>> Okay. Thank you all and thank you Kevin, Amy and Bill, and the rest of your colleagues for that kind introduction and for inviting me to present today. I hope everyone's able to see my title screen now. My presentation today is going to be on the area of cybersecurity to enable scientific research. Two things that are often seen to be somewhat in conflict with a rigor that can opposed by cybersecurity at times with the innovation that's required for scientific research, and what I want to talk through is how cybersecurity can be an enabler for open scientific research and talk about what the challenges are that open science has in terms for cybersecurity where cybersecurity can provide assistance. I will then talk specifically about the NSF Cybersecurity Center of Excellence, the role it's been playing to overcome those challenges and help the NSF and broader scientific community, and then conclude with some coming attractions in 2019 in terms of new activities by the NSF Cybersecurity Center of Excellence, as well as, the Research Security Operations Center that Kevin mentioned in my introduction. So, first I want to clarify what I mean by open science. Often times when I talk about cybersecurity and scientific research, that intersection of subjects leads to the research involving regulated data, protected health information and other forms of data that leads to cybersecurity by compliance programs such as HIPAA, FISMA, or NIST 800-171 for those of you deeply involved in that. On the other hand, the open science which constitutes the majority of NSF funded work is not guided by those compliance regimes, so when one thinks about the astronomy work or climate plants or physics or geology, there's not an off-the-shelf cybersecurity program for that type of scientific research even though at most of our universities this form of research is a sizable fraction and if a university doesn't haven't a medical school there's probably the majority of the work that is done at those universities. So, that is the form of open science and will be the focus of my talk today. I'm going to start with tackling a myth that we, myself and my team run into commonly in the community, which is we tend to equate the need for confidentiality, that is keeping something secret with the need for cybersecurity, and this is very understandable. If you think about a lot of the cybersecurity stories you read about in the media, it tends to revolve around financial data, credit cards, social security numbers and breeches that involve the loss of the secrecy of that data. So, I will start with talking about how open science doesn't fall into that particular paradigm, but just because it doesn't handle a lot of confidential data doesn't mean there isn't a need for cybersecurity. And let me start first with a key scientific goal of producing trusted and reproducible results, and this often starts with data integrity as being a key component of that trusted and reproducible nature of science. And in fact, for open science data integrity, unlike security and many other domains, tends to be the most important think that cybersecurity can provide for that science, and as we're seeing more political questions arising around scientific research, it becomes important not only to know that the data has integrity assurances, but to be able to assert that and prove that against those who might assert otherwise. The, a change that occurred over the last 5 years is what you may have seen a number of headlines is ransomware, and I'll put aside a lot of the technical aspects of that and focus on how it changed the landscape of cybercrime, in that before cybercriminals used to have to have some financial value in the data that they sold, credit card numbers are an easy thing to turnaround and capitalize financially, social security numbers always want to commit identity theft, but needs direct ties. With ransomware just like ransom itself in the physical world, it turns the value of its data to you into something that can be extorted by a criminal. So, it broadened greatly this fear of possible victims and

unfortunately roles open science medical number of different areas of research into now things that can be targets. Reproducibility is another key, and why reproducibility is still somewhat of an open research topic onto itself, it's clear that if you can't protect your infrastructure from uncontrolled, unpredictable changes by parties, then you're reproducibility is a risk because you don't know the state of the infrastructure that you're running your scientific research on. And this is going to be a particularly tough thing for cybersecurity to deliver, the cybersecurity has a need often to patch vulnerabilities often quite quickly and sometimes more frequently we would like which is at odds with a desire for stability to support reproducibility. So, this is another area that cybersecurity delivers on and is particularly challenged by in scientific research. The next area I want to talk about is about supporting scientific productivity, and this is another area where science has a need for the availability of unusual or even unique instruments in scientific datasets that often times can come at critical times in a scientific project lifecycle where these instruments are unavailable due to cybersecurity incidence. It can be very damaging to a scientific projects productivity. We also have a cultural difference in scientific projects. If we compare a scientific project compared to the institutions that often house them, the projects tend to be relatively short-lived starting with a starter's gun with funding of a proposal and then having a relatively short lifecycle that needs to progress quickly very focused on a particular scientific results, and this is a little bit different from our universities and other institutions of higher learning which have lifetimes of decades or even centuries and so this leads to some differences in risk tolerance and I would say, you know, the speed of which things are looking to be accomplished. These research projects are often very collaborative, spanning multiple universities and even countries and these collaborations are not defined with regard to the human resources databases or directories of the institutions in which they sit, so it becomes valuable and necessary for cybersecurity to support these sorts of distributed collaborations that don't fit nicely into the buckets provided by our local institutional organizations. We also have the challenge that there's a big difference between the cyber infrastructure that we use for our scientific research and Enterprise IT. We have some access paradigms in scientific research that are very different. We will often have shared access to computers across a distributed collaboration, again, this is something that's very unusual in Enterprise IT situation. We upload and run virtual machines or software and let people bring their own code to do a science project done on HPC or an HTC system and this, again, is unusual in an Enterprise IT environment. Specialized paradigms, science gateways and science DMZs represent specialization of information technology infrastructure to support with larger collaborations or a high-performance, and so far I've been to these differences, but these all become areas where cybersecurity is required to be tailored for these different information technology environments as compared to Enterprise IT and not only to sufficiently secure them for the sciences need, but to also convince our colleagues and Enterprise IT that it's an acceptable risk for the scientific community to be operating on their campus networks. Lastly in the productivity front, we have a question of reputational risk, and if we're seeing as having insufficient control over our research processes and workflows, this can invite outside certainty as to the autonomy of how we're doing things and the possible imposition of cybersecurity in forms that are not appropriate for the scientific workflow. So, we need to maintain not only the ability to appropriately manage our scientific research, but to make sure that that is apparent to others so that we can avoid having inappropriate cybersecurity pressed upon us. And I do

want to talk about even though confidentiality is not the predominant requirement in open science, it does still exist even in open science projects and a couple of examples of that tends to be the embargo period one sees in scientific projects, so this is a period between when results, data have been collected and before they're make public in order to do some quality assurance to make sure that data isn't wrong and causes some embarrassment or to control the dissemination of results for an announcement or other events. So, there are still there confidentiality requirements in periods of time where essentially there's a quality assurance or a quality control part of the project to prevent public relations mishaps and to ensure that what comes out of the project is actually accurate and integrity and trust can be maintained and there are also data confidentiality requirements unrelated to regulation, ethical concerns. The Wild Book Project is one such example. This is a project from Dr. Tonya Wolf-Berger and her team that does applies artificial intelligence to animal recognition, able to identify individual animals and photographs and even distinguish different animals in the same species and these species include endangered species and as I'll mention shortly, the NSF Cybersecurity Center of Excellence has engagements and we worked with Tonya in this case to ensure appropriate access control to these images out of their concern of not leaking images that might endanger these animals due to tourism or poaching. Moving onto the second part of my talk now I want to talk specifically about the NSF Cybersecurity Center of Excellence. That center often goes by the shorter monitor of Trusted CI or trusted cyber infrastructure and it has been in existence now for almost 6 years with the mission statement out there I won't read to you, but I'll just say in brief it's here to both lead the community in this area of applying cybersecurity to science and assist it in the application of thereof. And it's made up of not only myself and my team here at IU, but the other partners whose logos are displayed there on the bottom of the page, NCA of Illinois, Pittsburgh Supercomputing Center of Carnegie Melton and colleagues at the Computer Science Department and the University of Wisconsin, and it operates under a grant funding from the NSF. One thing I wanted to clarify about how Trusted CI operates is that it is not a technology-producing center, it's genesis goes back to a pair of workshops in 2010 and 2011, where myself and colleagues went to the NSF community and turned out to be naively asked the question, what technology do you need for cybersecurity? And we were pretty firmly told that a new technology was not at the top of their list. They were looking for leadership and guidance that they could trust to be unbiased and in their best interest, and we heard that in the first workshop, we heard it then echoed in the second workshop, and we listened and fortunately NSF also listened and, hence, the Trusted CI was warned as an organization that does not provide technology itself, it helps the community understand what to do with that technology, how to apply it and how to implement cybersecurity to advance their scientific research. So, the challenges being addressed by Trusted CI and this is a little bit of a summation here from my prior slides giving the motivation for cybersecurity for science, is it is looking to provide cybersecurity that meets the scientific community's needs for the trustworthy productive and reproducible science. That cybersecurity also needs to be broadly excepted so that when the community is applying appropriate cybersecurity that where developing is led by Trusted CI is seen by others as bring something that is appropriately acceptable and not needing a less appropriate regime cybersecurity program enforced in addition to what we're doing. We also recognize it needs to be reasonable to implement. As I mentioned before, science projects had to have short lifetime. They tend to be relatively compared to large

institutions smaller and that means they have ability to specialize in different areas of workforce, so where a large institution may have special people who could specialize say in network security, anything but the very largest NSF projects would struggle to get to that level of specialization in their workforce. The final point is around the National Science Foundation has a grant funding model to projects and it gives projects quite a bit of autonomy in how they conduct their scientific research. And that means unlike other federal agencies, they do not mandate from the top as much as you might see, for example, in a DOD context where cybersecurity would be seen as something that is mandated down from above and, hence, in the NSF context, one of the roles of the Cybersecurity of Excellence is not just to help with implementation, but to help with community consensus building and drive the leadership on understanding what cybersecurity for science means and tackle all of these things as bottom up from the community. A few months ago, part of the Trusted CI team then looked back over this 5 years that we have been working on this and compiled a report on our impacts which I'm very proud of the results in this report indicating, you know, we have now impacted nearly 200 NSF projects since our inception. Our Annual Cybersecurity Summit regularly draws a 120 people from the community and through our monthly webinars and training, we've delivered hundreds of hours of free education to the community on cybersecurity topics and I'll have URLs for these later on in my talk. The last bullet up on here is we've conducted 35 engagements including 9 with NSF large facilities and engagements are one-on-one collaborations where a project comes to us with a particular cybersecurity challenge and that really is quite broad. We have a lot of projects coming to us starting from scratch and looking to get going in cybersecurity, but not being sure where to start all the way to other projects that have very mature programs looking for an assessment to other projects that are looking for specific advice on software technologies or other technology features and how to best implement those. We've also worked in areas identity and access management being a common one, giving advice on that and we've also done work on privacy working, for example, with the array of things in the city of Chicago with rolling out their instruments across the city and helping manage the privacy aspects around that. We accept applications for these engagements every 6 months and I just wanted to mention we are currently accepting applications for the first half of 2019 on our website at the URL shown, and if your project you believe might be able to benefit from that, I would encourage you to apply or contact us before October 1st to discuss that application. In addition to the one-on-one applications, we produce a fair amount of community-driven guidance in the form of best practice documents. We do this very carefully. One of the principles of Trusted CI is all the members of the team spend time with other NSF projects doing operational cybersecurity or at their own institutions or otherwise have what I call "boots" on the ground, myself included if you reflect back on some of the other roles Kevin mentioned that I have. And then we often do this guidance in collaboration with members of the community to make sure that it's reasonable and applicable to the community and that we're not getting too high up in an Ivory Tower casting out advice that isn't, isn't practical and you can find this particular guidance is all freely available on the Trusted CI website. The Annual Cybersecurity Summit that we organized has turned out to be invaluable for community and network building not just with our projects, but also between community members. We hear great things about hallway conversation and folks meeting their peers at other projects, it turns out to be a really valuable service in light of these projects limited lifetime having giving folks

the opportunity to connect is really powerful. We start the summit with a day of training and workshops, so an opportunity for people to learn and dive more deeply into specific subjects and then we have an agenda that is almost entirely community driven through our call for participation. So, it is talks that were submitted by members of the broader community and then presented everything from their success stories to their lessons learned, so we really encourage it as a forum not just to standup and trumpet what's going well, but in times the challenges that we've been facing. In 2019, we're going to move the summit from its traditional home on the East Coast and have got an offer from colleagues out in San Diego to help us host it out there. So, we will be out on the West Coast in 2019. I encourage those of you who are interested to join our email list to stay abreast on announcements. I want to mention one more document that's of great importance to Trusted CI and that's we've put together a 5-year vision and strategic plan, both for a much broader NSF cybersecurity ecosystem and then the plan being focused on Trusted CI's role in that ecosystem, and this is very much a both a public document and a work in progress; it's intentionally labeled Version 1. We expect to evolve over time and very much want the community's feedback on this document to help guide us and the community as a whole going forward. I haven't touched on all of Trusted CI's services here in respect of your time, but I encourage you to spend some time on our website. The cyber infrastructure vulnerabilities is a new source of vulnerabilities related to security that's tailored for the NSF community in terms of filtering out technologies and in terms of giving additional guidance. The number, we have a number of specialized information for particular communities is listed on the lower left and then we also for the large facilities for being a cybersecurity team on a monthly basis discuss topics for them. And I encourage folks if you have any questions regarding cybersecurity to just please ask us. I mentioned the engagement process earlier which is our more longer term process for working on challenges, but we also get questions from the community that are just an email exchange or we'll answer them in a blog post and so I encourage you to send us anything, noting too big or too small. I'm going to conclude my talk now with some coming attractions for 2019, and first I'm very happy to announce that through a supplemental proposal Trusted CI has been extended through 2019 and also funded to take on an expanded scope. And so, I'm happy to welcome three new members to the team that are indicated there as underlined and with the logos on the right-hand side of your screen with the two logos and then a picture of Florence as an independent consultant. And I want to talk a little bit about three new activities that we're going to undertake in 2019. The first is the topic of transitioning cybersecurity research to practice. And as we've had a lot of experience now in the community with cybersecurity research, and yet we continue to see failures with regards to cybersecurity being very pedestrian. It's been, a lot of the community has started to recognize a number of challenges around getting the sophisticated research undertaken by NSF and other communities into practice, and there are technical human factors, usability and economic challenges all around that. And so, Trusted CI will be helping in this area by focusing specifically on the needs, looking at the needs of the NSF cyber infrastructure community identifying those we believe can be helped by research and then working backwards towards the research community to identify research that we think will fit in there and then hoping to match make the appropriate parties along the way to foster that research moving into practice, and we've already been reaching out to a number of members in the community particularly through Florence, but if you'd like to contact us if you believe you have a need that could be filled by

cybersecurity research or if you're a researcher that believes has research that can fill a need or looking for a need in the NSF community, we encourage you to contact us and we would setup a special email address TTP@trustedci.org for doing so. The second expansion for 2019, is we recognize that Trusted CI has had a good impact already across the NSF community, but as always there's room for improvement and we want to start a network of cybersecurity fellows to help u particularly extend our impact across a number of different axes such as, you know, the NSF science directorates we'd like see having more impact in some of those directorates than we've had to date. NSF now also has the 10 Big Ideas and having cybersecurity influence and improve the science for each of those 10 Big Ideas would also be a criteria for this, and then also improving the representation of the underrepresented groups and institutions and ensuring that cybersecurity for research has impact across those, that part of our community is an important goal of the Fellows Program as well. The Fellows Rules serve as liaisons between Trusted CI and their particular community, and it will be supported on the Trust of CI side with training and travel support, as well as, a prioritized support, communication from Trusted CI and this is a model that has shown success in other communities with [inaudible] our colleagues in the UK at the Software Sustainability Institute, the ACI-REFs Project, Campus Champions and so we're hoping that this will have similar impact within the NSF community for cybersecurity. The third and final area of Trusted CI for 2019, is I've mentioned now a couple of times the unique needs of the open science and cybersecurity and certainly Trusted CI has been putting guidance together on this over the past 5 years, and one of the points I've mentioned is we need to make sure that we're doing is seen as accepted by the broader community to avoid the influence from less appropriate cybersecurity frameworks. And so we see taking our guidance now and formalizing it into our framework which seems to be a common name these days for the Cybersecurity Program, such that it can standup to an be seen as serving on the same stage as other things such as demist of risk management framework and other frameworks of a similar nature. And Trusted CI is going to lead the development of this framework building off our existing guidance, but we also want community input to ensure that it meets all of Trusted CI's goals. And finally, I want to just mention a new second cybersecurity center that has been funded by NSF that I also have the pleasure of leading and that's the Research Security Operation Center. So, while Trusted CI is giving advice in community building and leadership, this as a security operation center, the research stock will be focused in providing cybersecurity services to the NSF community and in-turn helping them apply those services to bolster the community's resilience in the face of cybersecurity incidence. So, very focused on pushing technical services to the community supporting those and then also providing the appropriate training and tailoring of those services so that it meets the needs of the NSF science community as I've laid out in my talk, and this is with a slightly different set of partners besides myself and Pittsburgh Supercomputing Center colleagues at Duke and UC San Diego are a part of this effort. So, please look for that ramping up in 2019 and coming online in 2020 and I certainly welcome queries and thoughts as to these services as well. So, I'd like to summarize now and the leave some time for question and answer. So, I hope I have convinced you that cybersecurity is critical to open science research, particularly in ensuring that it's productive, trustworthy, and reproducible and Trust CI is here to help in that regard, both in leading our understanding of how cybersecurity can fill those roles and also helping projects with that implementation, and then hopefully I've wetted your appetite for some new things

coming down in 2019 [audio cuts out] position to practice the open science cybersecurity framework. And so, I want to thank NSF for their funding of Trusted CI. Once again, I thank the Trusted CI team across both my center here at IU and the now a 5 partner sites and I'll leave this slide up here with contact information during Q and A so that you can note any addresses you might like, and with that, I will thank you for your attention and let's see, I don't know if any closing comments by the colleagues of NSF or any questions have come in?

>> Yes, this Bill Norris. Thanks a lot Von. Again, if you want to ask a question because the audience is not able to use the audio, send me an email at wlmiller@nsf.gov and I'll relay it to Von. We'll wait a few minutes and in that interim time I'll just mention the fact that this and the other webinars are going to, they're recorded and their posted to a YouTube channel that we have at on YouTube. It's hard to read the, we have a short URL for that, it's Bit.ly/2JOw3Tq, but you can find the link to that on our website at NSF, so if you don't find it definitely write to us. You can write to me wlmiller@nsf.gov and we'll be happy to send that to you.

>> Bill do you have any questions ready? If not, I can ask Von a question.

>> Yeah.

>> Alright, hi Von this is Kevin Thompson again at NSF and thank you for that presentation. It was great. One of the new elements that you briefly described moving forward on this project is bit more active engagement with transition to practice activities and my question is, given that NSF has been supporting at a somewhat small scale transitioning to practice activities through small and medium sized awards in the basic research program at NSF for cybersecurity all the SaTC for short, secure and trustworthy computing, given the existence of some funded activities that underway in the component of SaTC that are all about transitioning to practice, do your plans include perhaps contacting some of those projects and getting to understand their activities better and potentially explore how they might fit into some of these scientific CI setting or not?

>> Yeah, thank you Kevin. Good question. And the, the short answer is yes. As you know, well for those of you who, some of you may know Florence over the past year posted a couple matchmaking workshops between SaTC PIs who had a transition to practice interests or award and then chief information officers and chief information security officers at various universities. And we saw some initial good successes with that. It also make it very clear there's still a huge gap even if you find somebody who's interested in deploying some research and the researcher himself and they seem to be a good fit, there's still a large way to go across that chasm to get that research there. And we're planning a workshop over the next year and what we want to do is take that previous model of matchmaking between researchers and potential employers and add some entrepreneurs into the mix. So, folks who have the interest of picking up research and, you know, commercializing it or putting it into practice in some way and seeing if we can make a 3-way match there and advance this particular area in terms of how to go about this. We're hopeful on this, but we recognize as I mentioned in my slide, transition to practice is a research challenge into its own right with these different aspects, so we're very

interested to see how this, this new look approach works and if there are folks; we're going to host this in the Chicago area is the plan and if there are folks out there that think they might be interested on any of those 3 groups that I met or have interest for other reasons I would definitely encourage them to contact me.

>> So, hi Von this is Amy. Let me echo Kevin and thank you. It's always a pleasure when I can really understand the cybersecurity presentations. Your model has been very much about empowering the people who come to you for help. Do you track that and can you talk to sort of what it's like to setup a program or to provide advisement to a program and then to watch it take on a life its own?

>> Thank you Amy. I appreciate that compliment. And to answer your question, yes we do track engagements after we've left following up every 6 months afterwards, year afterwards with a questionnaire asking how they go. And what we have found, we found a variety of things. It's really been interesting. I'll try to summarize it here. Sometimes it's clear that we're a catalyst and we're just the spark they need to start implementing and doing things internally and they will go gangbusters; other times it's clear that we uncover a deeper problem in that they're trying to essentially implement cybersecurity without the resources to do so, and so all the guidance in the world gets them back to the issue that there's nobody there with the appropriate time or resources and so we cause an initiation of a higher level conversation within the management of that project then to cause the resources to be done. We've also learned that different engagement tactics are necessary depending upon their lifecycle. Projects that are new and just starting up, which is often the case, you know, we'd like them to be considered security, often also tends to be one of the cases where they're most chaotic still trying to get their feet under them so to speak and they're designed [audio issues] from month to month and so there can be, what we found in those cases is we need sort of a longer, slower engagement of conversations rather than a deep focus dive. So, we've had those. We are gathering data longitudinally I think that's the heart of your question. We've had community surveys for a couple years now. I think it's too early to be saying much about the corpus of data we have. It's, as I conveyed earlier, it's a little bit all over the place for me to be able to make a grand sweeping remarks, but I think we've been learning quite a bit about how to shape our engagements to have the impact into what factors, you know, matter learning to ask questions about resourcing, management support, lifecycle and all these different aspects of the project instead of just going in and, you know, starting to do our thing.

>> So, not seeing any further questions. I thank Von on behalf of NSF for the webinar and thank everyone for connecting in and taking your time to listen to this. Just a quick final note, that we will running these monthly. Our next monthly webinar will be on October 18th at 2 PM Eastern and the speaker will be announced. So, stay tuned. Thank you very much and this not concludes the webinar. Thanks Von and have a great day everyone.

>> Thank you all.