Creating Sustainable Cyberinfrastructures

David P. Randall
The Information School
University of Washington
Seattle, WA, USA
dpr47@uw.edu

E. Ilana Diamant
Human Centered Design & Engineering
University of Washington
Seattle, WA, USA
ilanadm@gmail.com

Charlotte P. Lee
Human Centered Design & Engineering
University of Washington
Seattle, WA, USA
cplee@uw.edu

ABSTRACT
In this paper we report the results of a qualitative research study of the GENI cyberinfrastructure: a program of four federated cyberinfrastructures. Drawing on theories of stakeholder positioning, we examine how different GENI stakeholders attempt to enlist new participants in the cyberinfrastructures of GENI, and leverage existing relationships to create sustainable infrastructure. This study contributes to our understanding of how cyberinfrastructures emerge over time through processes of stakeholder alignment, enrollment, and through synergies among stakeholder groups. We explore these issues to better understand how cyberinfrastructures can be designed to sustain over time.

Author Keywords
Actor-Network Theory; Enrollment; Synergizing; Cyberinfrastructures

ACM Classification Keywords
H.5.3. Group and Organization Interfaces.

INTRODUCTION
The use of large-scale distributed scientific collaborations, powered by advanced computational and networking technologies has increased rapidly in recent years [21]. These endeavors, known largely as c*yberinfrastructures*, have wide potential to dramatically alter the way science is performed. Through utilizing such organizations, scientists can examine very large datasets from a wide variety of sources across geographic and interdisciplinary boundaries with relative ease, greatly increasing opportunities for effective scientific collaboration [36]. Despite this potential, and much prior research into the issues surrounding the design of such infrastructures in both the CHI and CSCW fields [5,6,12,28,30,46], creating effective and enduring cyberinfrastructure remains a complex challenge.

http://dx.doi.org/10.1145/2702123.2702216

Despite the challenges involved in designing robust and sustainable cyberinfrastructures, and understanding how they can be maintained over time, many funding agencies and governments continue to enthusiastically support and fund such infrastructure projects. However, the lack of understanding results in cyberinfrastructures breaking down, as the tensions between the various stakeholders lead to collapse. When cyberinfrastructures fail, science is impeded and so it is critical that we attain a fuller understanding of how we can better design cyberinfrastructures that will endure.

Cyberinfrastructures are huge, complex and multifaceted organizations, often embedded in and across many differing social, organizational and technological entities [5]. As a result of this complexity "nobody is really in charge of infrastructure" [38] meaning that their maintenance and longevity is the result of a great deal of negotiation between the various stakeholders. Many early cyberinfrastructure development projects were funded with little consideration about how the resources would be maintained at their conclusion [31]. Recent research has turned to investigating how sustaining cyberinfrastructure entails "continually realigning the relationships among people, technologies, and organizations." [6] However, despite this research, we still know far too little about how people, groups, and larger organizations work to continually reposition their relationships with themselves and with relevant technologies in order to sustain these virtual organizations. Through understanding these relationships, and how stakeholders reposition themselves, we can better identify how cyberinfrastructures can be designed to accommodate such repositioning from their inception.

In this qualitative study we use a grounded theory approach to draw on the frameworks of *Actor-Network Theory* and *Synergizing* in order to better understand stakeholder positioning practices in the development of a program of four interrelated cyberinfrastructures, known as Global Environment for Network Innovations (GENI). These positioning practices are undertaken to cultivate the relationships necessary to create and maintain a functional cyberinfrastructure. We examine how the different GENI stakeholders attempt to enlist new users to the four cyberinfrastructures that comprise GENI, and also how they use leveraging to realign already existing relationships. We focus explicitly on stakeholder positioning for two main reasons. Firstly because it presents a critical challenge for the sustainability and long-term maintenance of cyberinfrastructures and for the creation of scientific artifacts (data, software, algorithms, and research products) that cyberinfrastructures are designed to support. And secondly because there is still relatively little understanding of how the concept of synergizing works ‘on the ground’, especially in the context of cyberinfrastructure design. We then relate the implications of these results to the design on future cyberinfrastructures.

To this end, we focus on the following research question:

RQ: How do existing stakeholders use positioning practices to develop and sustain cyberinfrastructures?

LITERATURE REVIEW

Cyberinfrastructures represent a new front in both science and technology. They are comprised of cutting-edge information technologies with the explicit goal of supporting large-scale, collaborative, inter-disciplinary research.

The term cyberinfrastructure (CI) was first used by the National Science Foundation (NSF) in 2003 [1], in reference to so-called ‘collaboratories’, which utilized emerging high speed networks in order to support distributed science [39,14,17,28]. The NSF, along with many other US government agencies, has placed a high priority on CIs as part of their strategy for developing tools to promote scientific research and discovery, in both the natural and social sciences. Even though collaboratories and CIs were initially considered technological artifacts, more recent research has conceptualized them as virtual organizations that include both technical and social aspects [11]. The social and relational aspects comprise the “human infrastructure” [21] of cyberinfrastructure which is the complex aggregation of social forms and interactions.

Research on technology adoption [10,41], innovation [32], user participation [2], and information systems development [9,24] has tended to conceptualize users as a predictable group of entities, operating within the physical or cultural boundaries of an organization or a community with a common value system (e.g., open-source community). CI systems differ in that their users and stakeholders are not always homogenous, specified or known ahead of time [21,37]. However, research has shown CIs to be highly emergent—changing and evolving over time[37]

Star et al. [37] have shown that infrastructures are fundamentally relational, in the sense of being embedded in multiple organizations, communities of practices and also in multiple technologies. They describe embeddedness, as the way an infrastructure “is ‘sunk’ into, or inside of, other structures, social arrangements, and technologies.” More specifically, when we describe an infrastructure as embedded we mean that it is “situated within a network, the web, or other arrangement of relationships to other systems” [5]. The infrastructure’s relationships thus both enable it to provide valuable services, but also constrain the processes which supply those services. In addition an infrastructure may be embedded in multiple structures simultaneously, such as technological, social, or organizational networks. That relational aspect creates social and technical challenges, not only for the long-term maintenance and sustainability of CIs, but also for their design [6].

While many large scale systems have traditionally been the focus of a purely technical design strategy [35] [30], the unique intricacies of CIs require their design to be approached from a more socio-technical perspective. Temporal aspects also need to be considered, as such large-scale collaborative endeavors may be impeded by failure to plan for the alignment of the multiple temporalities involved in such infrastructure projects, and so it is critical to coordinate the discordant rhythms and events emanating from such organizations that could negatively affect sustainability [39]. In focusing cyberinfrastructure design around the interactions and relationships between both the technological components and the human infrastructure [21] involved, designers are in a position to better secure longevity and sustainability in these systems.

Studies of IT development outside the field of cyberinfrastructures have shown that user involvement is a critical element for the success of IT-development projects and for the long-term adoption and use of IT systems. However, the specific mechanisms through which user involvement can best benefit IT projects is unclear [2,3,25,29]. Similarly, studies of large-scale IT systems have discussed the need to identify actions of “user advocacy” whereby stakeholders of a system’s user groups promote the new system to other stakeholders and potential new users [4,23].

Below we discuss in more detail the stakeholder positioning theories in the field of cyberinfrastructure development we will be using in this study.

Theories of Stakeholder Positioning in Cyberinfrastructure Development

The design and development of large cyberinfrastructures involves complex decisions and tasks that require the input of experts with different backgrounds such as engineering, information technology and systems design. Past research has noted the challenge and need for positioning of the divergent goals, values and cultures of CI’s stakeholders [41, 31, 35, 36, 28]. More specifically, stakeholders need to be positioned in ways that promote both effective development and long-term use and sustainability of the CI resources. However, achieving such a positioning of diverse stakeholders across organizational and disciplinary boundaries is a challenge. Funding and sponsoring organizations (e.g., NSF) play a role by promoting workshops and other activities that bring together key actors, and by working closely with project leaders throughout the development and maintenance of the cyberinfrastructure.

Positioning diverse stakeholders is not only a challenge for the systems-development effort but also presents challenges for the CI's sustainability and long-term maintenance. Making a CI sustainable means making it last for long (or even indefinite) time periods, beyond the scope of a single event, site or research project. However, as Ribes & Finholt [30] note, the sustainability of CIs is characterized by tensions such as the "development vs. maintenance" tension which arises from the necessity of doing ongoing maintenance work even though developing new tools and software tends to be more rewarded. More recent research by Bietz et al. [6] has discussed the sustainability of CIs in terms of two specific goals: creating sustainable institutional resources and the preservation of scientific products. Creating sustainable institutional resources includes challenges of funding, long-term planning and resource allocation [15,48]. Whereas creating sustainable scientific products includes challenges of designing reusable technologies, "big data" models and artifacts that cross disciplines, and scientific software that can be used by newcomers and future scientists beyond the immediate research questions they were designed to address [6,22,30,33,40]. Below we discuss two theories that can help us understand how infrastructural sustainability and stakeholder positioning were achieved in the case of the GENI infrastructures.

Actor-Network Theory

Actor-network theory (ANT) originated in the 1980s through the work of the science and technology studies theorists Michel Callon [7] and Bruno Latour [18,19], as well as sociologist John Law [20]. Originally constructed as a concept for explaining how theories, techniques and ideas in science become accepted and adopted, ANT has since been refined and applied by many disciplines. In information systems it is often utilized as a sociotechnical perspective through which to view technology adoption [42]. Understanding how interests are positioned in a heterogeneous network is key to understanding how CIs are developed and sustained.

Actor-networks may comprise any number of human and non-human actants [19], these can be anything that has an impact upon what is being investigated. This can include people, organizations, tools and artifacts. Actants are recruited to actor-networks through a process of translation, whereby the values and interests of the outside actants are brought into line with those of actants inside the network; in order to create the actor-network [34]. Callon [7] lays out this process into four distinct stages known as: *problematization, interessement, enrollment,* and *mobilization.*

Through these translation actions and activities, actants join or form networks. The process starts with identifying the actant, engaging those actants by supporting them with their interests/perspectives, and making them part of the network by negotiating roles for them. At the completion of the process enrolled actors can be thought of as allied actors.

In that sense, translation can be seen in the formation of networks and their dynamics. Callon discusses the *enrollment* stage specifically as "the device by which a set of interrelated roles is defined and attributed to actants who accept them. To describe enrollment is thus to describe the group of multilateral negotiations, trials of strength, and tricks that […] enable them to succeed." [7]. Enrollment involves negotiation of roles and relationships. Thus enrollment is the process of translating the interests of a large body of allies onto an actant so as to bring into line their interests with those of the other actants. This is done in order to create a heterogeneous network of aligned interests who are "willing to participate in particular ways of thinking and acting which maintain the network." [42]

Synergizing

Observing the need to investigate how organizational sustainability and change is accomplished, Bietz et al [5] introduced the concept of *Synergizing.* Synergizing recognizes that infrastructures, rather than being stand-alone systems, are embedded into pre-existing sociotechnical structures, and other social arrangements and technologies [5]. Synergizing is not just a point in time, but rather, it denotes a process of creating interactions that generate a greater combined effect by successfully utilizing an expansive array of technological, human and organizational entities. As a theoretical concept, synergizing allows us to consider the relational structures of a cyberinfrastructure and how they affect the cyberinfrastructure's sustainability over time [6]. It includes two sub-processes:

Aligning – the work done by developers to bring entities together in such a way as to create a relationship that will "produce, and function within, the nascent cyberinfrastructure."[5].

Leveraging – the process of utilizing an existent relationship with an organization, actor or artifact so as to reinforce that relationship, or to create a productive relationship with some other organization, actor or artifact.

In this paper we report the results of a qualitative research study of the GENI cyberinfrastructure program, which is an innovative program that funded not just one, but four interrelated cyberinfrastructure systems for computer networking science. We draw on the concepts of *Actor-Network Theory* [18,19] and *Synergizing* [5] to aid in our understanding of how GENI stakeholders encourage others to use their technologies, how they enroll new stakeholders, and how they leverage the existing ties between stakeholders to reposition relationships and create stronger and more stable connections within the infrastructure; especially in the context of cyberinfrastructure design. These two theories are closely related - synergizing's concept of alignment is similar to ANTs translation process at a high level – and the actions they explain were evident from the grounded theory analysis of the data.

RESEARCH SITE

The Global Environment for Network Innovations (GENI) is a program of four interrelated cyberinfrastructure systems ("technology clusters") that exist within a federation. GENI is funded by the National Science Foundation (NSF) and administered by an independent organization known as the GENI Project Office (GPO), who provide project management expertise, administer grants to project teams, organize regular conferences for GENI stakeholders and provide oversight and support for workgroups within GENI [47].

Each infrastructure within GENI is developed by a different team of computer scientists and houses different groups or communities of experimenters. GENI was created as "a virtual laboratory" for "exploring future internets at scale". That involves experimenting with alternative network architectures to the Internet's current TCP/IP design.

Experimenters and developers may choose any of the four infrastructures, and research projects are assigned to (or join) one of those four technology "clusters". GENI was selected for this study because it is comprised of multiple experimental cyberinfrastructures (PlanetLab, ProtoGeni, ORCA and ORBIT), with a common management and funding agency. In that sense, it is structured as a federated socio-technical organization of cyberinfrastructure systems.

METHODS

We conducted a series of semi-structured interviews with 49 GENI stakeholders including researchers, lead scientists and/or engineers (known as Principal Investigators or PIs), designers, experimenters, GENI managers, instructors and students using GENI for research or coursework. Table 1 shows a breakdown of interviewees by role (note that some interviewees held more than one role). In order to conduct these interviews the researchers attended a number of GENI Engineering Conferences (GECs) during a period of 18 months starting in March 2013. The GECs occur 4 times a year and are used as a forum for participants involved in the various GENI projects to share details of their work. The interviews lasted around one hour on average, the audio was recorded and then transcribed. Participants were selected for the study from across the four GENI infrastructures using a purposive sampling technique based upon their role within GENI (ranging from masters students to senior researchers, developers, and managers). The advantage of purposive sampling is that it allows researchers to focus on particular characteristics of the sample that might not be typical of the population [27,44].

Interviews were supplemented with participant observations of demos, presentations and group meetings that took place at the GECs. That allowed for triangulation of the data, [13,16]. Interview data was coded using Atlas.ti software and analyzed using a grounded theory approach [8], we then drew on this data to identify actor-network theory and synergizing as frameworks from which to view the activities of the GENI stakeholders. The names of all participants and their projects discussed in this paper are pseudonyms.

Role	**Interviewees**
Experimenter	11
GPO Manager	4
Developer	12
PI (experimenters & developers)	9
NSF	2
Student	15

Table 1. Breakdown of interviewees

FINDINGS

Our findings revolve around three aspects of stakeholder participation in the GENI infrastructures. First we identify the major stakeholder groups and their role in the four GENI infrastructures. Then we describe the methods various stakeholders use to enroll other individuals in the GENI infrastructures or projects that they are a member of. Finally we describe how stakeholders leverage existing relationships to increase stakeholder participation and solidify the ties between them.

Major GENI Stakeholder Groups

There are a large number of stakeholders involved in the development and administration of GENI, these can be broken down into a number of sub-groups, with some individuals having roles that span one or more of these groups.

Funding Agency & GENI Managers

Stakeholders in this group administer the GENI infrastructures. The funding agency stakeholders are the NSF directors and employees that are overseeing the GENI projects. Those individuals have responsibilities for funding, allocating resources, and high-level design decisions. So far there have been two program directors at NSF working on GENI since its inception in 2005. This stakeholder group also includes the GENI managers, who are members of the GPO. The GPO and its managers are responsible for planning, coordinating and making design and architectural decisions for the GENI infrastructures.

Developers

This group consists of the GENI developers and architects, who are computer scientists/engineers responsible for building the four infrastructures. Developers are primarily affiliated with one of the four infrastructures, but they may be involved technology projects which span across multiple infrastructures. Also included here are users and developers of the PlanetLab GENI infrastructure, which are *outside GENI*: these are individuals who are not currently attending or participating in the GENI conferences but who have contributed to the development of Planetlab in the past.

Experimenters

Stakeholders in this group are mostly using technologies developed in a single infrastructure. They include computer scientists/engineers whose work is primarily focused on an experimental research study of GENI, funded by the NSF. The work of these experimenters also involves development and design of software, protocols, hardware and other technologies that are part of GENI.

Students and Instructors

Yet another stakeholder group consists of students who contribute to the development of GENI technologies or are developing a technology based on a particular infrastructure as part of their research/dissertation project. This group also includes Computer Science instructors/educators in universities and colleges who are primarily interested in developing/implementing/using a specific GENI technology in the classroom in undergraduate or graduate courses focusing on networking science (e.g., with the goal of demonstrating principles of networking science).

Enrollment and Aligning Practices

Our analysis reveals the various GENI stakeholders engage in five different positioning practices: collaborating with domain scientists, conducting outreach programs for experimenters, improving communications between experimenters and developers, technology integration, and tool creation. These practices demonstrate how stakeholders are actively involved in the enrollment practices of actor-network theory, as well as the alignment and leveraging practices of synergizing, as they reposition existing relationships and form new ones in processes which help to maintain the cyberinfrastructures.

Collaboration with Domain Scientists

Developers from GENI's ORCA infrastructure, using the ExoGENI test bed, conduct experiments in collaboration with a number of domain scientists from bioinformatics, earth science and astronomy. Through these experiments the developers build new relationships with domain scientists, adding new experimenters to GENI, who in turn help to expand the GENI infrastructure and its related technologies, bolstering its capabilities and further expanding its potential user-base.

Through these collaborations on experiments, developers move to enroll scientists from disciplines outside of computer science. One of the ORCA/ExoGENI PIs described these experiments as benchmarking tests of computational applications that are being developed specifically for the domain-science users. They also discussed the involvement of domain experts in another project in which developers have teamed up with the scientists from Information Sciences Institute-West (ISI) who brought in both other domain scientists as well as their own software for managing complex computational workflows. Existing collaborators sometimes bring in new experts:

"[The ISI scientists] *brought in their software that manages complex computational workflows, but they also brought in some domain scientists in astronomy and earth science"* (Issac, ORCA/ExoGENI)

In such cases, new users might be enrolled into an infrastructure initially as testers, and start collaborating with the development teams, and through that collaboration they can adapt/adjust the infrastructure to the specific needs of the domain scientists.

Much like the ISI developers brought in domain scientists from their institution, the ORCA/ExoGENI team similarly enlisted the help of local experts from computational biology and chemistry that collaborated with the developers to configure the infrastructure to the needs of domain scientists. The ORCA team views CI applications as needing to be dynamically adapted, based upon domain scientists needs.

"*Basically the idea is to learn about the behavior of these complex applications that heretofore are run on grid systems and static resources. We look at their behavior as a more dynamic environment that can be scaled and configured to their [domain scientists'] needs*" (Issac, ORCA/ExoGENI)

Through their collaboration with domain scientists, the GENI stakeholders are engaging in the translation processes of actor-network theory. Enrolling the domain scientists into GENI where they may in turn enroll new experimenters, or in ANTs parlance, actants.

Conducting Outreach Activities

Studies of large-scale IT systems have discussed the need to identify actions of "user advocacy" whereby stakeholders of a system's user groups promote the new system to other stakeholders and potential new users [4,23]. However, some of the enrollment activities we see go beyond mere advocacy towards user outreach.

GENI managers are actively involved in outreach activities. Originally they tried to attract new experimenters though a series of town hall meetings held throughout the U.S. Currently they use direct outreach activities that advertise new uses of the infrastructure's components at conferences and workshops.

"*We are really encouraging people who want to organize workshops, run summer schools, all that kind of stuff to put in proposals and we will fund them, but then they're in charge*." (GENI Manager)

The managers also use more indirect practices, such as funding current experimenters and developers to train new users and to demo/present their work, as well as funding travel grants to get new people to attend the GENI conferences. That includes putting the developers and existing users in charge of workshops and GENI-wide user meetings, such as trouble-shooting sessions and roundtable design-discussions about new components of the infrastructure.

The ORCA/ExoGENI team mentioned previously do not have a pre-existing plan or strategy for attracting new users, rather the PI for the project stated that instead they are actively involved in “user outreach”. This means that they try to adapt the technology/infrastructure to the existing user groups while at the same time participating in the experiments themselves to understand how the system works from the users’ perspective. This way they are adding new users in the form of themselves, as well as expanding the technology to work with new user groups, widening the potential user-base.

In using this approach the ORCA/ExoGENI team are utilizing already existing relationships to expand to new users, creating new relationships between the infrastructure and themselves, as well as expending to a new user-base and forming additional ties to other domain experts outside of GENI’s immediate focus of computer science experiments. Through these actions the ORCA/ExoGENI team are engaging in the leveraging practices of the synergizing framework, by working to solidify already existing ties between stakeholders, and then utilizing these ties to create new relationships with others; in this case domain scientists.

In addition to bringing in new users, stakeholders are also solidifying the already existing relationships between the participants of the GENI infrastructures, or using those relationships to create new connections with other stakeholders.

User-Support

Most interviewees (developers and experimenters) mentioned that they receive or provide some type of technical support. Through this user-support stakeholders leverage their existing relationships with the respective other parties involved. For example when experimenters take grievances to developers and see changes in applications or procedures, the relationships are solidified.

“You know, we have […] dialogue with users […] at, like, this GEC and at previous GECs as to what they want”. (Jack, ProtoGENI)

One experimenter talked about how experimentation is a function of academic training that is lacking in their field. Supporting experimenters is matter of providing training in how to run scientific experiments, a more complicated issue than technical user-support.

“*It’s not like physics or chemistry, any of those really established scientific disciplines where there’s a clear process for experimentation. There is not such a thing in networking. We’re still learning how to do this”* (Jessica, GENI Experimenter)

Other developers talked about trying to make their technologies and frameworks usable by providing user support through listservs, being accessible and responsive to users through email and in person during the GECs, and sometimes by trying to support multiple user groups such as experts and students. For example, Julia, the main GUSH developer, noted that the goal of GUSH was to support both advanced and novice users.

In some cases experimenters have been involved in the development of GENI since its early stages. One interviewee described their team’s process of adapting ProtoGENI for experimental purposes. It involved understanding the underlying platform/architecture of the infrastructure and its dependencies, and then discussing their coding with the ProtoGENI developers to find solutions and improve access.

“[user-support is an issue of] *basically understanding the platform, understanding the system dependencies that they have and what we wanna do and if we have issues. Then we sort of discuss with them for workarounds and sort of the network connecting issues of getting people to access our software through ProtoGENI.”*(Peter, ProtoGENI)

They further mentioned that their team had many early-stage interactions with experimenters and that because of their team’s long involvement with GENI they can provide expert level peer-support to other experimenters that are new to GENI.

In the PlanetLab infrastructure, developers, and also sometimes Principal Investigators, often help experimenters with their requirements-definition process during the development of experimental technologies, even without going so far as to collaborate on the same research project. For example, one of the main developers of PlanetLab, mentioned that they have not done any experimentation outside of development, but had collaborated with experimenters in determining system requirements.

“I’ve certainly in PlanetLab interacted with a lot of experimenters. Something that we believe in pretty strongly in PlanetLab, and I think this is part of the reason why PlanetLab is successful is really engaging experimenters to figure out what are their requirements if they need new capabilities of the platform, trying to help build and deploy that. Actually the best case is if we can get the experimenters to actually help us build the new things and test it. We’ve had pretty good success there.” (Alex, PlanetLab)

PlanetLab developers encourage experimenters to be involved in development by working together to determine system requirements for PlanetLab, or ways in which PlanetLab can be adapted to support experimentation. In doing so the PlanetLab developers work to leverage and solidify their existing relationships with experimenters while also forging new ties with them as they take on a small developer role. These efforts are intended to make experimentation across the GENI infrastructures consistent. Not only does this act to reduce overhead for experimenters, but it also encourages synergizing's alignment sub-process, where-by experimenters seamlessly switch between infrastructures, forging new relationships with those infrastructures which works to encourage user retention.

Common Tools

Another practice of positioning stakeholders and new users is by leveraging GENI's existing systems, which includes modifying existing technologies and applications to produce new common systems that are more widely usable, or that can be re-used to build and expand new systems.

When GENI started, PlanetLab, along with several other projects, was sought after to join the GENI community. Originally there were four clusters involved in GENI, over time it became clear that a common API was needed to talk between them. As a result GENI moved towards creating a common user interface for experimenters that spanned across the various clusters and their underlying frameworks – allowing experimenters and developers to access different project resources with ease.

This project was called the GENI User Shell (GUSH) and it was designed - by both PlanetLab and GENI stakeholders - to allow users access to common GENI resources, with minimal effort (usually requiring just a small XML header change to switch frameworks). GUSH was heavily based off the pre-existing PlanetLab User Shell (PLUSH) utilizing around half of its original code. In addition to the utilization of the PLUSH code, a number of projects that were once part of PlanetLab were rolled into GENI's GUSH.

In particular, the developers worked on improving the interoperability among the infrastructure's components. This is a process whereby a tool is created to allow individuals to utilize resources across several platforms seamlessly, for example GENICloud, which allows experimenters to easily deploy and configure slices that span PlanetLab and Eucalyptus (an ORCA project). Traditionally in infrastructure development, the job of creating such tools has been the developers' responsibility. Those trying to work with multiple incompatible components of the larger infrastructure would devote some part of their work to integrating those components, before being able to use them in their research projects. Occasionally the GENI Project Office (GPO) has taken charge of that task as well:

"*The GPO has gone so far as to write software to make things interoperable, so we've occasionally gone to that length.*" (GENI Manager)

Technology Integration

In an effort to attract more PlanetLab experimenters to GENI there has recently been a push to further integrate PlanetLab resources into the GENI infrastructure, and leverage the existing relationships, by placing PlanetLab nodes into InstaGENI racks.

"So this model of using InstaGENI we hope will be attractive to PlanetLab community. They're already familiar with PlanetLab tools going to the PlanetLab website, setting things up, and so we're hoping to steer some PlanetLab users into GENI in this way"(PlanetLab Manager)

Despite the existence of a shared GENI API and a drive for further GENI integration, PlanetLab still maintains its own API, allowing PlanetLab machines to run as part of GENI, or independently as PlanetLab.

Integration Issues in PlanetLab

As suggested in the previous sections there have been a number of issues with the integration of PlanetLab into GENI. Despite the push to further assimilate the GENI infrastructure with PlanetLab there is still some question of where PlanetLab sits within GENI today.

Much of these issues stem from PlanetLab's initial involvement with GENI, unlike many other projects that were incorporated into GENI, PlanetLab never fully integrated itself, instead maintaining its own API, user-base and applications to run alongside GENIs. Essentially PlanetLab was never fully enrolled in GENI and the resulting relationship is one of uncertainty and instability.

One interviewee even mentioned that they were unsure if PlanetLab would continue to support the GENI API (in some form), although they expressed a wish for such support to continue. It would also seem that even PlanetLab is unsure of its future direction, or if that will continue to include GENI, as one PlanetLab interviewee stated that:

"PlanetLab isn't just about GENI. We have our own vision, which overlaps a lot with GENI, but it's maybe not exactly the same, so on the PlanetLab side we're exploring our own kind of agenda with trying to reach out to the SDN community and try to figure out."

Despite the effort by various GENI stakeholders over the years to better integrate PlanetLab, the future of PlanetLab within GENI still looks uncertain. Continued efforts to integrate PlanetLab nodes and ProtoGENI racks, and the support of the GENI API by PlanetLab indicate a continuation of PlanetLab's incorporation into GENI, at least for the short term. However, it is apparent that PlanetLab has its own goals and there is a possibility that such goals could involve leaving GENI behind.

There was some initial success with GENI enrolling PlanetLab, after PlanetLab joined the GENI effort and contributed its resources to the development of GUSH, but over time efforts to further attract the PlanetLab community seem to stall, not only because of technical differences, but also because of slight differences in vision, and agendas. Leaving PlanetLab and GENI's combined future ambiguous.

Much of this ambiguity stems from the short-term vision of both PlanetLab and GENI during their original integration. From an actor-network perspective, the process of translation was never truly completed. Although GENI has a vision for what it wanted from PlanetLab, PlanetLab never completely accepted this, as evidenced by their decisions to keep much of their own applications and technologies active alongside GENI's. Both GENI and PlanetLab had their own initial goals when they came together, but neither truly had an idea for what the long term relationship between the two projects would be. The result has therefore been a continued ineffectiveness in attempts to reposition and leverage the existing relationships between PlanetLab and GENI and an uncertain future for the ties between the two.

DISCUSSION

Positioning stakeholders within a cyberinfrastructure involves complex and interrelated sociotechnical processes. We know from past research that cyberinfrastructures can be considered as "systems of systems" in the sense of including multiple inter-dependent technologies, software, networks and databases. This is certainly the case with GENI as it is a system of interrelated infrastructures, networks, and other smaller technical components.

Our findings present a number of implications for the design of future cyberinfrastructures. We can see from the example of PlanetLab's ineffective integration with GENI that not designing for the long term has negative consequences. In other GENI infrastructures the positioning practices they engaged in demonstrated how stakeholders are actively involved in the enrollment practices of actor-network theory, as well as the alignment and leveraging practices of synergizing. These infrastructures engaged in designed, conscious and forceful efforts to reposition existing relationships and form new ones in processes which helped to maintain and sustain the cyberinfrastructure. Whereas in the PlanetLab integration there was no design for these. Without such considerations for designing for synergizing and enrollment, the attempts to leverage existing relationships or further align/enroll new users from PlanetLab were continually ineffective; to the point where today the future of PlanetLab in GENI is extremely uncertain. This shows the critical importance of designing cyberinfrastructures throughout their lifecycles for the long-term, with particular emphasis on designing for the continual repositioning and realignment processes of actor-network theory and synergizing.

In GENI, stakeholders are constantly involved in a process of recruiting new users to their projects, be this through active recruitment or through indirect practices such as "demoing" their work. In bringing these new users on board, stakeholders engage in the translation processes of actor-network theory. This starts when new individuals (or actants) are identified by existing stakeholders - or display an interest themselves independently - the interests and needs of these new potential stakeholders are then teased out, and current stakeholders outline how GENI can be a solution to their problems; this is the *problematization* stage. A negotiation then takes place where changes are made and roles defined depending on an actants needs, once this negotiation is complete the actants become full stakeholders, free to bring new stakeholders into the project as they wish.

The enrollment of new experts and individual experimenters makes a difference for the long-term sustainability of GENI's infrastructures. By bringing in new stakeholders in this way GENI combats user-attrition and allows itself to become increasingly diverse, giving it a robust and sustainable user-base. Such enrollment also aided groups in gaining additional funding, and expanding their infrastructure by enrolling not only experimenters but also scientists from other disciplines such as computational biology.

Once stakeholders have been enrolled and brought into GENI there is still the risk of long-term user attrition as the needs of the enrolled user-base changes over time. In GENI however we see stakeholders engaging in activities whereby they utilize established relationships in order to adapt infrastructure to meet the changing needs of their users. This in turn allows them to solidify their already existing ties and open up new relationships as the changes attract new users to the projects. These actions are examples of the leveraging processes of the synergizing framework and not only is this occurring among a diverse group of individual stakeholders (or potential individual stakeholders) but it is also occurring at the higher levels of whole infrastructures. Stakeholder positioning in GENI also takes on the flavor of co-enrollment or of mutual alignment. Stakeholders try to align with other stakeholders, or potential stakeholders, incrementally and part-way by changing technologies or holding workshops. Potential stakeholders incrementally align with stakeholders by engaging in dialog about the potential of creating new system requirements or a common user "shell" to access resources. Common types of positioning practice entail promoting usability efforts requiring iteration between developers and experimenters who subsequently enroll still more experimenters.

Through these practices we can see how GENI stakeholders engage readily in the processes of synergizing. This enables us to show synergizing as a framework in action and also shows how actor-network theory and synergizing are complimentary and powerful theories involved in the design of successful and sustainable cyberinfrastructures. This is perhaps best demonstrated not by the success of the many GENI infrastructures but in the ineffectiveness of the

PlanetLab infrastructure's incorporation into GENI. Throughout the entire process of PlanetLab's participation in GENI it was unable to effectively leverage existing relationships or to sufficiently align its resources and practices with the rest of the GENI infrastructures. As a result, PlanetLab's position within GENI remains uncertain as it continues in a stagnant maintenance state, underutilized and with a shrinking user-base.

CONCLUSION

The design of cyberinfrastructures emerge over time through processes of stakeholder alignment, enrollment, and through synergies among developers and other stakeholder groups. This study examined how GENI stakeholders enroll new actors (individuals, groups, and technologies) into the cyberinfrastructures they are developing or working with, and how they leverage their existing relationships with other stakeholders to create sustainable infrastructure. GENI stakeholders (developers, experimenters, and managers) promote the development and use of the GENI infrastructures through various means and thus promote the enrollment of new stakeholders and technologies into their respective projects within GENI.

Further research is needed to build a more nuanced understanding of the sociotechnical processes underlying the development, adoption and use of cyberinfrastructures, particularly in complex scientific research domains such as GENI's "infrastructure of infrastructures". Consideration of the social and technical aspects of enrollment and leveraging practices can enrich explanations of how cyberinfrastructures can be sustained over the long term. We can then use this understanding to aid in the design of more sustainable cyberinfrastructures.

ACKNOWLEDGMENTS

This research was supported by NSF Award OCI-1220269. Thanks to our participants and to our anonymous reviewers for their feedback.

REFERENCES

1. Atkins, D. *Revolutionizing Science and Engineering Through Cyberinfrastructure: Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure.* 2003.
2. Barki, H. and Hartwick, J. Rethinking the Concept of User Involvement. *MIS Quarterly 13*, 1 (1989), 53.
3. Baroudi, J.J., Olson, M.H., and Ives, B. An empirical study of the impact of user involvement on system usage and information satisfaction. *Communications of the ACM 29*, 3 (1986), 232–238.
4. Bettencourt, L.A., Ostrom, A.L., Brown, S.W., and Roundtree, R.I. Client Co-Production in Knowledge-Intensive Business Services. *California Management Review 44*, 4 (2002).
5. Bietz, M.J., Baumer, E.P.S., and Lee, C.P. Synergizing in Cyberinfrastructure Development. *Computer Supported Cooperative Work (CSCW) 19*, 3-4 (2010), 245–281.
6. Bietz, M.J., Ferro, T., and Lee, C.P. Sustaining the development of cyberinfrastructure. *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work - CSCW '12*, ACM Press (2012), 901.
7. Callon, M. Some elements of a sociology of translation: domestication of the scallops and the fishermen of St Brieuc Bay. In J. Law, ed., *Power, action and belief: a new sociology of knowledge?*. Routledge, London, 1986, 196–223.
8. Charmaz, K. *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis.* Pine Forge Press, 2006.
9. Cockburn, A. *Agile Software Development.* Addison-Wesley, Reading, MA, 2001.
10. Davis, F. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly 13*, 3 (1989), 319 – 340.
11. Edwards, P.N., Jackson, S.J., Bowker, G.C., and Knobel, C.P. *Understanding infrastructure: Dynamics, tensions. and design.* 2007.
12. Edwards, W.K., Newman, M.W., and Poole, E.S. The infrastructure problem in HCI. *Proceedings of the 28th international conference on Human factors in computing systems - CHI '10*, ACM Press (2010), 423.
13. Eisenhardt, K.M. Building Theories from Case Study Research. *Academy of Management Review 14*, 4 (1989), 532–550.
14. Finholt, T.A. Collaboratories as a new form of scientific organization. *Economics of Innovation and New Technology 12*, 1 (2003), 5–25.
15. Galison, P. *Big Science: The Growth of Large-scale Research.* Stanford University Press, 1992.
16. Gallivan, M.. Value in triangulation: a comparison of two approaches for combining qualitative and quantitative methods. In A.. Lee, J. Liebenau and J. DeGross, eds., *Information Systems and Qualitative Research*. Chapman & Hall, London, 1997, 417–443.
17. Hey, T. and Trefethen, A. e-Science and its implications. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences 361*, 1809 (2003), 1809–25.
18. Latour, B. *Science in Action: How to Follow Scientists and Engineers Through Society.* Harvard University Press, 1987.
19. Latour, B. Reassembling the Social - An Introduction to Actor-Network-Theory. *Reassembling the Social - An Introduction to Actor-Network-Theory, by Bruno Latour, pp. 316. Foreword by Bruno Latour. Oxford University Press, Sep 2005. ISBN-10: 0199256047. ISBN-13: 9780199256044 -1*, (2005).

20. Law, J. Notes on the theory of the actor-network: Ordering, strategy, and heterogeneity. *Systems Practice 5*, 4 (1992), 379–393.

21. Lee, C.P., Dourish, P., and Mark, G. The human infrastructure of cyberinfrastructure. *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work - CSCW '06*, ACM Press (2006), 483.

22. Lee, J.W., Zhang, J., Zimmerman, A.S., and Lucia, A. DataNet: An emerging cyberinfrastructure for sharing, reusing and preserving digital data for scientific discovery and learning. *AIChE Journal 55*, 11 (2009), 2757–2764.

23. Mambrey, P., Mark, G., and Pankoke-Babatz, U. User Advocacy in Participatory Design: Designers' Experiences with a New Communication Channel. *Computer Supported Cooperative Work (CSCW) 7*, 3-4 (1998), 291–313.

24. Miranda, S.M. and Kavan, C.B. Moments of governance in IS outsourcing: conceptualizing effects of contracts on value capture and creation. *Journal of Information Technology 20*, 3 (2005), 152–169.

25. Mirel, B. *Interaction Design for Complex Problem Solving: Developing Useful and Usable Software.* Morgan Kaufmann, 2004.

26. Olson, G.M., Zimmerman, A., and Bos, N. Scientific Collaboration on the Internet. (2008).

27. Palys. *Research Decisions: Quantitative and Qualitative Perspectives.* Thompson, Toronto, 2008.

28. Procter, R., Borgman, C., Bowker, G., et al. Usability research challenges for cyberinfrastructure and tools. *CHI '06 extended abstracts on Human factors in computing systems - CHI EA '06*, ACM Press (2006), 1675.

29. Redish, J. Expanding Usability Testing to Evaluate Complex Systems. *Journal of Usability Studies 2*, 3 (2007), 102–111.

30. Ribes, D. and Finholt, T. The Long Now of Technology Infrastructure: Articulating Tensions in Development. *Journal of the Association for Information Systems. May2009 10*, 5 (2009).

31. Ribes, D. and Lee, C.P. Sociotechnical Studies of Cyberinfrastructure and e-Research: Current Themes and Future Trajectories. *Computer Supported Cooperative Work (CSCW) 19*, 3-4 (2010), 231–244.

32. Rogers, E.M. *Diffusion of Innovations, 4th Edition.* Simon and Schuster, 2010.

33. Schopf, J.M. Sustainability and the Office of CyberInfrastructure. *2009 Eighth IEEE International Symposium on Network Computing and Applications*, IEEE (2009), 1–3.

34. Şeker, Ş.E. Possible Social Impacts of E-Government: A Case Study of Turkey. 2004.

35. Shin, D.-H. A socio-technical framework for cyber-infrastructure design. *Technological Forecasting and Social Change 77*, 5 (2010), 783–795.

36. Sonnenwald, D.H. Scientific collaboration. *Annual Review of Information Science and Technology 41*, 1 (2007), 643–681.

37. Star, S.L. and Ruhleder, K. Steps Toward an Ecology of Infrastructure: Design and Access for Large Information Spaces. *Information Systems Research 7*, 1 (1996), 111–134.

38. Star, S.L. The Ethnography of Infrastructure. *American Behavioral Scientist 43*, 3 (1999), 377–391.

39. Steinhardt, S. and Jackson, S. Reconciling rhythms: plans and temporal alignment in collaborative scientific work. *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*, (2014), 134–145.

40. Stewart, C.A., Almes, G.T., and Wheeler, B.C. (eds. . Cyberinfrastructure Software Sustainability and Reusability: Report from an NSF-funded workshop. (2010).

41. Venkatesh, V., Morris, M.G., Davis, G.B., and Davis, F.D. User acceptance of information technology: toward a unified view. *MIS Quarterly 27*, 3 (2003), 425–478.

42. Walsham, G. Actor-network theory and IS research: current status and future prospects. In A.S. Lee, J. Liebenau and J.I. DeGross, eds., *Information Systems and Qualitative Research*. Springer US, Boston, MA, 1997, 466–480.

43. Wulf, W. The collaboratory opportunity. *Science 261*, 5123 (1993), 854–855.

44. Yin, R.K. *Case Study Research: Design and Methods.* SAGE, 2009.

45. Zimmerman, A. and Finholt, T.A. Growing an infrastructure. *Proceedings of the 2007 international ACM conference on Conference on supporting group work - GROUP '07*, ACM Press (2007), 239.

46. Zimmerman, A. and Nardi, B.A. Whither or whether HCI. *CHI '06 extended abstracts on Human factors in computing systems - CHI EA '06*, ACM Press (2006), 1601.

47. GENI » Meet the GENI Project Office (GPO). http://www.geni.net/?page_id=26.

48. National Research Council. Committee on Metagenomics: Challenges and Functional Applications New science of metagenomics: Revealing the secrets of our microbial planet0. 2007.