PROJECT SUMMARY

Overview:

As a NSF Cybersecurity Center of Excellence, the Center for Trustworthy Scientific Cyberinfrastructure (CTSC) will continue to provide the NSF community with leadership and support necessary to tackle the unique cybersecurity challenges of open science. Building on three years of cybersecurity leadership experience with the NSF community, CTSC will leverage its existing knowledge, relationships, and processes to further its work to meet its objective of ensuring trustworthy science in environments that encompass highly valuable scientific instruments, workflows, and datasets accessed by researchers who are geographically and organizationally diverse.

CTSC will support individual NSF cyberinfrastructure projects through collaborative engagements that address specific project needs. CTSC engagement activities include (but are not limited to) security audits, security architecture design and code reviews, and assistance with adoption of best practices. First year engagees will include Gemini, HUBzero, the United States Antarctic Program, and the Authorisation and Authentication for Research Communities project. Additionally CTSC will undertake an innovative collaboration with the proposed Science Gateways Community Institute to reach the large community of science gateway projects. CTSC will continue to innovate with engagements to reach greater number of NSF projects more effectively.

CTSC will perform outreach and dissemination of best practices via blog posts, email lists, and online chats, as well as providing cybersecurity training in person and via online courses. CTSC will provide cybersecurity situational awareness to the NSF cyberinfrastructure community through timely advisories and notices, and, in collaboration with ESnet, CTSC will publish an information security threat model scoped to the particular assets and interests of the open science community. CTSC will continue to organize the annual NSF Cybersecurity Summit for Large Facilities and Cyberinfrastructure, providing the community with the opportunity to share best practices, attend practical training sessions, and collaborate on solving common challenges.

Intellectual Merit :

NSF cyberinfrastructure must support science that is both trustworthy and open to a geographically and organizationally diverse community. Such science has rare, even unique, assets with particular security risks and exists in a world of constantly changing threats. Addressing these security risks coherently, while addressing the tension of maximizing scientific productivity and interoperability, is meritorious applied research with social, managerial, and technical facets.

Broader Impacts :

Through annual summits, high-quality publications, in-person and online training, student internships, and engagements with individual science projects, CTSC will help make NSF science more secure and productive. CTSC will provide significant value to the NSF research community through direct interactions with hundreds of cyberinfrastructure professionals and indirectly through the thousands of scientists who rely on NSF cyberinfrastructure.

CICI: Center of Excellence:
Center for Trustworthy Scientific Cyberinfrastructure

January 1, 2016 - December 31, 2018

Leadership: Von Welch (PI), Dr. James Basney (co-PI), Randal Butler (co-PI), Craig Jackson (co-PI), James Marsteller (co-PI), Prof. Barton Miller (senior personnel)

1. Introduction

Since 2012, the team submitting this proposal has been operating the NSF-funded Center for Trustworthy Scientific Cyberinfrastructure (CTSC). CTSC has been performing the core functions called out by the NSF Cybersecurity Innovation for Cyberinfrastructure solicitation as vital to a Cybersecurity Center of Excellence (CCoE): engaging with 21 NSF CI projects, including 6 Large Facilities, to aid them with their cybersecurity challenges; providing cybersecurity training to nearly 150 NSF CI professionals from over 70 projects; organizing and leading an annual summit building community and momentum in tackling NSF CI cybersecurity challenges; and contributing a section on cybersecurity to the NSF Large Facilities Manual [1] to guide the 28 large facilities representing NSF's largest investments in science.

Only CTSC brings the combination of cybersecurity expertise, history of collaboration with the NSF community, experience in managing the activities required of a CCoE, and recognition as a leader by the community which provide the needed foundation for a CCoE to address the cybersecurity challenges faced by the NSF community in delivering trustworthy science.

CTSC's current funding from NSF (award #1234408) is equivalent to the funding for a CCoE and will expire at the end of 2015. Without the continued funding as a CCoE, CTSC will cease to exist and the NSF community will lose a vital cybersecurity resource, and momentum and productivity will suffer.

Hence, we propose that NSF fund CTSC as a CCoE. This proposal describes why our CTSC team is uniquely qualified to serve as a NSF CCoE and is organized as follows. First, it describes why the NSF community and the challenges such a CCoE will face need a CCoE. Our Prior NSF Support Section then focuses on CTSC, describing our accomplishments in detail and why we are uniquely situated to meet NSF's cybersecurity challenges. The main section of our proposal then describes our future work as a NSF CCoE if so funded, specifically how we will continue and enhance current activities and add new innovations to exceed the requirements of the solicitation. The Key Relationships Section then describes how we have or will establish the relationships to foster cybersecurity interoperability and broadly disseminate results within and outside of the NSF community. The proposal concludes with a Broader Impacts Section and a section describing our Advisory Committee.

2. The Need for a Cybersecurity Center for Excellence (CCoE)

To provide context for this proposal, it is worth exploring why a Cybersecurity Center of Excellence (CCoE) is important to the NSF community. Two example engagements we undertook as CTSC (from the nineteen described in our Prior Work Section), both documented more completely in CTSC's annual reports [2][3], serve to illustrate this value:

> The Daniel K. Inouye Solar Telescope (DKIST) is a NSF-funded Major Research Equipment and Facilities Construction (MREFC) project whose construction began in 2010 and is planned to be completed in 2017. The DKIST will be the largest solar telescope in the world and will be located in Haleakalā, Hawai'i with an operations center in Boulder, Colorado.
>
> In 2013, Bret Goodrich, Senior Software Engineer for DKIST, was charged with developing DKIST's information security plan to comply with the NSF Cooperative Agreement Supplemental Financial & Administrative Terms and Conditions on information security. Seeking guidance on the matter, Bret found little. The NSF Large Facilities Manual at that time [1] contained only a brief paragraph of guidance.

> Fortunately, CTSC restarted the NSF Cybersecurity Summit for Large Facilities and Cyberinfrastructure that year, and Bret attended looking for guidance. There, he met the CTSC team and began planning an engagement. Bret sought expert guidance, but wanted a solid resource that would enable him to make well-informed planning decisions on his own. CTSC recognized that existing cybersecurity resources were likely either too minimalistic or too voluminous and labyrinthine (e.g., the NIST Special Publications) to foist on Bret. CTSC and DKIST embarked on a collaboration not only to solve the immediate challenge facing DKIST, but to develop a set of guidance, templates, and tools for future NSF projects with similar planning needs [4]. The resulting resources' utility was borne out when the NSF-funded Large Synoptic Survey Telescope subsequently used the guide with minimal help from CTSC to design its own cybersecurity program. CTSC is now producing a section based on this guidance for the Large Facility Manual at NSF's request.

A second example comes from the Dr. Ewa Deelman and the Pegasus Workflow Management project that she leads:

> Dr. Deelman and her team were struggling to support scientists with workflows that needed to coordinate multiple resources, e.g. a compute cluster and data storage, and that were so long duration that it was unreasonable to expect a scientist to be present to provide a password whenever the workflow needed to access a resource. She knew that embedding passwords or cryptographic keys in the workflow would greatly weaken security. She also knew that CI software existed to support such workflows securely (e.g. [5]), but not all the administrators of the resources the scientists she served had the flexibility and time to install this CI software.
>
> Dr. Deelman turned to CTSC and the two teams collaborated to understand her science use cases and the security goals. The CTSC team then examined options and, while no "silver bullet" was found, determined a number of practices using the standard Secure Shell software that would allow the Pegasus team to optimize for the combination of science productivity and security for her community [6].

It was stories similar to these two examples that motivated the 35 NSF projects represented at the 2010-2011 Scientific Software Security Innovation Institute workshops [7] to express the community's clear desire for "security leadership and guidance" and "documentation, training, recommendations, and consulting." Attendees at the 2011 workshop particularly emphasized the need for a cybersecurity center that provides leadership by "aggregating community needs and speaking on behalf of the community to external entities" as well as sharing of "successes and lessons learned from projects so that other projects can benefit." Likewise, attendees at the 2014 NSF Cybersecurity Summit for Large Facilities, which we organized and hosted, emphasized the need for "sharing materials, services, policies, practices, lessons learned, and collaborative/peer reviews" [8].

One could argue that projects should be responsible for obtaining and developing their own cybersecurity expertise. However, the workforce for cybersecurity in general is stretched thin [9], and science projects are challenged to find cybersecurity talent, particularly talent that is also familiar with scientific computing. A recent Department of Energy Advanced Scientific Computing Advisory Committee Workforce Subcommittee has documented the challenges in finding qualified workforce across a range of computing skills, including cybersecurity [10].

In addition, experience demonstrates that even when individual projects can find and retain cybersecurity talent, each would only be tackling its slice of the science cybersecurity challenge, in its own manner, leading to a fragmented, often not interoperable, set of solutions, reinvented repeatedly. The complicated NSF CI ecosystem brings significant challenges in cross-project collaborations and knowledge dissemination. Hard fought lessons learned by a project are shared haphazardly between projects, if at all. Additionally, important institutional knowledge is often lost when a project is completed or key personnel leave the community. Taken together, the foregoing factors combine to require each CI project to tackle cybersecurity independently, leading to mistakes being repeated and multiple implementations for such things as authentication systems

that do not interoperate, and confound the goal of scientific collaboration, data stewardship, and dissemination.

As acknowledged by this solicitation, hubs of knowledge, such as CCoEs, are important to the community to effectively advance and capture community knowledge, and to ensure that projects can effectively utilize prior cybersecurity lessons, freeing project personnel to maximize their effort on advancing science.

3. Challenges for a Cybersecurity Center of Excellence (CCoE)

Given the examples in the prior section, it is clear that a CCoE faces challenges beyond just being expert in cybersecurity. Even within the scope of cybersecurity, a NSF CCoE must understand the NSF science that it ultimately seeks to serve. Cybersecurity for science has its own tension to maximizing both the productivity and the trustworthiness of science. Scientific CI must provide science an environment that both supports open access and protects sensitive assets. Scientific CI is a complex weave of high-performance networks, computing services, and storage infrastructure that are often distributed across higher education institutions, research labs, and commercial providers. Specific NSF community challenges with cybersecurity, and how we will or are already addressing them, are described in our supplemental Project Plan.

Beyond the scope of cybersecurity, a NSF CCoE must also have strong relationships with the NSF community so that the community is aware of and trusts its activities are in the best interest of science. The CCoE must also use these relationships, along with skills in dissemination and outreach, to have broad impact, foster a steady stream of engagements, and ensure that engagement results have impact beyond the immediate engaged project.

The CCoE must have expertise in managing multiple simultaneous collaborative engagements. Engaged projects will fact their own internal shifting priorities and goals which make such engagements non-linear, requiring flexibility and persistence to ensure engagements complete and produce valuable results.

4. Results from Prior NSF Support: Center for Trustworthy Scientific Cyberinfrastructure

For the past three years, we have been working together as the Center for Trustworthy Scientific Cyberinfrastructure (CTSC) (#1234408, 9/2012-8/2015, $4,518,845) and accumulating a set of experiences and connections to the NSF community that uniquely prepare us to operate a CICI Cybersecurity Center of Excellence.

Intellectual Merit: Our accomplishments as CTSC include engagements with NSF Large Facilities and CI projects, organizing the 2013, 2014 and 2015 NSF Cybersecurity Summits for Large Facilities and Cyberinfrastructure, providing the community with a guide and templates for developing and maintaining a cybersecurity program, authoring the information security section of a forthcoming NSF Large Facilities Manual, and developing and providing training on topics including secure coding, incident response, and how to plan, establish and operate a cybersecurity program.

Broader Impact: To date, CTSC has engaged with 21 NSF projects (6 being high-investment Large Facilities), and trained over 130 CI professionals representing 30 NSF projects (including 14 Large Facilities). Nearly 150 individuals, representing over 70 NSF-funded projects, have attended one or both of the CTSC-led NSF Cybersecurity Summits. The 2014 Summit was particularly successful in building the community around a call for participation that resulted in the broader community presenting two training sessions and four experience reports. These metrics (projects assisted, individuals trained, and size of the developed community) demonstrate the positive benefit CTSC has already delivered to the NSF community.

Engaged NSF Project(s)	NSF Directorate	Status/Resulting Publications
LTER Network Office (LNO)	Environmental Biology	Engagement complete: Private risk assessment and cybersecurity plan delivered to LNO.
CyberGIS	Advanced Cyberinfrastructure	Engagement complete: Private risk assessment and cybersecurity plan delivered to CyberGIS.
DataONE	Advanced Cyberinfrastructure	Engagement complete: DataONE: Identity Management System Review [11]
DKIST	Astronomical Sciences	Engagement complete: Cybersecurity Program Guide [4]
Gemini Observatory	Astronomical Sciences	Pre-Engagement Analysis to start June 2015.
Globus	Advanced Cyberinfrastructure	Engagement complete: DataONE: Identity Management System Review [12]
HUBzero	Advanced Cyberinfrastructure	Initial Engagement Complete: Private reports and draft policies delivered.
IceCube	Physics	Engagement complete: IceCube Cybersecurity Improvement Plan [13]
LIGO	Physics	Engagement complete: Engagement reports [14][15][16][17][18]
LSST	Astronomical Sciences	Engagement complete: Private LSST cybersecurity plan submitted by LSST to NSF for review.
NEON	Emerging Frontiers	Engagement ongoing.
OOI	Ocean Sciences	Engagement ongoing.
Pegasus	Advanced Cyberinfrastructure	Engagement complete: Pegasus-CTSC Engagement Final Report [6]
Penn State University and University of Utah (CC-NIE peer review)	Advanced Cyberinfrastructure	Engagement complete: Peers delivered reviews by teleconference.
perfSONAR	Advanced Cyberinfrastructure	Engagement ongoing.
SEAD	Advanced Cyberinfrastructure	Engagement complete: Private cybercheckup report delivered to SEAD.
SciGaP	Advanced Cyberinfrastructure	Engagement ongoing: Interim products [19,20]
University of Oklahoma (CC-NIE)	Advanced Cyberinfrastructure	Engagement ongoing.
University of Pittsburgh and University of Cincinnati (CC-NIE peer review)	Advanced Cyberinfrastructure	Engagement to start July 2015.

Table 1: Complete list of 19 CTSC engagements with 21 NSF projects under current funding.

Other publications from the CTSC award, not listed above, are: [2,3,6,8,9,34,36,39].

5. Other Experience of our Team

Our success with CTSC is due to the combined expertise of its members, not only in technical aspects of cybersecurity, software development, and national scale CI development, but also with strong experience in engagement with the NSF science and CI communities. The leaders of CTSC, and their relevant proposal history outside of CTSC, are:

Von Welch is the director of Indiana University's Center for Applied Cybersecurity Research (CACR) and serves as the Director and PI of CTSC. Prior NSF PI experience include server as co-PI, until 2007, on the SCI TeraGrid Resource Provider project (#0504064, 8/2005-12/2012, PI:

John Towns, $35,067,160) which provided high-quality HPC resources and associated services to the scientific community, serving as senior personnel in the TeraGrid Grid Integration Group and publications reflect both of those roles [21, 22, 23-26], PI of SCI: Collaborative Research: NMI DEVELOPMENT: Policy Controlled Attribute Framework (#0438424; 12/2004-12/207; $396,060) [27, 28], co-PI of EAGER: Best Practices and Models for Sustainability for Robust Cyberinfrastructure Software (#1147606; 9/2011-8/2013; PI: Dr. Craig Stewart; $296,637) [19], PI of TWC: Medium: Collaborative: Foundations of Application-Sensitive Access Control Evaluation (#1228668; 9/2012-8/2015; $201,779), PI of A Cyber Identity Infrastructure for National Science (#0943633; 9/2009-8/2012, transferred to Dr. James Basney when Mr. Welch moved to Indiana U.; $1,757,640), co-PI of CILogon: Secure Access to National-Scale Cyberinfrastructure (#0850557, 6/2009-5/2011; PI: Dr. James Basney; $399,837) [29,30], and PI of Security at the Cyber-Border (#1158796; 12/2011-11/2012; $32,623) [31]. Recent PI experience for Mr. Welch outside of NSF includes serving as co-PI of the DHS Software Assurance Marketplace [32] and the DOE-funded XSIM project looking at identity management for virtual organizations [33] Mr. Welch also serves on the InCommon Steering Committee as an advisor for the research community.

Dr. Jim Basney is a senior research scientist in NCSA's Cybersecurity Directorate. Jim leads the CILogon project, which enables federated authentication to cyberinfrastructure [34]. Jim is also the security technical lead for XSEDE Software Development and Integration, and the identity management lead for the DHS Software Assurance Marketplace. Jim is an active participant in The Americas Grid Policy Management Authority and the InCommon Technical Advisory Committee. Jim is co-PI on CTSC (#1234408, 8/2012-9/2015, PI: Von Welch, $4,518,845). Additional PI and co-PI roles include: CC*IIE IAM: FeduShare: Bridging Campus and Research Identity and Access Management for Self-Managed Collaboration (#1440609, 10/2014-9/2015, PI: Jill Gemmill, $291,040), SDCI Sec: Distributed Web Security for Science Gateways (#1127210, 8/2011-7/2015, PI: Basney, $948,821), Scientific Software Security Innovation Institute (#1043843, 8/2010-7/2012, PI: Randal Butler, $49,862), A Cyber Identity Infrastructure for National Science (#0943633, 9/2009-8/2012, PI: Basney, $1,757,640), CILogon: Secure Access to National-Scale CyberInfrastructure (#0850557, 6/2009-5/2011, PI: Basney, $399,837), and Integration of the MyProxy Online Credential Repository into the NSF Middleware Initiative Software Infrastructure (#0222571, 7/2002-6/2005, PI: Basney, $598,343).

Randal Butler is director of the NCSA's Cybersecurity Directorate, Chief Security Officer for NCSA. Prior NSF PI experience includes serving as PI on Scientific Software Security Innovation Institute (S3I2) (#104383, 8/2010-7/2012, $49,862), a workshop series that addressed the potential benefits of a security focused software institute that would serve the entire NSF research and development community [35], PI on SDCI: NMI-NEW: Observatory Middleware Framework (#0721617, 1/2007-8/2011, $499,025), co-PI on A Cyber Identity Infrastructure for National Science (#0943633, 9/2009-8/2012, PI: Dr. James Basney, $1,757,640) and co-PI CILogon: Secure Access to National-Scale CyberInfrastructure (#0850557, 6/2009-5/2011; PI: Dr. James Basney; $399,837).

Craig Jackson is Senior Policy Analyst at Indiana University's Center for Applied Cybersecurity Research (CACR). He has led engagements, co-organized the 2013 and 2014 Summits, co-authored of the *Guide to Developing Cybersecurity Programs for NSF Science and Engineering Projects*, and has served as CTSC's project manager. Jackson is policy lead of the security team for the DHS-funded Software Assurance Marketplace (SWAMP); and he is part of the DOE-funded XSIM (Extreme Scale Identity Management) project. He is a graduate of the IU Maurer School of Law (J.D.'10) and IU School of Education (M.S.'04), and member of the Indiana bar.

James Marsteller is the Pittsburgh Supercomputer Center Chief Information Security Officer (CISO) and is responsible for securing PSC systems and operations. He has extensive security leadership experience with the TeraGrid and XSEDE security operations team and co-leads the XSEDE Security and Incident Response teams.

Prof. Barton Miller is a co-investigator of SDCI In-Depth Vulnerability Assessment of Middleware (#1032341, 1/2011-12/2015, $775,000). The University of Wisconsin is currently funded to

support the development of new techniques for performing in-depth vulnerability assessments based on their First Principles Vulnerability Assessment (FPVA) methodology. This project has included extensions to their FPVA methodology; the design, development and demonstration of tools to automate parts of the assessment process, and applications of these techniques to critical Grid middleware and commercial software (such as Google Chrome). They are currently in their last year of this effort.

6. Our Proposed Cybersecurity Center of Excellence

Our existing Center for Trustworthy Scientific Cyberinfrastructure (CTSC) has given us nearly three years of experience delivering the services called out in the solicitation and developing the required skills, community connections, and track record of leadership to be a Cybersecurity Center of Excellence. Hence we propose CTSC, with its existing expert team, established community relationships, and proven processes, as a NSF Cybersecurity Center of Excellence.

CTSC is already routinely performing the activities identified in the CICI solicitation, namely: providing leadership to the NSF research community, conducting security audits and architecture design reviews, ensuring adoption of security best practices in the NSF research community, advancing the security acumen of NSF CI project staff, and hosting an annual workshop in the form of the NSF Cybersecurity Summit for Large Facilities and Cyberinfrastructure. We have experimented with notifications to the community, which we will mature into a situational awareness service. We will similarly mature our work on secure coding (and relationships with the DHS Software Assurance Marketplace) in a Software Assurance strategy. We have not to date developed a threat model, but applaud it as something our work has determined is useful in the form of a recommendation from the 2014 Cybersecurity Summit [39].

The following sections describe our accomplishments, processes, and innovations for each of these activities from the past three years, as well as our plans for improving them as we evolve into a Cybersecurity Center of Excellence. Sections on Threat Modeling, Situational Awareness, and Software Assurance describe our plans to address these elements in response to the solicitation. To measure our overall effectiveness over time, we will conduct an annual community survey. Our supplement project plan describes milestones and additional metrics for each activity.

The funding for a CCoE is equivalent to the current spending rate for CTSC (the total budget for a CCoE is approximately 10% higher, but we increased spending over three years). After the first year where additional effort is needed to ramp up the new activities required by the solicitation, we believe we can continue our current activities as well as the new ones through increased efficiencies and innovation as we describe subsequently.

6.1 Engagements: Audits, Reviews, and Cybersecurity Program Development

The CICI solicitation calls for a CCoE to conduct security audits and design reviews. We have been conducting such activities for the past three years as CTSC “engagements”. Engagements are collaborations between CTSC and a NSF project, focused on addressing a particular cybersecurity challenge of that project. Possible engagement foci include audits and reviews, but are not limited to such and also include: developing a risk-based cybersecurity plan (DKIST, LSST, NEO), reviewing existing plans (HUBzero, IceCube, LNO, CC-NIE projects, PerfSonar), making recommendations on software security features (Pegasus, SciGaP), and reviewing software at the code (PerfSonar) or architectural level (Globus). The two stories in Section 2 are examples of such engagements. Descriptions of all of our engagements can be found in Section 4 on our Prior Work as well in our annual reports [1][2] (the reports also include statements from the engaged projects on impact of the engagements).

We have found that these engagements are non-trivial undertakings, requiring acumen beyond cybersecurity. Their exact format needs to vary depending on the projects culture, goals, challenge, and lifecycle stage. We have learned to innovate and be flexible with regard to engagements, experimenting to find ways of making them more efficient and with broader impact. Examples of this innovation include: the introduction of a brief “cybercheckup” at the start of an engagement to evaluate a project’s existing cybersecurity program and identity which aspects

would benefit most from attention; experimenting with a peer-to-peer review with projects reviewing each other that ultimately will allow for scale beyond what is possible for CTSC to do as an entity; and ongoing low-effort consulting for newly starting projects which are making a series of design decisions for which they need quick feedback with regards to cybersecurity.

Engagements are critical but also labor intensive, consuming the majority of CTSC's resources. Hence we will continue to mature our engagement methodologies to allow us to perform them more efficiently. For example, the CTSC-developed *Guide to Developing Cybersecurity Programs for NSF Science and Engineering Project*s [36], including over 16 related templates, tools, and resources, supports NSF CI projects in efficiently building a cybersecurity program to comply with the NSF Cooperative Terms and Conditions for MREFCs [37] . We developed it as part of our engagement with DKIST, and it was subsequently demonstrated by LSST as an effective tool for developing a cybersecurity program.

We have also, in consultation with NSF, engaged outside of the NSF community when it is of benefit to the NSF community. The Network Time Protocol project is a software component critical to the security of nearly all NSF projects and of broad impact to the Internet as a whole. As the reference implementation of Network Time Protocol, the NTP software project is the most widely-deployed time synchronization tool in the world: it provides accurate timing for scientific sensor readings, the stock market, medical and military systems, as well as infrastructure tools such as PerfSonar appliances, HUBzero servers, and Bro project installations. When we learned that NTP was dealing not only with security issues in their software but also with personnel and resource issues that prevented them from addressing those security issues, we stepped up and begin working with existing contributors to implement rigorous software assurance processes, refactor vulnerable code, and implement a transition plan in place to ensure the long-term viability of the software. This experiment is ongoing, but if successful indicates a path by which we can selectively engage critical software projects that lack access to cybersecurity resources and broaden our impact far beyond the NSF ecosystem. We anticipate such engagements being rare, but we are willing to undertake them in consultation with NSF when the benefit is warranted.

As a CCoE we will continue to perform these engagements and continue to advance our innovations. Two new innovations we propose to undertake as a CCoE are working collaboratively with NSF Software Infrastructure for Sustained Innovation (S2I2) institutes and tackling regulated data (e.g. HIPAA and FISMA).

Three initial S2I2 institutes are expected to be funded approximately the same time as a CICI CCoE [38] and will reach broad sectors of the NSF community. It is only natural that we collaborate with them to help with cybersecurity challenges in their sectors. CTSC expects to actively engage with the institutes to assist them in securing their own cyberinfrastructure, securely designing and reviewing their software development and integration processes, and providing security guidance to their communities. As part of our software assurance strategy, we will work with the institutes on a process for software vulnerability management for the software products they distribute. We already have an agreement in place to work collaboratively, through a jointly funded half-time security analyst, on cybersecurity issues in the context of science gateways with the proposed Science Gateway Community Institute under the leadership of PI Wilkins-Diehr and as described in her letter of support.

We have been seeing interest in regulated secure environments such as HIPAA and FISMA compliance among NSF projects as they are increasingly working with human subjects data and with federal facilities (e.g. NSF's Polar Program). The project team is actively developing experience with these compliance regulations (e.g. PI Welch oversees HIPAA and FISMA planning for IU's central IT unit) and we will bring this experience to bear in engagements when called for, and then subsequently include those experiences in the cybersecurity planning guide as well as working with the Coalition for Academic Scientific Computing (CASC) Regulated Data Working Group when appropriate.

We have the following projects committed as engagements in our first year as a CCoE:

- HUBzero: We previous engaged with HUBzero to evaluate their cybersecurity program and address some key issues. They solicited a more complete engagement with us to perform a full risk assessment of their complicated CI, in terms of its development and operations supporting dozens of projects.
- The Gemini Observatory: We will engage with the Gemini project to review their cybersecurity program and help address any issues.
- The United States Antarctic Program (USAP): The USAP operates CI at the South Pole for a variety of NSF projects and is seeking guidance on their cybersecurity. Our engagement is limited by lobbying policy dictating what we can do as a federally funded project working with a federal agency, but we will provide them with observations on their cybersecurity program.
- The EU-funded Authentication and Authorisation for Research and Collaboration (AARC) Horizon 2020 project: Recognizing that scientific collaborations are often international, our engagement with the GÉANT Authentication and Authorisation for Research and Collaboration (AARC) project enables EU-US coordination on federated identities for international science. The CTSC-AARC engagement will develop training, facilitate pilot projects, document requirements, and enable coordination with other InCommon and GÉANT interfederation efforts.

Historically we have averaged seven engagements per year, but we commit to fewer in year one as a CCoE to allow us to initiate the Threat Modeling and Situational Awareness activities required by the solicitation, and begin our collaboration with the S2I2 Science Gateway Community Institute. After year one, we plan to return to our current rate of seven solicitations per year, and attempt to increase that steadily through process innovation. We will solicit new engagements through the Cybersecurity Summit (see Section 6.2), attendance at the NSF Large Facilities workshop, advertisement on social media, presentations to the community, and word of mouth amongst NSF projects. Through these means we have an established track record of projects seeking to engage with CTSC and foresee no issue in continuing this rate of engagements for another three years as a CCoE.

6.2 NSF Cybersecurity Summits for Large Facilities and Cyberinfrastructure

In 2013, we re-launched the annual NSF cybersecurity summits after a five year hiatus and have continued to organize successful summits for the CI community on a yearly basis, introducing a highly successful Call for Participation process in 2014 to facilitate greater community involvement with the event To address the solicitation's requirement for an annual workshop, we will continue its community leadership role in organizing annual Cybersecurity summits as a CCoE.

The annual cybersecurity summits provide the community with a valuable opportunity to share best practices, attend practical training sessions, and collaborate on solving common challenges with regard to securing NSF-funded facilities and projects. Community evaluations have been overwhelmingly positive.

In line with developing a Threat Model, the 2015 summit will have the theme of "Understanding Information Assets that Enable Science" and the 2016 summit will have the theme of "Threat Models for Open Science." A call for participation will seek program content from the NSF community, industry, government and academia. The 2015 and future summit reports will be published and freely available as were those from 2013 [39] and 2014 [8].

The summits have become an increasingly important event for us. The interaction with the community helps CTSC to make new relationships and cultivate opportunities for collaboration with large facilities and projects. Additionally, the knowledge gained from past engagements is communicated at the annual summit amplifying the impact of the work to the larger community.

We will hold the Summit proximate to NSF to encourage presentation and participation by NSF program officers, something we have found to be critical in developing a sharing understanding between the community and NSF. We will follow NSF's move of its offices in 2017.

6.3 Outreach and Dissemination of Best Practices

It is our goal to encourage and assist the NSF research community in the adoption of security best practices as they are identified. The complicated NSF CI ecosystem brings significant challenges in cross-project collaborations and knowledge dissemination and more often than not security expertise is not present within the projects. Hard fought lessons learned within a project are shared haphazardly between projects, if at all. Additionally, important institutional knowledge is often lost when a project is completed or key personnel leave the community, necessitating that each CI project must tackle cybersecurity independently. This individualized and ad hoc method is inefficient, redundant, and distracting to the science mission. It leads to multiple implementations for things such as authentication systems that do not interoperate and confound the goal of scientific collaboration, data stewardship, and dissemination. It leads to common mistakes being replicated across the entire NSF CI community, leaving CI vulnerable.

As a CCoE, we will continue to undertake outreach activities both to disseminate our work and to advertise our services to NSF CI projects. We were highlighted, in collaboration with LIGO, in *International Science Grid This Week* [40] working with Internet2 on international federation - rare prominence for a cybersecurity project. Other outreach mechanisms include the CTSC website (trustedci.org), with an ongoing blog and twitter feed, covering our activities and cybersecurity news of interest to NSF CI projects. We also use the blog to disseminate best practices providing guidance and how-to guides that are broadly applicable to NSF CI, e.g. our series on identity management.

We will maintain a set of email lists [42] to foster collaboration and allow dissemination to the community. Topics for the email lists are: vulnerability announcements for software development projects, vulnerability announcements for infrastructure operations projects, federated identity, and a general discussion list. The lists currently have 24, 38, 15, and 48 subscribers respectively and are open to the public.

We are not yet satisfied with the impact of these lists and we continue to publicize them and build the trust of the community. We will publicize, and hold live open online chats featuring CTSC staff and special guests from the field monthly, similar to Internet2's IAM Online series. Each month of the calendar year will have a specific security theme relevant to the NSF CI community, CTSC staff members and guest speakers participating in these live online chat sessions will speak to topics tied specifically to the month's theme. All sessions will be captured and available to replay on the CTSC TrustedCI website. Each session will be advertised through the CTSC mailing list, the CTSC web site, and all social media outlets associated with CTSC. Assigning themes to each month in the calendar year will allow for the streamlining of all outreach efforts. Our initial three themes will build on our existing training products: building a cybersecurity program, incident response and software security. After those three months, we can choose more specific topics (e.g. risk assessment, code review) or target different populations and skill levels with the same topics (e.g. what a PI or manager may need to know about a topic).

We will also continue our in-person outreach to the community. We have presented at NSF PI meetings (SI2, CC-NIE), NSF project meetings, CASC and Internet2 meetings. We also regularly attend the NSF Large Facilities Workshop to both advertise and stay attuned to that community.

One means of outreach we will continue to strive for is having NSF solicitations direct the community to our efforts. One such success for CTSC came with the NSF Natural Hazards Engineering Research Infrastructure solicitation [41], which directed the awardees to attend the Cybersecurity Summit.

6.4 Training

A key component of our goal to achieve a more trustworthy NSF scientific CI and scale to reach the entire community is the development of new cybersecurity expertise through the creation, dissemination, and delivery of training and educational materials. As a CCoE, we will continue develop and offer training in two general themes: software assurance, and general open science cybersecurity planning and operations.

Software Assurance Training: Training topics will focus on software practitioners and managers, as well as spin-offs to create a software assurance course for advanced undergraduate and beginning graduate students. The training topics to be covered over the course of this project will include mobile code security, extended web security, and executive briefs for management around infrastructure projects and software development projects. All training will be developed initially for live presentation. We will continue our successful activity of teaching at workshops, conferences, companies, universities and research labs. We are aware of the instructional materials available associated with the CMU Software Engineering Institute's Software Assurance Curriculum. While our secure coding materials are more extensive than theirs, we hope to leverage materials that they have prepared on regulatory issues and executive-level briefings. Over the course of this project, we will take all of our existing and new training materials and record them in a form to be delivered online.

We will also develop a Massive Open Online Course (MOOC) covering the topics of secure coding, vulnerability assessment, analysis tools, and managing that assessment process. The development of the MOOC will include sharing materials with those under development for our course for advanced undergraduates and graduate students. There has been notable success by the CTSC team in presenting this material live in the past, and our goal is to broaden our reach and impact in this area.

General Open Science Cybersecurity: We have had great success to date with our two topic areas including “Developing a Cybersecurity Program”, and “Incident Response Training”, having offered those at the NSF Cybersecurity Summit as well through YouTube video clips. Those topics continue to evolve and we will update each of them as needed throughout the lifetime of CTSC. In addition we plan to develop three additional training programs on the topics of: 1) Security Analysis - will cover the entire workflow from monitoring and collection of intelligence, log management, analysis, and action, 2) Identity and Access Management - covering Identity and Group Management service options, configuration and setup, and integration from a project and national perspective, and finally 3) Software Development Security Best Practices - which will focus on the software development environment and best practice around facilities, software development, patch management, testing, and evaluation of vulnerabilities and patches. These courses will be developed at the rate of one per year. They will first appear at the annual NSF Cybersecurity Summit and then be followed by online “webinars” for wider outreach.

Evaluation and feedback: Feedback on training content and delivery mechanics from training audiences, as well as input from our advisory board, plays a critical role in the development of future materials, selection of topics, and in the use of tooling for their delivery. Every in-person and online training event includes an evaluation and assessment of the content and the delivery tooling, infrastructure, and experience (see our Summit Reports [39][8]).

6.5 Providing Cybersecurity Situational Awareness

As a CCoE, we will formalize the community notification process it has already begun under CTSC and provide a Cybersecurity Situational Awareness service for the NSF community as required by the solicitation. Situational awareness in a cybersecurity context is a critical operational awareness of potential threats and vulnerabilities that allow for proper remediation. Situational awareness is specific to a particular community’s IT infrastructure and assets. For example, 19 Incident Sharing and Analysis Centers (ISACs) exist to cover different national sectors. Situational awareness provided by a NSF CCoE needs to address threats and vulnerabilities that are relevant to NSF CI and in a manner appropriate to the NSF community.

Software vulnerabilities are a major enabler of cyber attacks and a pervasive problem. Software used by NSF CI is a complex ecosystem of NSF-developed CI, open source software and some commercial software. A NSF CCoE needs to intelligently filter the constant stream of new software vulnerabilities that arise and determine which ones are relevant to the NSF community. For example, most of the CI that CTSC has encountered in its engagements runs on Linux operating systems. Therefore, sending out numerous notices about Windows operating system vulnerabilities would be counterproductive - the community would begin to treat *all* notices as noise and ignore them. In addition to software vulnerabilities, there are other threats such as ongoing social engineering and distributed denial of service attacks that may be relevant to the community. Assisting with preventive cybersecurity, by making the community aware of new threats or vulnerabilities and providing guidance to reduce risk through actions to mitigate those threats, is the goal for cybersecurity situational awareness.

It is also critical for notifications to be clear and simple to understand, especially for the NSF community where not all projects have dedicated cybersecurity staff, . Community members need to be able to quickly ascertain if notifications are relevant to their project and, if so, how to apply a remediation.

CTSC has already experimented with filtering and announcing such vulnerabilities to the community. Over a nine-month period in 2014-15 CTSC sent out notices for seven vulnerabilities: OpenSSL Heartbleed (CVE-2014-0160) [43], OpenSSL CCS injection (CVE-2014-0224) [44], POODLE SSLv3 (CVE-2014-3566) [45], Drupal 7.x SQL injection (CVE-2014-3704) [46], Git/Mercurial (CVE-2014-9390) [47], GHOST/glibc (CVE-2015-0235) [48], and HTTPS spoofing (CVE-2015-2078) [49].

As a NSF CCoE, CTSC will mature this into a Situational Awareness service that the community can count on for high quality, easy to follow notifications on relevant vulnerabilities and threats. CTSC will track notifications from https://nvd.nist.gov/ and https://www.us-cert.gov/ncas/current-activity, and leverage our relationships with the NSF Supercomputing Centers (NCSA, PSC, and other XSEDE Service Providers). We will filter issues for those relevant to the community and then supply simple guidance to go with those notifications. CTSC will utilize its existing email lists [42] and encourage a dialog among those receiving the notifications for further discussions and feedback. All notices will be archived and searchable from the CTSC email archives.

CTSC will continue to innovate in this area based on feedback from the community and direction from the Threat Model once complete (see Section 6.6). We will also continue to pursue leveraging the REN-ISAC. The REN-ISAC supplies situational awareness to the higher education community. Currently their membership requirements prohibit NSF CI projects from joining except for those with well-established cybersecurity programs and identified cybersecurity officers. For long-term sustainability beyond a NSF CCoE, we are working with the REN-ISAC and their membership committee to establish a membership category suitable for all CI projects. We will evaluate our progress and effectiveness at six-month intervals.

6.6 A Threat Model for Open Science

As a CCoE we will produce, in collaboration with the Department of Energy's ESnet, an information security threat model scoped to the particular assets and interests of the open science community, including NSF. The model will address one of the most challenging aspects of conducting useful risk assessments and allocating limited resources to information security: understanding the nature and likelihood of threats to the organization's information assets.

An organization may be very familiar with its assets, however, it takes a great deal of expertise to build an understanding of varied threats to those assets and the frequency with which the particular organization should expect to face them. There exist some tools to assist with the difficulty of threat identification. In general, these are either terribly exhaustive (e.g. [50]) or very concise (e.g. [51]) *general* lists of possible threats. However, these tools do little to help specific organizations select the relevant threats, and nothing to gauge their likelihood/frequency.

Alternative products-- referred to as *threat models* or *threat profiles* -- produce tailored characterizations of feasible threats and anticipated frequencies to specific technologies,

software, or organizations. These tailored products have the benefit of cutting through the noise, but have the same limitation of any risk assessment tailored to particular asset or organization: They take a great deal of effort and expertise to produce. Thus, highly specific threat profiles -- like bespoke clothing -- are expensive to produce and their application is limited.

The threat model for open science will serve as guidance to risk assessments and risk-based cybersecurity planning as implemented and taught by CTSC. Hence we need it to be practical and strike a balance between the highly tailored threat models or profiles traditionally produced for specific organizations or technologies, and broader applicability found in general resources. This community-oriented deliverable will not simply be an organized list of threats, but an easy-to-use tool designed to facilitate efficient identification of feasible information security threats to open science information assets and a baseline likelihood/frequency of each threat's materialization in a project or facilities environment.

Our collaboration with ESnet is uniquely positioned to develop this model for open science, with both CTSC and ESnet having unique vantage points across their respective broad communities. And a model developed in collaboration, covering the NSF and DOE open science communities, will serve to foster interoperability between NSF and DOE open science. To ensure the threat model has maximum benefit to the community, we will integrate directions on using it alongside our *Guide to Developing Cybersecurity Programs* (see Section 6.1), Cybersecurity Program Tutorial (see Section 6.1) and future versions of the Large Facilities Manual information security section. It will be used directly by CTSC in our relevant engagements to provide ongoing evaluation and refinement.

To develop the model, we will convene a working group of CTSC and ESnet experts, plus select community members. We will develop the model over the first year of the CCoE project, with an early draft presented at a working session at the 2016 Cybersecurity Summit to solicit community input. After publication, CTSC will maintain the document, serving as the ongoing editor to make adjustments based on changes in the NSF and DOE environment discovered through our engagements or reported by the community. We plan to update the document on an annual basis, using a working session at each year’s summit for community input. Significant changes will be circulated to the community for feedback before publication. As described in our Data Management Plan, the Model will be published in Indiana University’s ScholarWorks system under open licensing to ensure it persists and can be sustained by any subsequent CCoE or member of the community.

6.7 Software Assurance

Software assurance can be viewed as the natural transition from a system-level risk assessment to a more detailed analysis of the software components that might be at risk. While the goal of a risk assessment is to identify and estimate the likelihood and impact of risks, software assurance attempts to identify actual weaknesses in the code, the vulnerabilities to which they might lead, and direction on how to remediate these vulnerabilities.

We will strive to motivate the adoption of software assurance practices into the design, development, and deployment of software systems used by NSF CI. In both the research and industrial communities, there remains a distressing lack of awareness of the need for such practices. In fact, we note that NSF itself has no such requirements on grantees who produce software that will be deployed as services or infrastructure.

Hence our goals in regards to software assurance are to educate the community and raise awareness for the need for software assurance activities, to provide the resources for groups to get started in such activities, and to help groups that need and desire higher levels of assurance.

To raise awareness we will leverage our Threat Model activity. We will identify threats related to software and from those threats, produce leadership briefings to articulate the risks to the NSF community, and convey how software assurance serves to mitigate those risks. From that we will work with the NSF community, particularly the Large Facilities, to drive consensus on expectations around software assurance - i.e., what software assurance processes and metrics

are expected for software in different contexts. We will then incorporate those expectations into our engagements and outreach with the software development community.

Our other software assurance efforts will focus on enabling software developers to implement good software development practices. We have already developed a detailed tutorial on the use of software assurance tools, including background into the capabilities and technologies in these tools, use of specific tools, and how to use these tools in the SWAMP. We will use this tutorial as a resource in our engagements under the CCoE.

Adding the use of these tools to the software development process is a small increment for a software team, though applying such tools to legacy code takes a bit of guidance and discipline. We will add the introduction of such tools to our risk assessment engagements. To help in this effort, we will leverage the DHS-funded Software Assurance Marketplace (SWAMP), which provides an open and free facility for applying a wide variety of both open source and commercial static analysis tools to software.

For software projects that are motivated to increase their level of software assurance, we will provide the additional service during our engagements of analyst assistance in interpreting the results from the suite of analysis tools. This activity provides the opportunity for increasing the understanding of how to interpret the results, how to determine when the results indicate serious vulnerabilities, and how to approach the remediation of these vulnerabilities. It also serves as an in-situ training activity, teaching the software developers how to more effectively use and respond to the software analysis tools.

For the most critical software, a more extensive analysis-driven assessment process is necessary. Such an activity is, by necessity, more time consuming and has the potential to find a greater spectrum of vulnerabilities in software. Over the period of this grant, we will select key software development projects and provide them the training (see Section 6.4) and assistance to conduct an in-depth assessment of their software, based on our First Principles Vulnerability Assessment methodology [52].

6.8 Annual Community Survey

As part of our maturation to a CCoE, we will augment the community input we obtain through the annual Cybersecurity Summit, our training evaluation and our engagements with a survey we distribute to the NSF CI community. The survey will ask for the project's current top cybersecurity concerns and also ask them to report on the current level of implementation of a cybersecurity program, and qualitatively gauge their level of comfort with that program. This survey will not only provide us with more data in shaping our efforts, such as selecting training topics, but start building a corpus of data on the maturity of the community with regard to cybersecurity.

7. Key Relationships

A crucial goal of cybersecurity is establishing trust and interoperability not only within the NSF community but also with collaborating communities. We seek both to leverage best practices from the broader community as well as to disseminate innovations from CTSC and the NSF community. Hence, we have already established relationships with important communities outside of our key NSF constituency to ensure the success of CTSC as a CCoE:

- Department of Energy (DOE): We are well connected with the open science community in DOE through the Energy Science network (ESnet). We already collaborate with them on projects with a science DMZ component. Greg Bell, head of the DOE, will join Nick Multari from PPNL our advisory committee, giving us broad visibility in DOE cybersecurity activities. Additionally, we are directly collaborating with ESnet on the Threat Model.
- Bro Center of Expertise: As the other large cybersecurity-related project funded by NSF ACI, collaboration between CTSC and the Bro Center is natural. With shared staff (Slagell is a co-PI of the Bro Center), we will continue collaborating on training at the NSF Cybersecurity Summits, engaging with NSF communities with large networking or Bro components, and developing documentation and materials where it makes sense.

- Higher Education, Internet2 and InCommon: The NSF Campus Cyberinfrastructure programs [53] demonstrate the continued growth in the portfolio of research support services on campuses, together with the importance of securely connecting campus CI with regional, national, and international CI. Internet2 and InCommon provide core research network and identity services to campuses and bring the community together to establish standards and share lessons learned. CTSC's representation within InCommon leadership (Welch on the InCommon Steering Committee, Basney on the InCommon Technical Advisory Committee) helps to ensure CTSC's continued positive impact in this community. Also, Tom Barton's participation on the CTSC advisory committee provides an additional valuable campus connection. We will continue to build on prior work (e.g., [54]) to make InCommon useful to NSF CI.
- Open Source Communities: NSF infrastructure and science projects rely heavily on the stability and security of countless open source software components. By collaborating with the open source community, especially through our relationship with the Internet Civil Engineering Institute (for which CTSC analyst Susan Sons recently took over as Director), we are increasing our awareness of security issues in open source components. These collaborations also provide opportunities for us to support improvements in open source software components when there is a clear need to make those components more securely and reliably support the NSF mission.
- NSF: We have learned there is great value in engaging directly with NSF directorates. We are working closely with the Large Facilities Office to produce cybersecurity guidelines for a future revision of the Large Facilities Manual (LFM) that is influenced by CTSC's cybersecurity planning guide. We will also engage directly with the United States Antarctic Program, which directly operates CI at the Antarctic. Perhaps the broadest level of engagement with NSF comes from the annual cybersecurity summit, where program officers from many different NSF Offices and Directorates participate and interact with members of the NSF community, industry, government and academia to better understand the community's cybersecurity challenges.
- International Science: Our engagement with the GÉANT Authentication and Authorisation for Research and Collaboration (AARC) project enables EU-US coordination on federated identities for international science. Neil Chue Hong, director of the UK Software Sustainability Institute, serves on our advisory committee, giving us a persistent liaison to science outside of the U.S.
- Other federal agencies and the commercial sector: Through a small number of selected invitations to the annual Cybersecurity Summit, we will maintain awareness of other federal agencies and activities in the private sector, as well as allow for the dissemination of our work. Several members of CTSC management serve as PIs on DOE and DHS projects.

8. Broader Impacts of the Proposed Work

As a CCoE, we will have numerous broader impacts. The annual Cybersecurity Summit brings together over a hundred members of the NSF community across dozens of projects each year to learn and share experiences. The formal training at the Summit (and in other venues throughout the year) impacts dozens of attendees and their respective projects in turn and the call for participation brings about training and sharing of experiences from the participants.

We conduct workforce development not only through training and engagement of the professionals in the NSF CI community, but also through engagement with students. We have a student intern who works under supervision of one of our staff members on our engagements. We also offer four student scholarships to the NSF Summit each year to expose and interest students in NSF cybersecurity challenges.

As described in our Data Management Plan, all of CTSC produced materials are freely available to the public. This enables our publication of engagement results to have impact beyond both the scope of those engagements and the lifetime of our efforts.

We recognize the lack of diversity as a problem in the cybersecurity workforce in general, and within NSF there is no exception. Starting in 2014, we made a concerted effort to practice inclusiveness in the invitations for the Summit program committee and speakers. With one-third of the presenters in 2014 being female, we believe we did well, but will continue to keep this a point of emphasis. All three of our student interns are female or members of other minority groups.

Finally, as described in our Key Relationships Section, we have strong connections outside of the NSF community, including DOE, DHS, Internet2, InCommon, and higher education, to disseminate our results broadly outside of NSF.

9. Advisory committee

As a CCoE, CTSC will continue to have an advisory committee to guide our strategy and keep us informed of opportunities and activities in the NSF and broader community. The advisory committee will meet once per year. The following community leaders have committed to serve (all but Greg Bell already serve on CTSC's advisory committee):

Tom Barton is senior director for architecture, integration and chief information security officer at the University of Chicago. He has strong ties to Internet2 and InCommon as the lead for the Internet2 Middleware Initiative's Grouper project. He is a member of the Middleware Architecture Committee for Education, the InCommon Federation's Technical Advisory Committee, and EDUCAUSE's Identity Management Working Group.

Gregory Bell is the division director for ESnet and holds experience in advanced networking, collaborative tools, sustainable IT, cloud computing services, high-performance computing, and security models for open science. Bell is founding chair of the Cyberinfrastructure Advisory Committee for the Deep Underground Science and Engineering Lab (DUSEL).

Neil Chue Hong is director of the Software Sustainability Institute (SSI) and is responsible for representing the SSI and UK researchers' software interests nationally and internationally. Within the organization, he oversees SSI operations, leads the community engagement, develops and manages collaborations, and acts as the principal liaison with stakeholders.

Don E. Middleton leads the Visualization and Enabling Technologies Section in NCAR's Computational and Information Systems Laboratory. Don currently serves as PI or co-PI on a number of projects, including the Earth System Grid, the Earth System Curator, the Virtual Solar Terrestrial Observatory, the North American Regional Climate Change Assessment Program, the Cooperative Arctic Data and Information Service, and NCAR's Cyberinfrastructure Strategic Initiative.

Nick Multari is the Senior Project Manager in Cybersecurity Research at Pacific Northwest National Lab (PNNL) in Richland, Washington. He has expertise in establishing the direction and leading the execution of various research projects resulting in a rigorous foundation upon which security concepts are matured and implemented.

Nancy Wilkins-Diehr of the San Diego Supercomputing Center has a breadth of experience in community engagement as co-lead for the TeraGrid/XSEDE Science Gateways and Extended Collaborative Support Service programs.

References Cited

[1] National Science Foundation. (2013, Jan.) *NSF Large Facilities Manual* [Online]. Available: http://www.nsf.gov/pubs/2013/nsf13038/nsf13038new.pdf.

[2] V. Welch. (2013). *Year 1 Report: Center for Trustworthy Scientific Cyberinfrastructure* [Online]. Available: http://hdl.handle.net/2022/17205.

[3] V.Welch. (2014, Sept.). *Year Two Report: Center for Trustworthy Scientific Cyberinfrastructure* [Online]. Available: http://hdl.handle.net/2022/20028.

[4] Center for Trustworthy Scientific Cyberinfrastructure. (2014, Aug.). *Guide to Developing Cybersecurity Programs for NSF Science and Engineering Projects, v1*. [Online]. Available: http://hdl.handle.net/2022/20026.

[5] V. Welch, F. Siebenlist, I. Foster, J. Bresnahan, K. Czajkowski, J. Gawor, C. Kesselman, S. Meder, L. Pearlman, and S. Tuecke. (2003). *Security for Grid Services, presented at Twelfth International Symposium on High Performance Distributed Computing (HPDC-12)*. [Online]. Available: http://www.vonwelch.com/pubs/SecHPDC12.

[6] R. Heiland, S. Koranda, and V.Welch. (2013). *Pegasus-CTSC Engagement Final Report*, Center for Trustworthy Cyberinfrastructure, trustedci.org. [Online]. Available: http://hdl.handle.net/2022/15562.

[7] National Center for Supercomputing Applications. (2011). *Report to the National Science Foundation Office of Cyberinfrastructure (OCI)* [Online]. Available: https://security.ncsa.illinois.edu/s3i2/S3I2WorkshopReport2011Final.pdf.

[8] C. Jackson, J. Marsteller and V. Welch. (2014). *CTSC 2014 Summit Report* [Online]. Available: http://hdl.handle.net/2022/19244.

[9] M. Suby and F.Dickson. (2015). *(ISC)2 Global Information Security Workforce Study* [Online]. Available: https://www.isc2cares.org/IndustryResearch/GISWS/.

[10] Advanced Scientific Computing Advisory Committee. (2014, Jul.) *ASCAC Workforce Subcommittee Letter* [Online]. Available: http://science.energy.gov/~/media/ascr/ascac/pdf/charges/ASCAC_Workforce_Letter_Report.pdf.

[11] J.Basney, P. Duda, V. Welch, and C. Jackson. (2013). *DataONE identity management system review*. Center for Trustworthy Scientific Cyberinfrastructure, Center for Trustworthy Cyberinfrastructure, trustedci.org. [Online]. Available: http://hdl.handle.net/2022/16926.

[12] R. Heiland, S. Koranda, and V. Welch. (2014). *Globus Data Sharing: Security Assessment*, Center for Trustworthy Cyberinfrastructure, trustedci.org. [Online]. Available: http://hdl.handle.net/2022/19165.

[13] J. Marsteller and R. Heiland. (2014). *IceCube Cybersecurity Improvement Plan*, Center for Trustworthy Cyberinfrastructure, trustedci.org. [Online]. Available: http://hdl.handle.net/2022/17364.

[14] J. Basney and S. Koranda. (2013). *Center for Trustworthy Scientific Cyberinfrastructure Engagement Plan: Final Report for LIGO Engagement*, Center for Trustworthy Cyberinfrastructure, trustedci.org. [Online]. Available: http://hdl.handle.net/2022/16689.

Data Management Plan

Description of data to be generated in this project:
During the course of the proposed project, data generated that should persist will be training materials and other public documentation (e.g., best practice guides, lessons learned educational curriculum, engagement reports). CTSC does not expect to generate or capture experimental or other data that would necessitate a relational database or specific data file formats for programmatic access from computer models. We expect that all of the data generated by this project can be projected into the Adobe Portable Document Format (PDF), and preserved as described below. PDF documents are commonly full text indexed by search engines, and are available to text mining and natural language processing systems. The project team expects that the ability to consume and manage content in the PDF file format will outlive the meaningfulness of the data generated by CTSC.

The project will generate some data, related to its work in software assessment and engagement activities, which will not be immediately public until we have had a chance to work with involved parties to perform responsible disclosure, after which time the data will become public. The PIs are familiar with this process and our cybersecurity plan includes our processes for handling this sort of sensitive data.

Responsibility for data management:
Ultimate responsibility for data management within CTSC will reside with PI Von Welch; however, CTSC team members will each be responsible for the management of data within activities they lead. This responsibility includes ensuring that the materials have appropriate search terms and metadata; have project, grant, and partner attribution; have the CTSC license declaration; have been cataloged as a project artifact and preserved according to the policies within this data management plan.

License for data generated as a result of this project:
All materials *de novo* generated as part of this project that will be distributed will be distributed under the Creative Commons Attribution-NonCommercial 3.0 Unported (CC BY-NC 3.0). The full terms of this license are available at http://creativecommons.org/licenses/by-nc/3.0/. This license includes the following terms: You are free to share – to copy, distribute and transmit the work and to remix – to adapt the work under the following conditions: attribution – you must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work). For any reuse or distribution, you must make clear to others the license terms of this work.

Data preservation, dissemination, and public use:
CTSC will leverage the Indiana University ScholarWorks system (http://scholarworks.iu.edu/) for data preservation. IU ScholarWorks is a set of services from the Indiana University Libraries and Indiana University Digital Library Program to make the work of IU scholars freely available and ensures that these resources are preserved and organized for the future. CTSC will also make its products available on the center's public website, and will take steps to ensure the NSF community and public are aware of these products.