This paper is a preprint of an article to appear in CloudCom2014.

A Framework for Measuring the Impact and Effectiveness of the NEES Cyberinfrastructure for Earthquake Engineering

Thomas Hacker
Computer & Information Technology
Purdue University
West Lafayette, Indiana, USA
tjhacker@purdue.edu

Alejandra J. Magana
Computer & Information Technology and Engineering Education
Purdue University
West Lafayette, Indiana, USA
admagana@purdue.edu

Abstract— **Many cyberinfrastructure and cloud computing systems have been developed and deployed over the past decade. Although use metrics are collected by many of these systems, there is not a clear link from these metrics to the ultimate effectiveness and impact of these systems on science communities. This paper describes a framework we developed that seeks to provide context for use and impact metrics to facilitate understanding of how these systems are used and ultimately adopted by science and engineering communities. We use this framework to present metrics of use, impact, and effectiveness collected from the NEES cyberinfrastructure.**

Keywords—Cyberinfrastructure, cloud computing infrastructure, assessment, earthquake engineering, experimental data

I. INTRODUCTION

The ability to perform detailed computational simulations of physical and social systems has transformed research over the past 60 years. In the current decade, the growth of tremendous amounts of "big data" from sensors, simulations, and analysis is providing a new paradigm for research that adds to the traditional trio of theory, experiment, and simulation [8]. This new paradigm is built on the integration of data with cloud computing technologies, high performance computing capabilities, sensors, and visualization systems to create an always-on cyberinfrastructure accessible anywhere. Cyberinfrastructure (CI) refers to a distributed information technology infrastructure comprised of systems, software, databases, and visualization facilities, all interconnected with high-speed networks [7][8][12][20]. One of the key characteristics of CI it can leverage existing computing technologies to provide a fully integrated system that can be tailored to the specific needs of the users and researchers. CI includes *science gateways* that provides an easily accessible web interface to data, cloud computing resources, applications, and collaboration facilities. The George E. Brown, Jr. Network for Earthquake Engineering Simulation (NEES) project is a network of 14 civil engineering research laboratories that conducts experiments that generates digital sensor data. NEES relies daily on CI to globally distribute earthquake engineering data, knowledge, and software applications.

Cloud computing technologies and cyberinfrastructure have fundamentally changed the nature of research. However, it is not well understood to what degree these technologies have directly impacted science and engineering. Certainly, if the user community does not use the CI, it will not make any meaningful impact on the way in which the community conducts research. An early report by Atkins [1] describing the need for CI found that CI systems needed both powerful technical capabilities as well as a broad suite of functionality to provide the tools needed by researchers to more deeply explore a problem. Ultimately, the CI must be adopted and integrated into their work by the community to make an impact on research outcomes.

Considerable effort has been expended over the past decade to develop CI and cloud computing systems. While many of these systems collect metrics and measurements that aim to describe use, there is often no clear context or framework in which to present these measurements to understand use and impact. What is needed is a framework to provide a context and a basis for the collection and interpretation of metrics to provide insight into how cloud computing and CI systems are used and their impacts on research. We present such a framework in this paper for the NEES cyberinfrastructure.

II. RESEARCH COMMUNITY CI ADOPTION PROCESS

The adoption of CI and cloud computing technologies by educators and researchers is necessary for the technologies to influence education and research. Rogers (*Diffusion of Innovations* [14]) provides a useful framework for the technology adoption process based on consecutive stages: (1) knowledge, where a user discovers an innovation; (2) persuasion, where a user assesses the innovation to determine its suitability for their work; (3) decision, where a user considers an innovation to decide if they would benefit from adopting an innovation; (4) implementation, where a user uses the innovation and judges the benefits versus costs of the innovation; and finally (5) confirmation, when a user finally adopts the innovation.

This process can be used to understand how research communities discover, trial, and adopt new cloud computing and CI technologies. We developed a framework based on this process to understand and measure the use and effectiveness of cloud computing and CI technologies for science communities. In related prior work [20, 22], we explored how users adopted a cyberinfrastructure for engineering education based in part on an early draft of the adoption framework described in this paper. In this paper, we apply this framework to assess use and adoption based on detailed use measurements and prior work in developing and operating the NEES cyberinfrastructure [4-6], and demonstrate the long-term usefulness and validity of the adoption framework for understanding the use and adoption of cyberinfrastructure over an extended project period for a large user community.

A. *Cloud Computing and CI Use and Effectiveness*

NEES is a *community of practice* [12] wherein civil engineering researchers collaborate and share information toward a common goal. The domain of the NEES community, the members' common interest, is to improve the resilience of buildings, bridges, and other structures by engaging in research into new materials, improved construction techniques, and new types of structures that can better withstand damage from earthquakes and tsunamis. If the CI can support and extend the community of practice [14], an effective CI can make a positive impact on the overall total count of users and artifacts generated by the community. We argue that when these technologies are adopted by the community of practice, an effective CI will produce positive changes in community productivity.

To assess this, we define two terms: use and effectiveness. *Use* of the CI is a count of individuals using the offered cloud computing and CI technologies. *Use* is an indication of the potential of the technologies to influence the community of practice. This influence results from the interaction of the community with the technology that changes the practices or thinking of the community. *Effectiveness* is the degree to which the CI and cloud computing technology directly affects the productivity and impact outcomes of the community.

Use can be quantified by website visitors, software downloads, tool invocations in a cloud environment, data uploads and downloads, and survey results. Measuring effectiveness is more complex, as it involves productivity resulting from several years of use by the community. Examples of effectiveness include the rate of research paper production, course evaluations, references to published datasets and software applications. Clearly *use* of the cloud computing and CI technologies must precede *effectiveness*, and measures of use and effectiveness are qualitatively and quantitatively different.

B. *A framework for measuring CI impact*

We can use Roger's five stages to define a framework and process to measure and understand the use and adoption of cloud computing and CI technologies. The adapted stages (in sequence) are: (1) *knowledge*, in which an individual learns information about the CI along with a basic understanding about the benefits of the system; (2) *persuasion*, where the user forms a favorable or unfavorable opinion about the CI as the result of an initial assessment of the system; (3) *decision*, in which the user tests the CI to determine its suitability for their work; (4) *implementation*, where the user uses the CI to a greater degree and further evaluates its usefulness; and finally (5) *confirmation*, in which a user commits to use the CI. This stage has the greatest potential to ultimately affect productivity and research outcomes.

III. Describing the Use and Assessing the Effectiveness of the NEES CI

The 14 NEES laboratories include facilities such as a wave basin for studying the effects of tsunamis, geotechnical centrifuges, a facility to assess displacement damage to underground structures, and shake tables. NEEScomm has operated a comprehensive cyberinfrastructure for nearly five years that serves the needs of civil engineering researchers and laboratory staff. The NEES practice includes objectives to operate a national network of civil engineering research laboratories that can facilitate this research as well as engage in undergraduate and graduate education, and help inform the community of civil engineering practitioners.

To track user traffic to the NEEShub, NEEScomm uses Google Analytics [2, 3]. Figure 1 shows the monthly count of visitors to nees.org from returning and new visitors. Each event shown in this figure, a pageview, is an individual visit by a

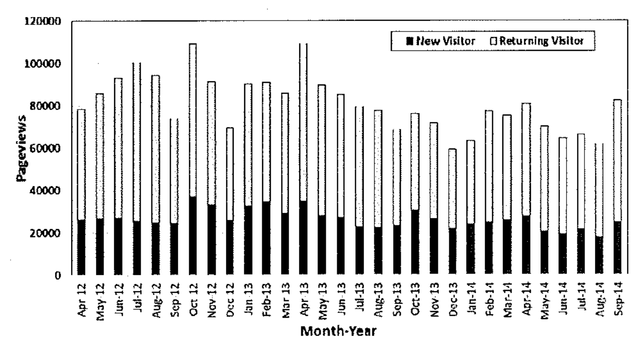

Fig. 1. NEEShub visits.

user to a webpage detected by Google Analytics somewhere on the nees.org website. We can also determine the count of users (defined as active users in Google Analytics) who access the NEEShub. The count of active users for whom Google Analytics measured at least one activity over a 30-day period totaled over each month is shown in Figure 2. We can use detailed measurements such as the data shown in Figures 1 and 2 to create metrics to measure progress for each of the five technology adoption stages.

The level of detail of use measurement depends on the degree to which a system can measure user activity. Google Analytics [2] collects the IP address of the user and uses tracking cookies. These data can be used to measure the user's activity over time as well as establish the user's location. The NEEShub, since it allows users to register and login, can more closely measure user activities such as user login events, tool use, file download, and content creation.

The next section of this paper presents detailed descriptions of each of the five CI technology adoption stages along with detailed statistics reported by Google Analytics and NEEShub that describe the use and impact of the NEES CI by the earthquake engineering community.

A. Knowledge Stage

In *knowledge*, the first stage of adoption, users discover the CI and learn about the basic advantages and functions provided by the CI. The trigger for entering this stage occurs when a user searches for the CI using a concept (such as NEEShub), directly enters a uniform resource locator (URL), or downloads software they have learned about. Users in this first stage are driven by curiosity about the CI and seek to "check it out" for a

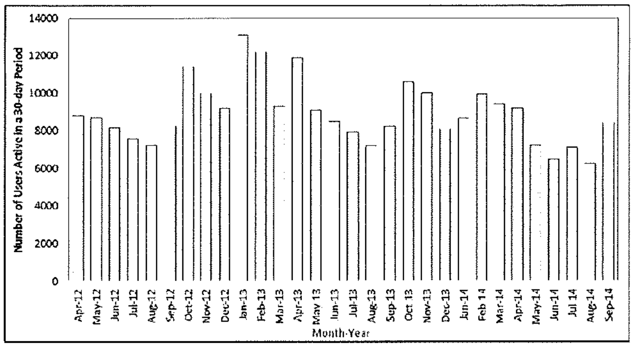

Fig. 2. Count of Active Users.

first exposure to the content and to seek out additional knowledge about the features and resources provided by the infrastructure.

The first level of use measurement is the amount of traffic to the NEEShub, which indicates the count of visitors to the site. For these users, we can use information gathered by Google Analytics. The count of monthly pageviews from visitors during the period April 2012 through September 2014 is shown in Figure 1. The total count of pageviews over this period was 2.42M. One clear trend in pageview activity is a seasonal periodicity in the use of the NEEShub with around two peaks per year. The average count of new visitors per month over months shown in Figure 1 was 26K with a standard deviation of 4.6K. The count of pageviews for returning visitors was an average of 54K with a standard deviation of 10K. Over this time there was approximately the same variation in pageviews for new visitors (COV of 0.18) and returning users. The decline in the count of users from May 2013 to December 2013 was likely due to the recompetition for the NEES project as well as the seasonal load cycle. The reduction in visits is also influenced by projects at NEES laboratories finishing up as the community reaches the end of the five-year project. We conclude from these data that over time the rate of new researchers has been consistent, and that large changes in webpage visits over time come mostly from returning visitors.

Another measurement is the count of active users measured by Google Analytics. Figure 2 shows this information from users who visited the NEEShub at least once in a 30 day period at the end of each month from April 2012 through September 2014. The count of users peaked in January 2013 at 13,095 users. The seasonal periodicity of pageviews can also been seen in the count of users over time.

Using information about user browsing gathered by Google Analytics, we make two observations. First, a fairly consistent number of new individuals have been and are visiting the NEEShub every month. The flow of new community members interested in the CI is a critical factor, since it is the initial population from which users are selected, and represents the entry point of users who go through the CI assessment and adoption process. This process leads to a population of users who decide to sign up for a NEEShub account and use the CI on a regular basis. The second observation is that in later months the CI seems to have reached a saturation point, where annual cycles of use dominate changes in visitor count.

The potential of the CI to influence research is limited if visitors are from a limited geographic scope, such as a single institution or state. From April 2012 to July 2014, Google Analytics recorded visitors from 203 countries. 65.6% of the traffic was from the Americas, 18.4% from Asia, and 12.1% from Europe. From these data, it is clear that the NEES CI is serving a global community of users interested in earthquake engineering

B. Persuasion Stage

The next stage of the CI adoption process is persuasion, in which a potential user forms a favorable or unfavorable judgment about the new technology from an initial assessment of the CI resulting from activities during the knowledge stage. In this stage, an individual seeks additional data about the CI to decrease their uncertainty about the innovation and the consequences of its use. In our framework, return visitors to the NEEShub represent users who are more carefully and thoroughly exploring the system to gather additional information and knowledge about the system. Figure 1 shows that a large fraction of the total NEEShub pageviews are due to return visitors. We computed the monthly ratio between new and return visits, and found the average was 2.08 with a

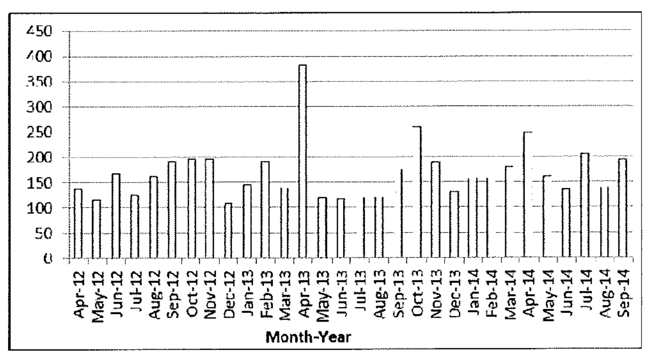

Fig. 3. Monthly count of new NEES user accounts.

standard deviation of 0.35. Thus, on average for every webpage visit from a new user, there are approximately 2.1 webpage visits from returning users.

C. Decision Stage

The decision stage is the third stage in adoption. At this stage, a user tests the CI to determine if they will adopt the technology. Users are required to establish a NEEShub account to run software applications on the NEEShub, upload content and project data, and to create groups. Users who take this step have transitioned into the decision stage through the act of creating their own user account.

The monthly count of accounts created on the NEEShub from April 2012 to September 2014 is shown in Figure 3. The count of new accounts varied per month, but stayed within a steady range with an average of 169 per month after April 2012. From this, we infer that the NEES CI is able to consistently attract new users who are willing to take the step of creating their own user account and entering the decision stage. The continuing increase in the sum of user accounts demonstrates the size of the community that have taken a clear interest in the NEES CI and are willing to explore it to some degree.

The NEEShub collects information on the activity of people who visit and use the CI, which is partitioned into three sets of users by the measurement system. The top area in Figure 4 represents download users who are unregistered. These users do not login and only download material. The middle area represents interactive users, who use the NEEShub as unregistered users. The bottom represents users with an account who logged into the NEES CI. Figure 4 shows the count of NEEShub users per month from April 2012 through August 2014. Overall user activity has grown over time until the most recent months, with the vast majority of activity arising from unregistered users. From April 2012 onward, the

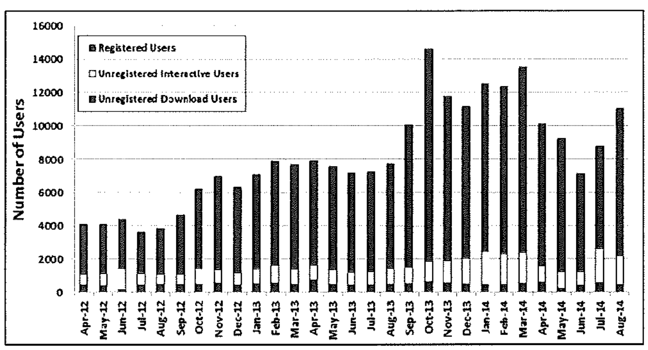

Fig. 4. NEEShub User Count.

average count of registered users was 509, and for unregistered users there was an average of 1.1K interactive and 6.6K download users.

Figures 3 and 4 show a growing flow of unregistered users accessing the NEEShub. A declining fraction of registered users are using their NEEShub accounts. These users may be accessing what they need without signing in, or may simply have never signed up for an account. To more closely track download activity, the NEEShub recently changed to require accounts to access some of the material on the NEEShub. From the information shown in Figures 3 and 4, we infer that the overall count of users who go through the persuasion stage is continuing to increase on an annual basis.

D. Implementation Stage

After a user has assessed the cyberinfrastructure, they may decide to implement the innovation for their work. In this stage, the user uses the innovation and further evaluates its usefulness. In our framework, when a NEEShub cyberinfrastructure user has reached the implementation stage, they have tested the usefulness of the NEEShub and are prepared to use the CI.

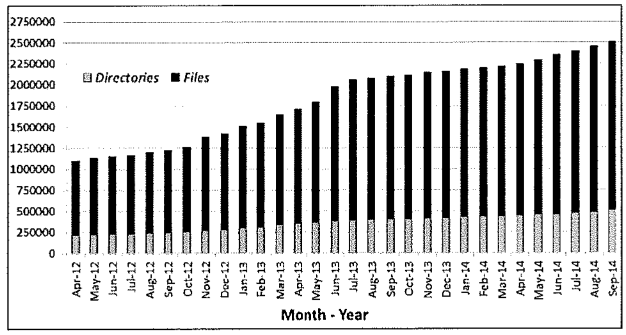

Fig. 5. Count of Files and Directories in the Data Repository from April 2012 to September 2014.

One measure of active users is the overall aggregate size of the content provided by users. Figure 5 shows a monthly cumulative total of files and directories deposited to the NEES Project Warehouse from April 2012 to September 2014. These files and directories are created and uploaded as a part of projects to the NEES data repository. As of the end of September 2014, there are 132 completed projects, 32 active projects, and 23 incomplete projects. Each project may contain thousands of files. As of September 2014, there are 2M project files and 511K directories containing data uploaded to the NEEShub by the community. These data represent the lasting legacy of the NEES project, as they contain valuable civil engineering data for projects as far back as 1976. Thus, users who have created and populated projects have made a significant investment in effort in the NEES CI. The data repository now contains a significant amount of valuable data for earthquake engineering, and has evolved over the past five years to become the most heavily used element of the cyberinfrastructure.

Above and beyond the project content, users can also upload additional resource elements (such as files, links, and videos) to the NEEShub that can be shared with others in the community. Table I shows an annual sample of the elements uploaded to the NEEShub beginning in January 2012. There were over 4,592 elements uploaded beyond Project Warehouse content over this time. Figure 6 shows a breakdown by type of element. A final measure of the use of the NEEShub by the community is the count of collaborative groups created in the NEEShub. Table I shows the annual count of created groups from January 2012 to October 2014, with a total of 369 groups created over this time.

To measure the use of NEEShub modules, we gathered the reported unique pageviews from Google Analytics on returning and new visitors from July 2012 through June 2014. The top 25

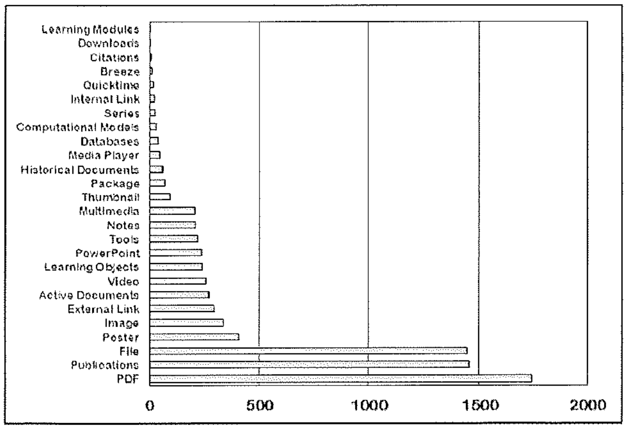

Fig. 6. Contributed elements (by type) from January 2012 through October 2014.

webpages visited each year were categorized by the section of NEEShub visited. The breakdown of pageviews by category of the NEEShub module visited is shown in Figure 7. The definition of a unique pageview in Google Analytics is one or more visits to a webpage by a user over a period of time over which the user is not idle on the NEEShub more than 30 minutes (defined as a session by Google Analytics).

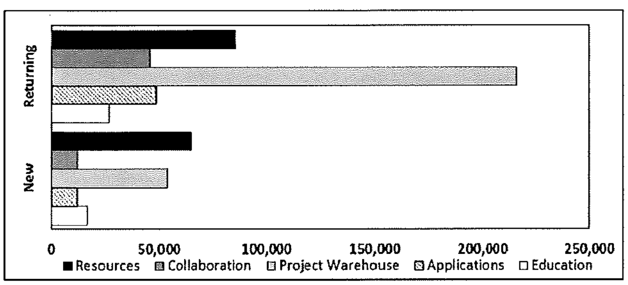

Fig. 7. Count of NEEShub module visits from July 2012 through June 2014.

This figure shows some characteristics of the use of the NEEShub during the implementation stage. First, the Project Warehouse, Applications, Collaboration, Education, and Resources modules are used primarily by return visitors who are working on the NEEShub for at least a 30 minute period (a session). The Project Warehouse is used primarily by returning users. The Collaboration and Application modules are used the least by new visitors, and the Resources module is accessed more frequently by new users than the Project Warehouse. Thus, new users are most frequently seeking out new resources, and returning users often return to the Project Warehouse.

Figures 5 through 7 and Table I shows that there has been an overall increase in contributions and frequent use of the NEEShub by the community. These trends demonstrate the individuals from the community are engaged in studying and using the NEEShub cyberinfrastructure. Based on these trends, we conclude that the set of users making contributions have transitioned to the implementation stage, assessing content to finally decide if they will integrate the NEEShub into their work.

E. Confirmation Stage

The confirmation stage is where there is the greatest potential for cyberinfrastructure to make a lasting effect on research outcomes and the pace of research. In this stage, the user has committed to adopt and use the cyberinfrastructure.

Understanding and quantifying the productivity of a research community is a complex issue that is beyond the scope of this paper. One quantitative way in which we can get a sense of the pace of research is the annual publication count of the community. Publication is a lagging metric, since the time between the start of a project and the final publication of project results can be several years. The time between manuscript submission and final publication in a journal can often take as long as two years.

TABLE I. CONTRIBUTED ELEMENTS AND GROUPS CREATED IN THE NEESHUB

Year	Annual Groups Created	Annual Elements
2012	170	929
2013	109	1289
2014	90	2374

The NEEShub allows the earthquake engineering community to quickly disseminate experimental data and electronic documents. Consequently, material uploaded by users such as documents, experimental data, software, and learning modules represent a leading metric of the size of the community who will make long-term use of the cyberinfrastructure. Consequently, it is a factor that indicates the future number of publications that ultimately will be produced by the community. At the end of 2012, 3,309 references were identified that resulted from work in the NEES network (https://nees.org/resources/7260). At the end of 2013, there was an increase of 653 references over the year for a total of 3,962 references. Publications have increased continually over the past five years, and we expect a continued increase in publications and citations to those publications.

IV. DISCUSSION AND LESSONS LEARNED

Tremendous investments have been made over the past decade to develop cyberinfrastructure systems to meet the needs of science and engineering communities. Although a vast number of measurements and metrics have been collected and disseminated, it has not been clear how individual metrics fit within a clear and cogent framework that describes how the efforts measured by these metrics ultimately lead to improvements in research and educational outcomes. Many metrics such as the ones we show above can be confusing when presented alone. For example, pageviews from Google is one view of activity from users visiting the NEEShub. While this number is useful, it doesn't capture the count of users who actually use the NEEShub in a more significant manner. This can only be measured using login statistics. With these two figures (since Google pageviews also captures activity from logged in users), it becomes possible to develop a rough estimate of the fraction of users who are going beyond browsing to more thoroughly interact with the CI. Rogers' Diffusion of Innovations model provides a sound structure to define a technology adoption framework to understand and measure the rate of adoption of cyberinfrastructure innovations and how efforts measured by these metrics lead to research and educational outcomes.

Using the cyberinfrastructure adoption framework and the metrics we have collected, we can make several observations about the cyberinfrastructure elements that have an effect on earthquake engineering research outcomes.

First, there have been a steady or increasing number of users interacting with the NEES cyberinfrastructure. Rogers describes the ideal types of adopter categories by their level of comfort and willingness to go through the adoption process, which include innovators, early adopters, majority, and laggards. Considering this distribution, if the count of new visitors and annual rate of registered users decline, the pipeline of new users will decrease, and consequently growth in users from the early and late majority categories will be limited. This decline can seriously impede the adoption and use of the system in the research community. Consequently, providers need to closely monitor the count of users who are progressing through the knowledge stage and focus on advertising the system to the community of practice to continue to attract potential new users.

The second observation is that the cyberinfrastructure adoption process can be thought of as a pipeline, in which new users go through the stages of the process to ultimately adopt (or potentially reject) the cyberinfrastructure. At any stage in the adoption process, users may decide to abandon the system. Providers need to monitor this rate of abandonment and diagnose and repair any element of the system that impedes adoption. To address this, NEEScomm is working to attract and convince new users to try the NEEShub and initiate the cyberinfrastructure assessment and adoption process. The NEEScomm group has been working to do this by holding training "boot camp" events for graduate students and workshops, advertising the NEEShub and its capabilities, and disseminating information about the NEEShub in trade publications [13]. NEEScomm also involves civil engineers in every aspect of system design and operation to ensure that any impediments are quickly identified and corrected. We believe that it should be possible to measure the effectiveness of these campaigns and efforts throughout the knowledge stage by monitoring the Google Analytics and NEEShub metrics. One item that we are working to better understand is how to quantify the point at which users transition from one stage to another. This may require a range of values to capture the behavior of a group or percentile of users. Another critical factor is maintaining persistent use of the NEEShub by users who have made a lasting decision to use the NEES cyberinfrastructure.

The third observation is that the destructive nature of the experiments in the NEES network makes the sharing and reuse of data an essential component of the NEES cyberinfrastructure. The growing number of data deposits over the past five years in the Project Warehouse shows that earthquake engineering users see the NEEShub as a viable repository for valuable and irreproducible earthquake engineering data. Ultimately, the Project Warehouse is likely to become the most enduring asset of the NEES system over the next decade. As in the case in attracting new users, we should be able to monitor the effectiveness of efforts to increase the reuse of data by monitoring the metrics described for users proceeding through the implementation stage.

Through these observations we can say that the NEES cyberinfrastructure is providing an environment and is meeting the challenges faced by the civil engineering community of practice. Wenger [15] describes the challenges that drive *digital habitats* as: rhythms (time and space separating members seeking collaboration; interactions (producing artifacts and engaging with others in sharing and discussing the artifacts); and identities (engagement as an individual or within a group). To support rhythms, the cyberinfrastructure provides group spaces, access to web conferencing systems, and projects spaces to facilitate in-person online group work as well as offline activities. In terms of interactions, the tremendous amount of contributed content demonstrates the level of involvement of the NEES community. The CI has also allowed users to interact as through identities as individuals or as part of a group. The large number of projects demonstrates the impact of the NEES CI on allowing individuals to establish identities.

Kim and Crowston [18] provided a summary and survey of technology adoption theories that seek to understand the factors that influence the probability of adoption of CI systems by scientists. Their survey of work exploring the Rogers model found that the relative advantage provided by the technology had a beneficial effect on user's attitudes to adoption. Complexity, however, had an inhibitor effect on adoption. Kim's model focused on factors influencing the initial and post adoption stages. In contrast to Kim's approach, the one proposed here does not address intrinsic and extrinsic motivation factors, but focus instead on quantitative measurements of use and productivity. The approach described in this paper complements Kim's model.

Similarly, Shin [19] described the development of cyberinfrastructure in Korea as a "socio-technical artifact" and describes the social and technical challenges related to designing and developing a CI in Korea. Shin concluded that CI development in Korea will require a sociological element to successfully integrate multidisciplinary teams. He noted that a top-down approach to developing CI left important voices from the user community out of the process. The experiences of the NEES team in designing the NEES cyberinfrastructure and cloud computing capabilities needed for the civil engineering community echoed Shin's experiences. The NEES team found engaging the research community was one of the most important aspects in the design and operation of the CI. Carefully addressing the social aspects of the project was as important and necessary as addressing the technological elements.

Barjak, et. al. [21] described cases studies of the how 'e-Infrastructure' (similar to CI) is adopted in the social sciences. Barjak did not utilize a framework inspired by Rogers, but instead described factors that influence the design and implementation of e-Infrastructure systems. Many of the influences he described were also the experience of the NEES team. For example, the need to forge a close link between software engineers and users. Our approach differs from Barjak in that we seek to quantify how the CI is used, and the effectiveness and impact of the CI over time.

Fry and Thelwall [23] measured hyperlinking patterns for managed text and language materials within a federation of archive centers (CLARIN) that provided access to resources through a web service. She sought to measure the diffusion of resources through the pattern of interlinking among the resources, and found a low degree of interlinking. The approach presented in this paper provides a greater level of detail and measurement, and greatly extends Fry's approach.

V. IMPLICATIONS FOR TRAINING, OUTREACH, AND ASSESSMENT

The development and use of cyberinfrastructure capabilities is a means to positively influence the research and educational effectiveness of a community, rather than being an "ends in itself". That is, to allow the community to learn from each other, the NEES cyberinfrastructure facilities information and experience sharing [12]. The clear vision communicated from civil engineers of the ways in which the NEES CI should be implemented and deployed has avoided the risks of creating a system that serves no use purpose for the community. A recent exploratory survey study of NEEShub users that we conducted suggests that that there is a strong positive correlation (r=0.83) between users' perceptions of NEES attributes and their frequency of use [16].

To encourage the adoption of this digital habitat, NEES provides training, information, and materials to inform the community about the CI capabilities (to encourage users to enter the knowledge stage), and to progress through the adoption process to the final confirmation stage. Rogers [14] identified several characteristics of innovations that influence adoption: *relative advantage*, *compatibility*, *complexity*, *triability*, and *observability*.

To allow individuals to gather detailed information about the relative advantages of the CI, the NEEShub allows users to engage in *legitimate peripheral participation* [15] to build their knowledge about the community and capabilities before establishing a visible identity and making a commitment to go through the adoption process. This step is necessary to allow individuals to understand the new advantages and increased value for their work provided by the CI to motivate them to invest effort in working through the adoption process. Another characteristic important to users is compatibility and alignment with the ways in which research is conducted in the community of practice and the NEES research laboratories. To ensure this alignment, the NEES team included civil engineering researchers at every step and level of the software development and prototyping process. This has been one of the most essential elements of the effort to provide a system that would be useful and relevant for the community. Another aspect is the ability of the CI to reflect the standards and values of the community. An example to illustrate this is curation activity. Researchers who upload data from a NEES project must provide an adequate level of additional description (metadata) for the data and files they provide for the benefit of future researchers. The curation status of their project and all of the projects from the same laboratory are shared publicly on the NEEShub. The sponsor of NEES research projects has communicated to the community that their projects must complete data curation before the projects are considered to be completed. The third element that affects adoption is the level of complexity presented to the user by the CI or cloud computing system. The NEEShub has designed to be simple, easy-to-use and intuitive for civil engineers. The interfaces and information display designs were based on formats conforming to the ways in which civil engineers represent and seek out information. The fourth element, triability, was another factor considered in the design of the system. This allows users to test an innovation (such as a new software tool that can be run in the NEEShub) with very little effort. This allows users to quickly assess the innovation without making a large investment in time and effort. The final element, observability, represents the externally visible results from the use of cyberinfrastructure on the community of practice. The NEEShub allows users to easily share data, projects, educational materials, and software they development with the community of practice.

By being cognizant of these elements that influence adoption in the design and development of the CI, individuals from the community of practice are empowered to easily identify and understand the benefits and advantages provided by the CI that go beyond their existing working environment [9][10].

VI. SUMMARY AND CONCLUSIONS

NEEScomm developed and operates a successful and well-used cyberinfrastructure to aid the research efforts of the civil engineering community. The NEES cyberinfrastructure effort was focused first and foremost on meeting the needs of NEES researchers. However, we found that the cyberinfrastructure was also able to meet a global need for accessing and dissemination research results for the international community of researchers, and was successful in attracting a growing community of users. As a part of measuring use and collecting metrics for the CI, we found that standalone metrics and measurements describing user activity made it difficult to synthesize a comprehensive picture of how the CI impacts the research and education activities of the community. For example, measuring pageviews alone did not provide insight into how the users were making use of the CI. By comparing the pageview measurements with the count of users who created their own user account, we were able to understand the level of effort users will willing to make to use capabilities provided by the CI. This was the motivation for the five stage diffusion process described in this paper.

Several important lessons were learned over the course of the project about how to best organize, manage, develop, deliver, and operate an effective system to serve the needs of a science community. The first and most critical factor is the need to solicit and support the integration of civil engineering graduate students and faculty into the overall process. This provided instant feedback and a "reality check" on software and cyberinfrastructure features. The second lesson learned involved exploiting the benefits of centralizing software development efforts. Distributing development over 14 sites would have led to tremendous communication and management overhead. The siting of development staff with the NEES headquarters helped to keep efforts focused on the most essential elements of the system. The data curation effort was also located at headquarters to provide a consistent set of policies and standards for project data. Beyond the curated data repository, the community is now working to create collections of community databases and data from articles [17].

Finally, over the course of the project there has been growth in the use of the CI by the global civil engineering research community beyond the NEES community from which the CI was developed. We assume that eventually the NEES CI may reach a saturation point in the number of civil engineering researchers who are interested in using CI. We thus need to ensure that the system continues to evolve and expand to increase the potential community of users. The tremendous amount of information in the data repository of content contributed by civil engineering researchers is a clear indicator of the utility of the CI. The diffusion of innovations approach we described in this paper is a useful framework to understand the degree to which the CI is used, and to begin to understand the impact the CI on the research community.

In conclusion, this paper described a framework to understand the impact, use, and effectiveness of CI systems for research. We focused on the NEES CI, and described a detailed set of measurements collected from two independent sources. Our measurements show a growing community of users who use the NEES CI throughout the entire CI adoption process. We found that the CI seems to be having a positive influence on the rate of research production measured by the annual rate of publications produced by the community.

VII. ACKNOWLEDGEMENTS

We would like to acknowledge the efforts of the NEEScomm IT and HUBzero teams at Purdue University in developing and operating the NEES CI and the NEEShub. The work in this paper is supported by NSF grant CMMI-0927178.

REFERENCES

[1] Atkins, D. A Report from the U.S. National Science Foundation Blue Ribbon Panel on Cyberinfrastructure, 2003.

[2] Clifton, B. Advanced Web Metrics with Google Analytics, 2nd Edition: SYBEX Inc., 2012.

[3] Google_Analytics. What is the difference between clicks, visits, visitors, page views and unique page views?. Retrieved from http://www.google.com/support/analytics/bin/answer.py?hl=en-GB&answer=57164, August 2014.

[4] Hacker, T. J., Eigenmann, R., Bagchi, S., Irfanoglu, A., Pujo, I., Catlin, A., & Rathje, E. The NEEShub Cyberinfrastructure for Earthquake Engineering, Computing in Science and Engineering, 13, 67-78., 2011.

[5] Hacker, T., Eigenmann, R., & Rathje, E.. Advancing Earthquake Engineering Research through Cyberinfrastructure. Journal of Structural Engineering, 139(7), 1099-1111, 2013.

[6] Eigenmann, R., Hacker, T., & Rathje, E.. Nees cyberinfrastructure: A foundation for innovative research and education. In Proceedings of the 9th US/10th Canadian Conference on Earthquake Engineering , 25-29, 2010.

[7] Hacker, T. J., & Wheeler, B. C. Making Research Cyberinfrastructure a Strategic Choice. Educause Quarterly, 30(1), 9, 2007.

[8] Hey, A. J. G., Tansley, S., & Tolle, K. M. The fourth paradigm: data-intensive scientific discovery: Microsoft Research Redmond, WA., 2009

[9] Jackson, M., & Woodward, R. EarthScope: A distributed, multi-purpose geophysical observatory for the structure and dynamics of the North American continent. Geophysical Research Abstracts, 12, 2010.

[10] Jackson, S. J., Edwards, P. N., Bowker, G. C., & Knobel, C. P.. Understanding infrastructure: history, heuristics, and cyberinfrastructure policy. First Monday, 12(6), 2007

[11] Jankowski, N. W.. Exploring e Science: An Introduction. Journal of Computer Mediated Communication, 12(2), 549-562, 2007.

[12] Lave, J., & Wenger, E. Situated learning: Legitimate peripheral participation: Cambridge university press, 1991.

[13] Rathje, E. NEES News: NEEShub Project Warehouse Data, EERI Newsletter, June 2011.

[14] Rogers, E. M. Diffusion of innovations. Simon and Schuster, 2010.

[15] Wenger, E., White, N., & Smith, J. D. Digital Habitats; Stewarding Technology for Communities: CPsquare, 2009

[16] Taleyarkhan, M., Alabi, O., Magana, A.J., & Hacker, T.J. An exploratory survey on user perceptions and adoption of NEES.org. Proceedings of the 121th ASEE Annual Conference & Exposition. Indianapolis, IN. June 2014

[17] Browning, J., Pujol, S., Eigenmann, R., & Ramirez, J. A. NEEShub Databases. Concrete international, 55., 2013.

[18] Kim, Y., & Crowston, K. Technology adoption and use theory review for studying scientists' continued use of cyber-infrastructure. Proceedings of the American Society for Information Science and Technology, 48(1), 1-10., 2011.

[19] Shin, D. H. A socio-technical framework for cyber-infrastructure design: implication for Korean cyber-infrastructure vision. Technological Forecasting and Social Change, 77(5), 783-795, 2010

[20] Malik, Q. H., Perova, N., Hacker, T. J., Streveler, R. A., Magana, A. J., Vogt, P. L., & Bessenbacher, A. M. Creating a virtual learning community with HUB architecture: CLEERhub as a case study of user adoption. Knowledge Management & E-Learning: An International Journal (KM&EL), 3(4), 665-681, 2011.

[21] Barjak, F., Lane, J., Kertcher, Z., Poschen, M., Procter, R., & Robinson, S. Case studies of e-infrastructure adoption. Social Science Computer Review, 2009

[22] Hacker, T., Magana, A. A Framework for Measuring the Impact and Effectiveness of the NEES Cyberinfrastructure for Earthquake Engineering, unpublished.

[23] Fry, J., & Thelwall, M. Measuring the Impact of e-Research: Accounting for Disciplinary Differential in Patterns of Diffusion. In Proceedings of the Fourth International Conference on e-Social Science, University of Manchester (Vol. 32), 2008.