, "So if you do not have a trust factor, if you don't have skin in the game, and I say—Oh, yeah, you should go try this, do that—And then it fails, or it doesn't work … That's the part where you start losing users, and they feel that you're not vested in their growth …." Furthermore, an administrator from Massachusetts explained that it is not that users are resistant to SG/CI, but—"A lot of times … whoever leads the research leads from what they know best … they just stay doing the way they've done it before" (5 November 2018). Then he continued by stating, "[E]ither the faculty [as SG/CI user] has to trust … the postdoc or any incoming grad students can do it in some new method or methodology." This excerpt shows that users need to have someone they could trust to carry out SG/CI implementation in their research. If it is not the SG/CI developer team, it needs to be someone they trust within their research team. He concluded by stating, "That has nothing to do with what tools and what capacity of computation or storage that you're providing for somebody."

The notion of trust from the informants suggests that trust means someone is there to support successful completion of the research when new SG/CI is being used, not that SG/CI is a better tool for their research. A pure user, technology developer, and an administrator from Wisconsin shared that for researchers, "[W]hat really matters is will the transfer happen and not how fast it will be. It should not take 500 years, but if they want to do something, they want it to be dependable" (26 April 2016). This excerpt suggests that relative advantage of SG/CI in terms of speed may not be always critical. What is critical is that the support is always there. He continued, "And they want to trust that if something goes wrong, somebody will be there to help them." With the same insight, a pure user who later became an administrator from California shared, "I would try to create a situation where people really trusted us to help them and took our advice … I think we could build a much nicer synergy between the people who run the cyberinfrastructure and the people who use it" (14 July 2016).

Building relationships with trust is a long-term commitment beyond simply adopting the tool. A scientist-developer and coproducing user from Utah told us, "[C]ollaboration involves quite a lot of trusting in the people you're working with that, when you come to publishing, you're both going to be appropriately listed as authors for reward and credit for your contribution" (7 April 2017). He suggested an interesting observation that we believe has not been documented in previous studies—along with users' success in implementing the technology within his/her workplace, the 'marketer' also shares in being recognized for the success. Although getting published is not formally a part of many developers' performance metrics, the credits certainly speak to the depth of the kind of relationships perhaps required in SG/CI adoption. The informant concluded by stating the importance of working with people that are interested in the same problems, approaching them in a similar way, and having respect and trust so that collaborators can have confidence in the results generated by the adopted SG/CI tool.

4.1.6 | Stimulating word-of-mouth persuasion

In addition to building trusting relationships with SG/CI users to reenforce adoption, another reason to build trust is to harness the relationships with users for word-of-mouth persuasion. Recommendations by peers in existing relationships encourage diffusion. A pure user, outreach educator, and administrator from Indiana explained, "[W]e try to build the relationship with the faculty long-term … we hope [through word-of]-mouth the faculty can tell their colleague[s] how the CI can help them [i.e., the colleagues] in terms of research … " (13 March 2017). This strategy again emphasizes the need to build long-term relationships with users. The strategy in fact has observed to be working, as the informant continued, "We have some faculty … they heard [that] their colleague[s] … can do something [with SG/CI], and they start to get interested."

The excerpt above reminded us of a study of hybrid seed corn in Iowa in the 1940s, where Ryan and Gross[30] discovered that sales agents provided the information, but it was the interpersonal conversations with neighboring farmers that promoted the adoption of the agricultural innovation. An administrator from Texas with a background in visualization technologies provided an analogy to describe word-of-mouth persuasion in SG/CI,

One can argue that class recommendations work especially well among peers that have the same major. The role of having a shared domain adds to this interpersonal diffusion. A technology developer from Massachusetts stated, "[A]s far as advertisements—a lot of these scientific applications are really domain specific, so there isn't as much advertising as much as colleagues providing the information to other colleagues they know would be utilizing it" (5 November 2018).

An important aspect of persuasion and influence by peers is that potential users can avoid wasting time in adoption SG/CI if it is not a wise investment for them. The people who can best assess if SG/CI is a wise investment are the peers of the potential users. A technology developer and an outreach educator from California shared, "Faculty and researchers don't have a lot of time to just experiment with whatever the latest or greatest thing is" (5 November 2018). Instead, he continued explaining that what works is "having somebody in the community that's local to where they are, somebody they've built up a trust around, that they can go to and say - Is it worth my time to invest in maybe experimenting with this?." Furthermore, a technology developer from Pennsylvania added the notion that if the SG/CI are produced or coproduced by their domain peers, it can be even more persuasive. He suggested, "[B]ring the domain scientists closer to the computational resource with tools that are trusted because they are developed by peers in the field, who are the innovators and who are funded and are motivated to develop those things" (12 July 2017). He continued, "[Y]ou need the domain scientist that speaks the domain science stuff among their peers." Collectively, either the goal is to build long-term relationships with users directly, or rip from users' organic conversations with their near peers about SG/CI to further spread the innovation, focusing on building trust in relationships is critical.

4.1.7 | Keeping reliable documentation

The final strategy explores the value and benefits of record keeping and documentation. A reliable documentation includes consistent and dependable record keeping of various versions, iterations, and updates over time, with source codes, user guides, instructional materials, and good examples, so potential users can learn the tools and imagine how they could use the tools in their own research. Reliable documentation signals to users that the developer teams are keeping up changes and addressing bugs, so users increase their trust and confidence in the tool that it will continue to be maintained for the long-term.

Documentation is a professional practice. It is often not incomplete in projects by developers who build a tool for their own science. A scientist-developer and a pure user from Arizona talked about his own tool, "We do have a wiki online, but … it's not very user friendly … [I]t's more sort of a kind of improvised way of doing some internal documentation" (19 January 2019). However, poor documentation can impede the adoption of a tool beyond one's own project.

Documentation is important for training, a strategy previously described. An administrator and technology developer from Canada suggested, "write better documentation for end users to get some training" (12 July 2017). An informant from Iowa described documentation as an ongoing practice, "[O]nce we work on certain projects, we write tutorials and wikis … So documentation and wiki is important" (17 July 2017). Good documentation also includes good examples, as an administrator and technology developer from California shared, "getting real world use cases well-documented and disseminated would be a powerful sales pitch" (3 May 2016).

Good documentation signals to users that the developer team is responsive to users' input, especially when there are bugs in the tool. Incorporating users' feedback will help improve the tool in its next iteration, thus increasing adoption. A coproducing user in California shared that lack of documentation can be an obstacle for potential users to adopt SG and CI:

> [P]oor documentation is a barrier to entry … [D]ocumentation is the unsung hero of most cyberinfrastructure and also the sort of unrecognized weakness that a lot of cyberinfrastructure has … [H]aving complete documentation that [is] written hopefully with the input of someone who's a user and not just the maintainer [and] the developer of the package, is super important … (11 July 2017)

As a summary, we identified seven strategies for external communication for SG/CI adoption and diffusion. Upon deeper reflection, we note a distinction between the first four strategies and the last three that we discussed. The first four strategies (i.e., raising awareness, personalizing demonstrations, providing training, and networking with the community) are practices aimed at actively promoting SG/CI, and these activities tend to be short-term oriented. The next three strategies (i.e., building relationships, stimulating word-of-mouth persuasion, and keeping reliable documentation) are aimed at organically diffusing SG/CI, and these activities tend to be long-term oriented. Thus, we further organized the strategies into two categories of aiming at short-term promotion and long-term diffusion, as labeled in Table 1. Given these strategies for RQ1, next we continue by investigating developers' perceptions of external communication.

4.2 | Developers' perceptions of external communication

In this section, we explore the findings to RQ2, *"How is external communication perceived by the developer teams behind science gateways and cyberinfrastructure?"* Two findings emerged: the mindset that external communication is mainly for attracting new funding and the belief that external communication is a skillset that belongs to nondevelopers.

4.2.1 | External communication is mainly important for getting continuous funding

External communication is mainly important when the developer team is looking for continuous funding. Specifically, an outreach educator from Texas shared, "If we're just talking about completing a project that's already been bought and paid for, you could live without [external communication]" (26 November 2018). And she continued, "[I]f you want to continue to have funding … there's no way you can live without all of those [communication] outputs … All of those outputs are hugely important for being able to do the next project." Another informant, a technology developer from Massachusetts concurred, "It's not necessarily the PR until it comes to funding. As soon as you're starting to do your next iteration of funding, then suddenly mass media and public relations becomes a little more important … " (5 November 2018). He also explained that external communication is important "in order to promote it to show it as a wider range and useful in other dynamic." These excerpts suggest that SG/CI developers tend to believe that external communication is separate from the technology itself, and communication is mainly for getting continuous funding for the technology in its next iteration.

The distinction between technology development and external communication is further distinguished as the two funding criteria of the National Science Foundation (NSF). A technology developer from Kansas explained, "So, I would say that external communication, particularly when you are starting the project—I would say it's useful, but on the other hand, it's the sort of thing that it can frequently be ignored for moderately extended period … " (24 January 2019). Interestingly, the informant suggested that external communication is not required or necessary throughout the entire duration of the project. In fact, it is something that can be overlooked, at least for a period of time. He continued, "[I]t's … like broader impacts versus the intellectual merit. If you don't have the intellectual merit, it doesn't matter how good your broader impacts are, you're going to get toasted in your idea … [E]xternal communication is the broader impact[s]."

NSF has two funding proposal review criteria—intellectual merit and broader impacts. Intellectual merit refers to the scientific outputs of the proposed research, and broader impacts refer to how the scientific outputs can benefit the larger community and society at-large. The informant's excerpt suggests that external communication fulfills the second criterion of how much a SG/CI tool can positively impact the STEM community overall. Here, the informant deemphasizes external communication as ancillary.

4.2.2 | External communication is a skillset for nondevelopers

External communication is a different type of skillset than technology development. Informants first distinguished the two skillsets of technology development versus external communication. An administrator and technology developer from Wyoming shared, "[E]ducation and evangelization of technology is a different skillset than designing and operating technology … [W]hen you deploy [a] new technology, you staff up by people that can design and implement" (21 April 2016). Not only did he distinguish the two skillsets as technology development (e.g., designing and operating technology) and external communication (e.g., education and evangelization of technology), he also implied that the skillsets belong to two distinct groups of professionals. He continued, "[A]nd you tend to bring the evangelizers and the educators along last … " Not only are the two skillsets distinct and belonging to two different groups of professionals, the sequence is for external communication to follow technology development.

While external communication and technology development are believed to belong to two different groups of professionals, some believe that communication is a teachable skillset. An outreach educator with a doctorate in engineering from Michigan shared, "We've been trying to train CI professionals how to communicate … They've been teaching them in business for years. We're just moving them into the CI …

> It's not just solving a technical problem with cyberinfrastructure, it's the engagement, the person constructive engagement between cyberinfrastructure engineers and team members and the researchers. And that engagement piece—that personal engagement piece—has to be part of the solution. You need to hire people who are very technical to do these jobs and the technology is their bread and butter … [T]hat's the hammer that they have. But, to be really effective, you need certainly to have that hammer, but you need to have this sort of mindset of—We're going to personally, constructively engage with researchers and kind of do the legwork to build relationships and relationships of trust. (12 July 2017)

It's been shown over and over again that it's a teachable skill" (24 July 2018). Although this quote suggests that communication is a teachable skill, the informant is passionate about it because he recognized that majority of the technology developers do not naturally see this skillset as fitting their professional role in the field. The punchline of the article can be summarized by the block quote by an administrator and an outreach educator from California below:

The premise of the article is that for SG/CI projects to have the greatest impacts, the tools need to be widely adopted in the scientific community. However, diffusion activities are often an afterthought in SG/CI projects. While we foreground the importance of external communication, SG/CI tools are fundamentally technical. We simply argue not to leave external communication to promote adoption as an afterthought. Therefore, we revised the pop culture line from "Field of Dreams" to *"If You Build It, Promote It, and They Trust You, Then They Will Come."*

5 | CONCLUSION, LIMITATIONS, AND IMPLICATIONS

We identified seven external communication strategies to promote SG/CI diffusion. While we present the themes as distinct, they are intertwined in practice. For example, we noted that documentation on wiki is for long-term references, but it is also related to providing materials online for short-term training. Furthermore, word-of-mouth persuasion can cooccur with networking with the community at conferences. We suggest carrying out all seven strategies sequentially and concurrently in a combinatorial fashion, adapting to community responses at any given time. Of course, qualitative social science cannot claim causality. This is a limitation of the present article. The strategies identified came from the informants' responses, suggestions, and reflections, which helped us formulate on what it may take for developer teams to promote SG/CI diffusion through external communication. These practices cannot guarantee success because an unfavorable outcome may occur due to reasons beyond anyone's control. However, in future work, we plan to quantitatively (in)validate and/or expand the findings by involving more participants in several platforms, which is possible by utilizing big data cloud services in conducting a survey/interviews as well as collecting and analyzing the data. We envision using NLP and other big data processing frameworks to help us perform such analyses.

Another limitation is recruiting from XSEDE/PEARC conferences for the study. Because of this focus, informants were primarily administrators, outreach educators, and technology developers. Users are not the biggest attendees and participants of XSEDE conferences in recent years. Furthermore, those who responded and agreed to be interviewed about SG/CI adoption and diffusion were individuals who likely already valued external communication is SG/CI projects. Another limitation of recruiting informants from these conferences was that projects represented in this study were active and funded projects, given their participation at the conferences. Therefore, the sample does not specifically include projects that did not have continuous funding at the time of the interviews. Given this, one can perhaps argue that the strategies suggested (or discussed to be responsible for successes) by informants in active and funded projects at the time of the interviews likely associated with their active and funded status. However, such an association is best confirmed by a follow-up quantitative study (such as using a survey questionnaire) to statistically test the hypothesized relationship. Furthermore, future research could actively seek out informants with experiences involved in failed projects, although recruitment of such informants willing to openly and honestly share about their failed projects may be difficult.

We balanced being descriptive (summarizing what informants told us they are doing and why) and prescriptive (identifying best practices that we can document and pass on to optimize external communication in future projects). In the process of ensuring that the key strategies are based on what informants described in their work during the interviews, we noticed that no one talked about having a formal website for SG/CI tools. Perhaps it is a given that there will be a website, or perhaps tools are best disseminated in an organized repository such as GitHub.

Furthermore, we maintain that user support is key to sustaining adoption. This may also be implied by the discussions of training and building relationships with trust, especially during the implementation phase. However, we can see the use of community forums,[23] chatbots[31] to aid in user support. In addition, as we write this article during the COVID-19 crisis, we recognize that tools designed for (or that can be repurposed to address aspects of) a major crisis will also receive a sudden surge in awareness and publicity, because of the crisis. We sprinkled these observations as additional efforts summarized in Table 1.

Along the line of trust, we note here that trust is especially important during relationship building for the long-term, our fourth strategy. However, a closer look at the excerpts summarized in Table 2 reveals that the notion of trust was also mentioned in "Raising Awareness with Success Stories of Reputable Peers" and "Stimulating Word-of-Mouth." Although the key word of "trust" was not mentioned in excerpts illustrating "Personalizing Demonstrations," "Providing Online and Offline Training," "Networking with the Community," and "Keeping Reliable Documentation," one can argue that trust is perhaps necessary and implied in all seven strategies. Furthermore, the notion of "reputation" is also found to be symbiotic

with trust, especially in the first strategy of "Raising Awareness with Success Stories of Reputable Peers." Reputation enhances trust, and vice versa. Although the key word of "reputation" was not specifically mentioned in excerpts related to the other six strategies, one could also argue that a negative reputation would probably undermine all of them.

Although not always the case, analysis did suggest that developers may become coauthors in joint publications with users in their research. This finding struck us as interesting for several reasons. First, it means the external communication efforts for SG/CI are different from traditional PR, advertising, and marketing work in the general public marketplace, where interactions between a "marketer" and a user conclude after the purchase transaction. Second, it suggests that external communication is necessary to be carried out by developers, as they have the technical knowledge to help users implement SG/CI to generate findings worthy of publications. Interestingly, findings suggest that what drives developers to engage in external communication is mainly to obtain the next round of funding, while many also believe that communication is not the skillsets of developers. We note this as potentially the explanation for why some SG/CI tools lack external communication, and the efforts are often an afterthought. If communication only represents episodic and occasional efforts, the developer team may risk not being able to sustain continuous funding. Furthermore, we note that external communication only makes sense when there is a solid SG/CI tool to promote and diffuse. Otherwise, promoting a weak tool is both unlikely to be successful as well as ethically questionable.

While there is a belief that communication is a teachable and learnable skillset for developers, we suspect that it will be difficult to prioritize it in practice. The low prioritization of external communication can be further jeopardized by the observation that many informants play multiple roles in the SG/CI community. In presenting their interview excerpts and quotes, we intentionally chose to list all the professional roles they indicated during the interviews. Many SG/CI professionals have to juggle between being a specialist in one field, and a master of another task. In fact, many talked about playing more than two roles during the interviews. Therefore, we can see why external communication is often carried out as a secondary or tertiary responsibility. While it is admirable, the juggling of multiple roles may compromise the quality of external communication efforts. Overlooking diffusion strategies is a common pitfall because developers are usually not trained in external communication. Their job should focus on the technical dimensions, rightfully so. Moreover, carrying out effective external communication is extremely time-consuming. Developers are often already overwhelmed by their technical responsibilities. The ideal situation would be for there to be a full-time communication professional on every SG/CI team. In this ideal case, a team would be putting in practice Kee and Schrock's[24] principles of seeking multidisciplinary expertise (to bring on communication professionals), having bridging liaisons (between users and developers), setting shared goals (between developers and the communication professionals to seriously promote SG/CI). However, it is often difficult to build into the budget of project to support full-time personnel to engage in external communication. Given this, we warn against the fallacy of "*If You Build It, They Will Come,*" lest that good SG/CI fails to gain traction due to practices that can be proactively engaged to promote uptake. We write the present article as another piece of evidence for projects to reference in funding proposals to request a bigger budget for external communication and strategic diffusion to increase broader impacts.

Given the longer-term and systemic nature of the challenges identified, we present some short-term and low barrier solutions for readers to consider, in the event that SG/CI developer teams would like to increase external communication and address the fallacy discussed in this article. First, developer teams can work with the 'strategic marketing and communications' (SMC) office (or an equivalent of such an office, whose name may vary across universities), and get a blogpost out about the project, so to increase the awareness of SG/CI within and across campuses. Then the developer team can further share and tweet out the blogs for further diffusion through their social media networks. Along the same line, SMC offices are usually responsible for their universities' social media presence (in addition to the blogpost). A developer team can request additional social media PR and advertising efforts to get the word out. Second, the SMC office can also sit down with SG/CI developer teams and help them explore strategies to access local resources developers are unfamiliar with or unaware of. For example, the university's vice president for research's office may put out monthly/quarterly research newsletters or magazines. This additional venue, in addition to other types of local resources, may be helpful.

Third, developer teams could consider consulting with books such as "Marketing for Scientists" by Kuchner,[32] a staff scientist at NASA's Goddard Space Flight Center, who learned about marketing from his successful country music songwriting career. Fourth, developer teams can identify multiple tool repositories and community forums, such as GitHub, StackOverFlow, and other venues where there is an active dialogue, and introduce their tools appropriately. Fifth, domain specific tools, such as the CyNeuro Science Gateway at the University of Missouri-Columbia, has seen successes also with partnering with neuroscience faculty in introducing CyNeuro in classes, thus influencing adoption of the tool by the next generation of scientist users early in their career. Finally, developer teams can invite the involvement of graduate students (or undergraduate students, faculty) at their home campus from departments such as communication studies (for interpersonal communication), mass communication (for science communication), public relations (for campaign strategies), advertising (for marketing strategies), and so forth, in exchange for a research opportunity for thesis projects, class projects, and so forth (although logistics of this suggestion may require some efforts).

To conclude this article, we propose an intervention at the funder's level. Given the funding assumption for external communication, the NSF could consider accepting a "communication management plan" either as an optional supplement or potentially a requirement in the long-term during funding proposal submissions. Such a plan, similar to the current practice of requiring a "data management plan," would encourage principal investigators (PIs) and their collaborators to be more thoughtful about external communication at the proposal stage. This could encourage PIs to consider budgeting for a communication professional to join the team or to budget appropriate time for technical personnel to have protected time to

prioritize external communication throughout the entire project. This may increase a PI's chance of continuous funding by design, while increasing the adoption, acceptance, and diffusion of the SG/CI tool the developer team produce.

ACKNOWLEDGMENTS

This article is based on work funded by the US National Science Foundation (NSF), under NSF CAREER Award OCA #1453864. We thank Mona Sleiman, Lydia Benjamin, and Jamie McCain for their assistance in data collection, and Faith Escalera for her contribution in the early conference version of this article.

ORCID

Kerk F. Kee https://orcid.org/0000-0002-0543-5009
Kulsawasd Jitkajornwanich https://orcid.org/0000-0002-6926-7577

REFERENCES

1. Schroeder R. e-Research infrastructures and open science: towards a new system of knowledge production? *Prometheus*. 2007;25(1):1-17. https://doi.org/10.1080/08109020601172860.
2. Atkins D. *Revolutionizing Science and Engineering Through Cyberinfrastructure: Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure. Technical Report*. Arlington, VA: National Science Foundation; 2003.
3. Stewart CA, Simms S, Plale B, Link M, Hancock DY, Fox GC. What is cyberinfrastructure. Proceedings of the 38th Annual ACM SIGUCCS Fall Conference: Navigation and Discovery; 2010:37-44.
4. Kee KF, Cradduck L, Blodgett BM, Olwan R. Cyberinfrastructure inside out: definitions and influences shaping its emergence, development, and implementation; 2011.
5. Wilkins-Diehr N. Science gateways-common community interfaces to grid resources. *Concurr Comput Pract Exper*. 2007;19(6):743-749.
6. Science Gateways Community Institute; 2007. https://sciencegateways.org/.
7. Kee KF. Adoption and diffusion. *The International Encyclopedia of Organizational Communication*. Atlanta, GA: American Cancer Society; 2017;1-14. https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118955567.wbieoc058.
8. Rogers E. *Diffusion of Innovations*. 5th ed. New York, NY: Free Press; 2003.
9. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quart*. 1989;13:319-340.
10. Min S, So KKF, Jeong M. Consumer adoption of the uber mobile application: insights from diffusion of innovation theory and technology acceptance model. *J Travel Tourism Market*. 2019;36(7):770-783.
11. Ooi KB, Tan GWH. Mobile technology acceptance model: an investigation using mobile users to explore smartphone credit card. *Expert Syst Appl*. 2016;59:33-46.
12. Yoon HY. User acceptance of mobile library applications in academic libraries: an application of the technology acceptance model. *J Acad Librariansh*. 2016;42(6):687-693.
13. Choi B, Hwang S, Lee S. What drives construction workers' acceptance of wearable technologies in the workplace? indoor localization and wearable health devices for occupational safety and health. *Automat Construct*. 2017;84:31-41.
14. Makki TW, DeCook JR, Kadylak T, Lee OJ. The social value of snapchat: an exploration of affiliation motivation, the technology acceptance model, and relational maintenance in Snapchat use. *Int J Hum–Comput Interact*. 2018;34(5):410-420.
15. Robinson T. Using the technology acceptance model to examine technology acceptance of online learning technologies by non-traditional students. *i-Manager's J Educ Technol*. 2019;16(1):21.
16. Melzner J, Heinze J, Fritsch T. Mobile health applications in workplace health promotion: an integrated conceptual adoption framework. *Proc Technol*. 2014;16:1374-1382.
17. Lane MS, Stagg A. University staff adoption of iPads: an empirical study using an extended TAM model. *Australas J Inf Syst*. 2014;18(3):53–74.
18. Zhang Y, Luo Y, Zhang X, Zhao J. How green human resource management can promote green employee behavior in China: a technology acceptance model perspective. *Sustainability*. 2019;11(19):5408.
19. Gesing S, Lawrence K, Dahan M, Pierce ME, Wilkins-Diehr N, Zentner M. Science gateways: sustainability via on-campus teams. *Future Generat Comput Syst*. 2019;94:97-102.
20. Katz D, Proctor D. A framework for discussing e-research infrastructure sustainability. *J Open Res Softw*. 2014;2(1):e13
21. Gesing S, Krüger J, Grunzke R, Herres-Pawlis S, Hoffmann A. Using science gateways for bridging the differences between research infrastructures. *J Grid Comput*. 2016;14(4):545-557.
22. Stewart CA, Barnett WK, Wernert EA, Wernert JA, Welch V, Knepper R. Sustained software for cyberinfrastructure: analyses of successful efforts with a focus on NSF-funded software. Paper presented at: Proceedings of the 1st Workshop on the Science of Cyberinfrastructure: Research, Experience, Applications and Models, Portland, OR; 2015:63-72.
23. Kee KF, Sleiman M, Williams M, Stewart D. The 10 attributes that drive adoption and diffusion of computational tools in e-science. Paper presented at: Proceedings of the XSEDE16 Conference on Diversity, Big Data, and Science at Scale, Miami, FL; 2016:1-8.
24. Kee KF, Schrock AR. Best social and organizational practices of successful science gateways and cyberinfrastructure projects. *Future Generat Comput Syst*. 2019;94:795-801.
25. Le B, Escalera F, Jitkajornwanich K, Kee KF. External communication to diffuse science gateways and cyberinfrastructure as innovations for research with big data. Paper presented at: Proceedings of the Gateways Conference; 2019; San Diego, CA.
26. Baxter LA, Babbie ER. *The Basics of Communication Research*. Belmont, CA: Wadsworth/Thomson Learning; 2003.
27. Kee KF, Browning LD. Challenges of scientist-developers and adopters of existing cyberinfrastructure tools for data-intensive collaboration, computational simulation, and interdisciplinary projects in early e-Science in the US. Paper presented at: Proceedings of the Data-Intensive Collaboration in Science and Engineering Workshop, Computer Supported Cooperative Work (CSCW 12); 2012; Seattle, WA.
28. Corbin JM, Strauss A. Grounded theory research: procedures, canons, and evaluative criteria. *Qual Sociol*. 1990;13(1):3-21.

29. Owen WF. Interpretive themes in relational communication. *Quart J Speech*. 1984;70(3):274-287.
30. Ryan B, Gross NC. The diffusion of hybrid seed corn in two Iowa communities. *Rural Sociol*. 1943;8(1):15.
31. Chandrashekara AA, Talluri RKM, Sivarathri SS, et al. Fuzzy-based conversational recommender for data-intensive science gateway applications. Paper presented at: 2018 IEEE International Conference on Big Data (Big Data). Seattle, WA: IEEE; 2018:4870-4875.
32. Kuchner MJ. *Marketing for Scientists: How to Shine in Tough Times*. Washington, DC: Island Press; 2012.

How to cite this article: Kee KF, Le B, Jitkajornwanich K. If you build it, promote it, and they trust you, then they will come: Diffusion strategies for science gateways and cyberinfrastructure adoption to harness big data in the science, technology, engineering, and mathematics (STEM) community. *Concurrency Computat Pract Exper*. 2021;e6192. https://doi.org/10.1002/cpe.6192