Keith Marzullo: Thank you for coming. This is our third, it is our third, yes our watch lecture. Just to remind you this is a lecture series that we are having on every first Thursday of every month. These are a set of talks. We bring in {indiscernible word} of people in the area of trustworthy computing and cyber security. This is a corporation between SPE, CISE, and OCI. In the effort to try to build a broader understanding of the issues that lead towards the cyber secure trustworthy cyberspace. So, today's lecture is Klingenstein, he is the director of the Internet to Middleware and Security Initiative. His internet to middleware role, he is responsible for posturing the development and insemination of middleware inoperability and best practices, through partnership efforts of leaders among campus IT architects, corporations, and government agencies. {Indiscernible word} results of this activity are cabalist, SEML, and federated identity. From 1982, 85 to 1999 excuse me, he served as a directorate computing and network services at the University Colorado at Boulder. A beautiful place! His responsibility included overall management for media and networking and computing at the University. Ken has been a leader in National networking for the past 20 years and received his PhD in applied mathematics from the University of California at Berkley. Go Cal!

Klingenstein: Thanks Keith! I'm very honored to be here. This is a very special institution in my life and the world and I'll come back to mention that in just a second. I'm really thrilled to be here. So, the theme here is going to be. I want to begin with by 20 minutes of context, the emergence of the internet identity and trust layer and then I want to get to the premise of the talk, which is that this layer offers much in leverage and helps the solution to other problems in internet security. So it's not just an end in itself. It is a real layer that new solutions to other chronic issues in security have the opportunity to leverage by this. I'll give you some examples in turns of wide-aperture security tools, role based access controls, zero knowledge identity, etc., Tao of attributes and to take away for this a gust audience is that you have the ability in your solicitation and your invention going forward to create a whole new raft of cyber security tools by leveraging the infrastructure that some of us in the mist of deploy, so that's the context for the book. So an apology about the title, it use to be esoteric. Does anybody not know the story of Turtles all the way down? {Indiscernible phrase} Anybody not know? Ok, so very quickly {indiscernible first name} Russell told me this, not really! This is a { indiscernible first name} Russell story he was given a talk in 1905-1906 and talking about the nature of the universe and at the end of the talk a little old lady from the audience stood up and said "Well that's all very good Dr. Russell but that's not how it is. We all know that the world is carried by a turtle." And Dr. Russell said but madam what's the turtle walking on and she said another turtle and he said but madam and she interrupted and said don't go there Dr. Russell its turtles all the way down. So it use to be esoteric and then Steven Hawking that popular book writer used the story in one of his books and so now its mugs, t-shirts, a song, and most importantly an achievement in World at Warcraft. So there's some of the t-shirts and there's a mug. Here's the song by the New Junk Aesthetic just in case you haven't heard it or aren't tapping your toes to it. With the chorus of "When we think we've reached the end we're only back where we begin superman, superman, superman." Truly, truly a popular tune and then finally the World of Warcraft achievement, which you only get by riding a sea turtle and I, I've read more about this then I should have and it took 3350 clicks for somebody to get this World of Warcraft "Turtles all the way down." So because it becomes so popularize let's talk about trust and infinite regression because that's really what Turtles all the way down is the embodiment of. There's a couple of interesting analogies between trust and infinite regression. Trust builds on itself. It's very much of a feedback loop that as you have trustworthy interactions with individuals you intern build more trust in them. Transitive trust has always been a very difficult subject. In union context in PKI we spent years trying to limit the chain of trust that can be built in constructs of {indiscernible word}. So you've got to be very careful about how you work that space. Trust can move from one context to another. I can trust you as a neighbor and then discover that you have a certain kind of business and I can move my trust in you as neighbor to a trust about your business. We get into infinite regression because the other choices are not good in { indiscernible word} and so we don't like circular arguments which is an alternative in infinite regression and axiomatic use is an alternative in infinite regression will try to avoid those in our space and then the point of this talk. Trust can be leveraged to enhance other technologies. I'll come back to in a few minutes after I talk about the context. It's a small enough room that if people have questions please interrupt along the way. Raise your hands. So I want to talk about what's been happening in terms of Internet identity and attributes over the last several years. There's been an emergence of internet identity. It's just gone explosive. In two flavors, there's the flavor that we typically have been promoting in the R&E sector which is federated identity. I'll come back and talk about that in some depth. It's well anchored in the US, it's extremely well anchored oversees. So the state of federated identity for our colleagues in Europe and in Asia is much more advance, much more comprehensive -large because we're trying herd cattier, I'm talking to a room full of cat tiers, you know how hard that is but what if the sea changes that's happening and I'll mention this just a little bit, is that we're starting to see federated identity become a normalizer of behavior among institutions and we never thought we would get there. And so we're seeing MIT and Stanford come to me and say can't you get these other universities to populate all of the attributes, can't you get all these other universities to put consent mechanisms in front of users. This is wonderful stuff, we did it in '85 and '86, I remember when we first deployed {indiscernible phrase} first campus ticket on the internet, our NSF net in those days and what, what, what number should we pick for our various ports on our router. Well our external port, let's name it dot 19. What's wrong with dot 19? Well normalization set in. Dot zero for going out, dot 1 for this interface, dot 2, so that same normalizing behavior is starting to invent itself in federation, the results of a set of theoretically interoperable social identity providers serving large masses of social and low risk applications. How many of you have a Gmail account? Welcome to the social identity world. I will try to distinguish social identity from social networking. Social identity is valuable, social networking get a life! But that's my personal geezer opinion here you know about this. The federated use values a lot by, by, by country, in some countries its 100%, Denmark is a notable example. Everybody in Denmark has a federate identity. People got it from their universities to some degree and then the Danish government, said good thing because I intended to give federated identity to every citizen that can do as one buy a small company that got purchased by banks the next day and so the banks of the identity providers for citizens in Denmark. You cannot get welfare in Denmark unless you have an identity because it's all electronic transactions, so you can be homeless, you can be without shelter, but you have to have a federated identity. You can see that pattern in some other countries as well and then it tails off in some other places. In the US the R&E sector is vibrant. I'll get back to that in a second in its intensive federal and state government use. And now we're finally starting to see the verticals begin to form federations as well, multi-liable federations in particular. So you see it real estate, you see it in medical industries; we see it in security industries and a few other places as well. It's ironic that I often site places like real estate and security as where you wouldn't want to emulate them today, but that's a separate issue. Social identity a large scale phenomenon beginning around 2007, there's a number of major players currently sharing a set of non-interoperable deployments of weak protocols. By weak I mean, cryptic-graphically weak, privacy preserving weak, but Chu-ching rich in that the opportunity to put advertising around these exist. So there's two to three clusters of players here, Google, yahoo, and PayPal is a major set. They talk, they work together in standards processes and { indiscernible word } standards processes and that the rest of us aren't allowed in the room but those three or four companies are able to do some normalization and then facebook plays by itself. For a couple of reasons not the least at which is when you click on that terms of service when you get your facebook account, you have signed over everything you upload, all your preferences, etc. to facebook. They are definitely afraid that any kind of scrutiny of this will lose that term of service and they want to keep them terms of service in place. There may be convergences happening in the social identity space, so these are not interoperable today. There is a standard called Openid- Connect, it is advancing forward, again {indiscernible word} standard and that is several of these companies working together. Ironically their using a large amount of federated identity approaches, so attributes are getting into that, privacy preserving is entering their mind set, { indiscernible word} is entering their mind set. So the good news in this is that when we did SAML and Shibboleth we have the idea pretty quickly and then we was sent out by our fearlessly leader for a year to make sure there was no simpler way of solving this problem. And after a year, we were satisfied, that if you wanted to preserve privacy, if you wanted to avoid correlation attacks, if you wanted to carry attributes, what we were doing sure looked minimum in that phase. What we're seeing now with these companies, with the same elevated goals finally of privacy preserve, etc., are coming back and saying "Wow it looks a lot like SAML doesn't it". So it's a reassurance that what we did ten years ago, twelve years ago, is truly a minimum set of complexities necessary to put this in place. So the Openid stuff, I'll just mention one more thing about this, which is that when you go out to Silicon Valley. I was there last week. The application developer is key—it's all about the coolest new application. Application developers do not want to deal with systems people. They don't like moves, they don't like dealing with those people who worry about security. And so they've liked Openid because boy I didn't have to deal with the systems people, I didn't have to go down to the Java stuff. I could just sit here in application space and do identity and now we’re starting to show them some of this complexity in the Openid connect space, we’ll see if they buy into it or not. It is a broken market place, whether it will heal or not is a big question. That’s not to say though that we don’t want to embrace social identities, we need to. Not everybody has a federated identity and everybody has a Gmail account. So cool things are happening like social indiscernible word} gateway, where I can go out authenticate using my social identity and a SAML assertion comes out the back end to a relying party that says, this user has authenticated in twitter, you can decide how much trust you want to put into that, how much indication. But we are starting to stitch together these worlds, which is good because we need a comprehensive environment. R&E SAML federations, this is an old chart. It’s expanding rapidly; there is now a federation in Saudi Arabia, which is a shocker to us, because when we first started this stuff, we exchanged emails one night, where Bob Morgan, my fearless leader and I both came up with the name Shibboleth and we said “well you know if this stuff is successful we can’t use that name because it won’t play everywhere and it’s not going to be successful, so we went ahead and said good lord its successful their using Shibboleth somewhere in the middle east, I don’t know what they call the software. So very comprehensive, Eastern Europe is now well covered. China has a very growing federation. What privacy means in China isn’t critical but privacy preserving is but the tools are available to them. Yes? {Question from the audience- What’s Shibboleth?} Shibboleth is our open source of—{what does it mean?} Oh so oh, I should have pulled up that website. So Shibboleth is an Old Testament phrase. It was used to separate the Afrimmites from the Gladites at a border crossing because the Gladites could not say the sh sound. So they would say Shibboleth at which point we thought they were stopped from going across the border, actually they were killed. So if you want to follow this, some people know this, there’s something called the brick testament.org. It is the entire Old Testament done in Lego blocks. If you go to, well I don’t remember which part of the Old Testament it is, you will see little Lego blocks with blood scattered all over their stuff from the Gladites who couldn’t get through there. So yeah it’s a great website. Probably the first time that was ever mentioned in this room I would guess. So in common is the federation that we’re building here in the US, it’s got over 250 universities, 400 total participants, the growth continues rapid, it’s roughly expediential. I got to ride the expediential curve once before in ‘85 it’s a thrill. 10 million users, none of those people know they’re using you, but all of that’s changing shortly. Traditional uses, outsourced services, government applications, access to software, over 50 percent of the educational software that Microsoft provides to students is downloaded using Shibboleth { indiscernible phrase} so welcome industrial use of that. New uses bloom: access to wikis, I’ll show you that in a second, shared services, command line finally, calendaring. If you do doodle you should know about foodle, if you know about foodle which is federated doodle, then in fact you should know about, well and so it goes. I’ll come back to UHC in a second and then finally we’re getting back into the certificate business. This time for sure as Rocky and Bullwinkle would say. Um some of us are OPKI warriors but we’re never going to do PKI the way it was done originally because it was a broken paradigm and that I would leave with the certificate that had my identity in it, so privacy was gone the moment I authenticated. What’s happening now and it’s happening in DHS and lots of other places is I’m using {indiscernible phrase}, high assurance certificates but I’m using that for a local authentication and then what’s being passed to the other side is we have a very strong identity here, you don’t need to know who it is, I’ll come back to that in a second. So important services, research.gov including NSF Fastlane and I’ll get to that in a second. Electronic grants administration from NIH. They have a killer app in the clinical trials wikis the CTSA wikis that is driving massive use of federated identity in NIH and I’ll show you some pages about that in a second. The Mayo Clinic is now a member of in common. UHC is the University Health Care contortion; it’s the 95 biggest research- medical research centers in the country. They have joined in common long lost. It’s a very interesting phenomenon not just because their big dogs but because they in turn have tentacles reaching out to community clinics and other kinds of resources within the health care sector and so we’re about to, you know once you start to tug on the medical research center, all of these other medical facilities will start to come in and I’m not sure we really want that many in common but we do want to promote federated identity and so we’re letting them in. IEEE, the National Student Clearinghouse, a very interesting phenomenon that’s emerging is the College Board wants to be an identity provider for large numbers of students who don’t happen to have federated identities. College Board does lousy identification { indiscernible word}, you just, they’ll send you another user name and password, but then they have this business process when you show up to take the college boards, that checks two forms of photo id and all kinds of other things. And so that lousy identity suddenly becomes a higher assurance identity, what can we do with that? Now that’s where the interest of some parts of the federal government going forward. Some other possibilities within that. I want to give credit to NSF. Yes... {Indiscernible question}Yes, Yup, and that’s um, you know we’ve worked closely with them, they’ve drunk large amounts of Kool-Aid, so in fact some additional things will be coming down that will be very cool in terms of attributes and group memberships from GENI. So, NSF funding um, was critical to this and I just want to thank this building that led to SAML and Shibboleth that led to development of access control, attribute ecosystems, leading to internationalization, commercialization, and infrastructure. This is a lot like TCPIP layer and somehow this organization desires a lot of credit that it’s not getting for having to shape the commercial market place and international market place. There’s a story here to deliver. I should also mention that in 1964, I went to NSF Science Honors program at Rutgers University so I am dyed in this wool folks, and those were formative days for me. So, I don’t think can afford these kinds of science honors programs anymore, but my god for an eighteen year old kid, it was transformative as I learned how to program in P01. Ok, NSF as a deplorer. So NSF joined InCommon in 2009. Research.gov is now accessible, and the intentions to have NSF become an IdP as well, so that you can use your login at NSF to get to your partners around the world, to get to some projects that you might be supervising, you deserve to be full class citizens in this space and have your identities, which are actually stronger identities because if your identity proofing mechanisms be useable in the external world, that’s coming down the road. That will be a small bit of payback for the supporting of this development and OCI supports things like {indiscernible word} and etc. Just want to cover a couple of other identity in service to the federal government other folks are here from other agencies. So we talked about some of the scientific resources, over there in the department of energy they are building a federation among their labs. That federation rides on top of InCommon, so it includes the InCommon levels of assurance and the attributes we use and the contracts, but then they add additional attributes that are organized around energy specific topics. And then major virtual organizations such as LIGO and iPlant are leveraging the infrastructure to great effect. It's really cheap thrills, I'll come back to that in a second, but it's… NSTIC, the National Strategy for Trusted Identities in Cyberspace, its major White house initiative, when President Obama came in within 30 days of him taking office he ordered a comprehensive cyber security review. That cyber security review identified identity as one of the major areas. And so the government wants to get citizens to be able to use identities for government services but they also want to anchor a commercial market place. So, their dancing with the big dogs of Google and Yahoo, and Facebook doesn't dance it turns out for anybody. But their working with the commercial market place, boy I sure hope it works! On the other hand it's not clear it will. Some—again some of these big companies don't see the value of it, limited revenue stream. Until NSF put advertising on its website you will not be the cash cow that Google wants in this space. The other thing that matters is whether or not there will be a charge for transaction. So five years ago, in the PKI space we danced with the banks because they have good mechanisms for identity proofing and they have good order processes. And we ran some tests with the federal government, where you could use your bank account to be able to look at your social security information. And at the end of that six months test, which worked really well, the bank said so what are going to pay us per transaction, end of story. For those of us who are veterans, this was a lot like – Are we going to build per packet in 1985, and for that story the answer turned out to be, we—I was working with Steve { indiscernible last name } and others at that point and we talked to the Telco's and 52% cost of their service was billing. And that's on phone calls... That's not even on per packing basis, sooo, why? By the way if NSTIC makes it, it works for this community. In fact, there's an animation on the NSTIC page, I won't pull it up now, but it's exactly a picture we drew on Ethiopian restaurant place mat in 2001, so Division has been captured and is being implemented. DHS and DOJ, there driven by the high need of security and secrecy, so its maybe the world's largest federation in terms of systems and scope. They have gone beyond anything that I can understand in terms of attributes, so I'll show you the attributes we passed in our R&E community. We have six attributes we passed. There's 350 attributes in the DHS, DOJ scheme. A lot of attributes! They do it for secrecy; they don't do it for privacy. They do it because the operatives need to work with each other, but they don't want to reveal who their operatives are to each other. So, basically we say we have a very high assurance authenticated person on this end of the pipe, and we can tell you who it is if that's what's needed down the road, but we're not going to tell you now and let them into the content, and it works if they just { indiscernible word }. Standards processes, I think this is really important for success. So, we did the SAML spec via OASIS, with three of the seven members of the technical committee engaged with Shibboleth. We're now starting to the non web federated identity work in the IETF, and then we're doing a lot of stuff in the policy space with NIST guidelines etc., and it's important to scale and create commercial services, you guys know that. Just a little bit more about InCommon, it's a work in progress; so, we're trying to manage the growth, we're getting up to higher levels of assurance, we're putting end users in charge attribute release and their privacy, we're expanding mega participation. This is a very powerful normalizing force. When MILT and Stanford say, "gee Cal Tech is not populating all the attributes that we need, would you talk to those people?" Wow! Yeah we'll talk to them. Boy if we can get—Identity management should not be a cottage industry, it's not what makes us distinctive as a community. If we can normalize this {indiscernible phrase}. Federation is a work in progress itself, so this interfederation, think of BGP, one level up the stack. In fact I have some interesting conversations with Scott { indiscernible last name} about BGP and whether it's a metaphor for interfederation, and you know BGP works on the basis on the number of hops and things like that. When I said, "Oh the perimeters here Scott are locust of a judication," he said "I don't want to touch this." -- Privacy management, application to non-web apps, and collaboration management platforms. So this is more than identity, in fact identity is becoming less critical because we like to lead with attributes, cause if we lead with attributes we have at least reserved preserve privacy as a stock. We have also got into scaling capabilities, etc. So, the attributes are connected to identity and they need trust. They do syntax and semantics both and there's lots of processes in the identity and attribute echo system. You'll see this reference in the NSTIC video and other places as well and it is an area of future resource. Give you a sense of attributes we're talking about, but there's a lot of them that are institutional attributes. Your role here at NSF, my role on internet two -- It means that I have authenticated against that organization, there's organizational roles. One of the other things we do however is – on my campus we catch a citizenship of foreign students, it's required by service legislation. We're not definitive on that, except we -- we're given guidelines about how we capture that stuff, so is citizenship a useful attribute? Let's look at Crackn, which is this super computer at the University of Tennessee. It requires citizenship; it uses a self asserted value for citizenship. I think I'll be {indiscernible word} today; wouldn't it be nice if we had authoritative sources for that one? There's the governmental attributes, if we want to enable – in fact I was just talking to { indiscernible first name} Warren about, wouldn't it be nice if Congressmen know that the warm notes that their getting from their constituency are really coming from their constituency. Wouldn't it be nice to keep you at anonymity, but have something that says yes you're in this particular precinct. There's the temporal attributes – geolocation, who owns those? Who owns your cell phones location? Well, it looks like android thinks they own it, looks like the Telco that's providing the android service thinks that they own it, everybody's clear that you don't own that information, however that's not the way it should be going forward. There's the community and collaboration asserted attributes, we're working a lot with virtual organizations, I'll come back to that in a second – and then the informal stuff for friends of friends and reputation systems, and finally there's the self-asserted attributes. And typically people think of theses as junk – these are my sports or my hobbies. On the other hand one of the biggest attributes in use in Europe is preferred language. There were countries like Switzerland, where choice of language is rich, wouldn't it be nice. And you know how that's solved, typically they throw up a splash page with enter in five different languages and whichever one you click on, they deduce that that's your preferred language. Well wouldn't it be nice if we could do some other ways on doing that, same with accessibility. So would cut the pipes of our federation and say what the hack is flowing through those pipes. You would see these attributes. There's a large number of just member of the community, and that's good for our content and a lot of other purposes, faculty and staff good for other purposes. We also have an Opaque persistent non-correlating identifier. Opaque 32 bits can tell a person's identity, we use a different identifier for every site you go to, so you can't do correlation attacks. This was stuff that was a complexity that we thought -- Oh can't we avoid this, we came back and said no we can't avoid it. Now Google is using this technology, others are using this technology, it looks like it's a minimum set. There is a persistent and human-usable identifier, like kjk@internet2.edu, a name – wouldn't it be nice when I hit that wiki it doesn't say – hello 326513 but says Hi Ken! Wouldn't it be nice if we had that display – email addresses and the open-ended set of entitlements. These are the pay loads that we pass in our attribute echo-system. So where we may be headed in R&E is towards an international peering of SAML R&E federations, with common attributes and LOA, so edge-u-persons, so of those attributes you just saw are adopted globally. And then some – careful integration of other identity approaches, and then on top of that I want to start to talk about the collaboration VO identity management overlay because your business here is to support science and research, and we now have some tools that we think are going to enable that in dramatic fashion. So, I went – well let me hold off for a second. So collaboration management platforms, this is the work that's been most recently supported by NSF that we're doing and it's a collaboration identity management system, and a lot of it goes back to a scaring workshop that I did – participated in and some of you did – BEVO (Building Effective Virtual Organizations), so I hope the funder wasn't in this room, I'm not about to say nice things. But he went into the weeds really quickly. The first speaker got up there and I think he was talking about neisgrid, which wasn't really a great success at that point. And he's going – so he's taking about some of the things they learned about virtual organizations and at the end he said but we learned one thing for sure, whatever you do stick your collaboration inside Saki. So the second speaker came up { indiscernible word} their entire talk and said I got to take issue with that Saki it's really modal and we disintegrated into a day and a half arguing what application did you want to be locked into. Why would you want to be locked into any application? Why wouldn't you want your identity and access control to be independent and feed domesticated applications with information about identity and access control? And so out of that traumatic experience came the idea of building collaboration identity -- management platforms and we're rolling it out and we can just tell you that the scientist in LIGO and iPlant, the unionist in Bamboo, the women in the earth sciences women network all say – this is really good, its increasing our productivity, we can bring a new application in and make it subservient to our groups, make it subservient to identities. What applications can we do, well we can do wiki lists cal ending {indiscernible phrase} and I'll show you a whole raft of them and we can do the domain apps as well. And in particular we can span command line and web, and one of the things that I just saw in some report that NSF issued, the challenges of getting access control to work on both command line and web is not a challenge. So what can we do? We can create and delete archive users, accounts, keys, group management, blah blah blah. And we've done this by basically cobbling together a whole lot of our existing tools but building them - rebuilding them for VOs verses for enterprises – well go into all that. So if this is so big why can't I see it, well first of all my totals in infrastructure is not suppose to be visible, but the visibility is growing. And so here's the research.gov website as of last week. Notice that I can authenticate as a visitor, user, staff, or {indiscernible word}, thank you NSF. Umm there's now eleven universities that are enabled, there's another thirty in the cue. Um my guest is by the middle of next year we'll have 150 universities using local credentials to be able to get to fastlane, that's good because I don't even remember my fastlane password. This is the Pub med page, after you get to Pub med you click on sign-in and you get to this site – and you can see that I can sign-in by a Google, NIH login, or a pole analysts that looks a lot like the InCommon membership, it is the InCommon membership. So if I click on that pole analyst I get this list of, well actually that's one quarter of the universities that can access NIH facilities using their local logins, so that goes on for 4 pages. Unless you think this is all good here's hell. So this is the terraino.org website, I know that some of you have worked with terraino in the past. Their trying to service an international community, so my god, select your identity provider; there's what 7—35 languages up there including RJ {indiscernible word}. Their languages I don't know, this is not where we want to be down the road, but you can see a hint of a better world and that this is a sticky experience. Notice this line that says, you are previously chosen to authenticated internet2 {indiscernible word} internet2. So rather than strolling down 875 sites in 1 point font to find internet2, I get a sticky experience up at the top. So we're working the union factors piece of this, though those of you in union factors land we need help because we're geeks. This is another thing that you'll start to see if you haven't already, which is the privacy managers. This has been an early version done by the swiss for what attributes has been released about you. And so your {indiscernible word} name is being released, your given name, a couple of other things from entitlement, the faithful don't show me this page again box. And you can confirm and this is compliant with the EU privacy directors. Here's a version that's in deployment at Brown University, same kind of idea, digital ID card, different attributes out there but it's visible stuff, I want you to know that. And then finally the collaboration management, this is quite of the tools we're developing for LIGO and iPlant which is the group piece, but let me escape from this and play a video from the Dutch that really explains well what collaboration management is. (Video plays) It's not a Dutch accent, but it is in fact. Ok back to the turtles. So, that's basically what's happening in the world. What I want to spend the remanding minutes on is talking about that we can leverage this stuff in other security areas. There's a new turtle in this stack. So I want to talk just briefly about each of these topics; wide-aperture security tools, role-based access controls, zero knowledge identity providers, federated network control, privacy management, and understanding the Tao of attributes. So, wide-aperture security tools, real-time correlation of security data across a federation of organizations and devices, stuff that we couldn't do a couple years ago because we had no trust. So I couldn't release all this information from my log files to be shared automatically with your log files because we didn't have an environment that allows that, we have an environment that allows that now. How do I begin to tap those tools, and it looks like a very rich space of tools. So we're doing some nice analytics, on that and then fordrop.org very interesting spot, which combines social networking with security analysis. Almost crowd sourcing security tools, where you can link files together in the nimble space that social identity allows you to do and then have people explore it but everybody's trusted inside that space because the identities, even if their anonymous are actually auditable and log back, so I can function anonymously but with security and people knowing that its really me. Role-based access controls – it isn't 5 minutes after federated identity when some wiki only goes – oh great I don't have to do accounts anymore that's really handy – but these are the people who want to access the page, do I have to list them all out, there's 700 people out there, 700 identities, no maybe group membership is a way of doing this, maybe scalable access control of it and then you'll replace your apple list of 700 names with a member of a group. We're able to start to develop those capabilities now; we're at the very early edge of that, we want to be able to extend that. I'm running through these for time. Zero-knowledge identity – so if you went to the NSTIC video and listened to it, it's a nice video, they make one claim that we can't do today. They make the claim that the identity provider doesn't know what you're doing, does not know where it's releasing attributes to or what attributes its releasing, that's not the way it works in our world today. But there are good technologies – Stephan {indiscernible last name} early work in PKI and zero-knowledge identity. We need to stitch that back into our fabric and then you'll have true anonymity in this space. So one of the things I sure like to see out there is real tools and library stuff. Federated network control, so {indiscernible name} spent years trying to get firewalls open so that grids and – we all know this one you know – and I was CIO when a scientist came up to me and said, "so would you open up the following 800 ports on your router?" and I'm going – yes and you go no no I'm not going to do that -- but maybe aww, maybe there's a procedure in the process. Well so we talked about the port poker years ago, the port poker is utterly viable at this point – you know trusted environment, can we get to it. Same with 323 or video conferencing makes it another thing that firewalls tend not to leap through. Wouldn't it be nice if on the basis on identity or even better on the basis on an attribute, I'm allowed to be able to open ports? Help us get those tools. Federated base level network access – some of you know about edge-u-roam. Wouldn't it be nice if we had edge-u-roam in this building around the US? We're working on it; we're working real hard on it. And then federated software defined networking – so back to some of the GENI references, back to some of the Open flow stuff. Are these people thinking about a scalable trust factor to underpin this kind of programming for routers? Are they? (Person from audience: Considering yes). Ok we need to talk as they say. Computer-aided privacy management – so one of the things about privacy is people want it but they don't want to manage it. And so um you saw those little boxes— um oh my you know – we're going to be faced a lot of here’s the minimum set of attributes I need from you, but if you give me your email address, I’ll give you a {indiscernible phrase}, I’ll give you extra storage in drop box, I’ll make some other inducements in return for giving you identity. Well, we need ways to have your past preferences in this space, generate suggestions about what the privacy release mechanism should be for this new site that you encountered. So we need little {indiscernible word} out there basically. And what we‘re learning is that attributes tend to travel in bundles, if I need an identity attributes, I probably also need a display name, I probably also need a few other attributes. So, we’re learning a little about the ecosystem, which leads to this last place understanding the Tao of attributes. So we ran a workshop in September 2009 that was just an existence proof. Yeah there’s an art to this stuff, there’s going to be ways of managing attributes in this system. The LOA of attributes – level of assurance in confidence that the attribute value that was assigned for you is really valid. There’s a lot of stuff for us to understand here, and we call it Tao because it’s not going to be a set of precepts, it’s going to be a way of approaching an ecosystem, we need to make that work. So moving forward – here’s the pitch. There’s lot of challenges out there; technical, policy, social and we need inspirational research grounded to the new infrastructure. We don’t want people to be inventive about identity management, we’ve done that, we want people to be inventive about the use of identity management to enable new capabilities in security, collaborations, etc. We’re going to need a connection of the research to some of the policy issues that are still out there, this is sticky stuff, I’ll come back to that in a second. And we’re going to need some innovative public-privacy partnerships to make it work. So technical stuff – I’ll just confess that I have not talked to a network security researcher in years and none of them have talked to me. We need to change that model, we have stuff to share with each other, tell me how to do that. Virtual organizations don’t know anything about identity management at all, and yet directorates in this building spend off new VOs and send them out into the world and these people go – oh I guess we have to do identity and access control – well golly in college I learned about something {indiscernible phrase} that must be what we were going to use. We need to cross pollinate, we need to work together. Um one of the things that would be really funny here is that um that um as we get into other federated security services, we’re just doing identity federations, but there the right campuses, there the campuses where GENI terminates, there the campuses where other kinds of services terminates. Once we get the CIO to sign a document, my god lets milk that for all we can. Let’s have appendix X which says we can share this information, and appendix Y that says we can share this information. I won’t get into policy because its alligators all the way down, and I won’t get into social because I sure don’t understand that but it was a great graphic from making it all work for the consumers. So to wrap up, my plea to this grand organization that has developed some – has allowed us to develop some useful tools, some programs in this space, some informed solicitations as I said before, wouldn’t it be nice if the solicitations are anchored in the infrastructure. Why did we build the cyber infrastructure for if not to support new innovation? Work across directorates and agencies – coordinating the deploying federal infrastructure. So I know there’s a star metrics project going on that is trying to think about some kinds of profiles for research consistently across agencies, we really need that. I went to a – one of the data net projects a while ago about a year ago at a workshop, and there were all these important data set owners sitting around the room and their talking about their data set, and each one begins with this following sentence—I'm going to ask researchers to fill out a profile of what their interested in. 17 profiles right in that room alone, and maintain it, oh yes, wouldn't it be nice if we had a single profile. That is the new single {indiscernible phrase}. There is work going on to make that happen, I sure hope it gets traction. We need schema consistency—NIH went out and tried to build an NIH person across their institutes as we all know. 330 attributes later, 17 email addresses attributes later, they put it up on the shelf and they said maybe we're not ready to coordinate yet. Well I need that to work not only at NIH but I need that to work across agencies because if I'm going to ask federated partners, the campuses to provide this information they don't want to keep 17 email addresses just for NIH and then ask NSF so what attributes would you like me to capture. So we need that coordination, we need risk assessments strategies for science. People are not doing good risk assessments on their instruments, on their applications to know exactly what level of assurance you might need. And then lastly, how do we support science projects that need to be {indiscernible word}. And I've had this conversation in the past about do you guys give money out to projects that say oh we're not doing identity management right, we want to do it better, we want to make it secure in privacy, this wont advance the science per se, it's not an instrument but it will enable us to advance the science moving forward. So wrapper—the internet identity total is new and distinctive—where's Waldo. It rebuilds some trust that the internet lost. This is the oldest picture I could find of turtles all the way down. Beautiful! Coming back full circle—it may not be turtles all the way down, I don't know the answer to that but there's enough turtles. That's the punch line, there's enough turtles and one of them is a federated turtle. It's got a big back, what can we put on it um what can we put on it, what can we leverage. From where I sit its turtles all the way back up, I sure hope we can have that. Thank you.

Audience member: indiscernible question

Klingenstein: So the question was where does the tortoise of DNS sec? Um and there's one other tortoise story I will mention after that. So um its very interesting and that, I think DNS sec kind of independent of this but once its n place, boy can we start to use DNS in ways that we haven't done before. You saw the {indiscernible word} where we were able to sign and secure email messages coming from this proxy etc. So I would say DNS sec is not a federated technology per se but once it's in place we can start to use DNS in a much more secure fashion to leverage federated approaches.

Audience member: So I would recommend taking a look to see how far along it is because it's well past the starting block {indiscernible word}, but well past the starting block. The other question I have is, what is the cost of using this technology, these providers where do they get their income and how does those cost get past to the users?

Klingenstein: So, I'll give {indiscernible phrase} an example, we charge about $3,500 a year for enterprise. If you take Penn State – Penn State's doing 80,000 transactions a day that is signed by the InCommon nevadator. If you {} the cost of $3,500 over 80,000 transactions times 365. This stuff is not costly, in fact we had a number of campuses come to us and say, if we charged more for InCommon could we get more campuses to fully implement InCommon, would you have more staff to do this. So, the cost is not being passed back on to the end users, the cost is low. I work for a networking organization that knows about real cost, that knows about leasing fiber, that knows about {indiscernible word} upgrades. They constantly look at InCommon and they go—well that's making some money, what does it cost? Well it doesn't cost anything, you have to provide {indiscernible word} to the federation and that's it. So the federations operation is very low cost. Um we're seeing that { indiscernible word}, even at Google, if you take Gmail as an example, 90% of the cost of Gmail to Google is not hardware, not software, not servers, not network, its password reset. When there's a privacy spill and city bank spills all those passwords on the floor, it's the same passwords that people use at Google, so Google does a password reset for all those things. That's where they hit the cost, so it's generally seen as a cost saving mechanism. So the return of investments, Penn State did a study early on, on this, and actually the sweeds did an excellent return on investment study here. And the return on investment is typically between 6-9 months, in terms of saved cost at the help desk if nothing else.

Audience member: in auditable question

Klingenstein: Well, so I was out at box.net last week, which services enterprises with um drop box types services. And they said so our business is enterprise and I said I'd like to changes that so your business is federated enterprise and the go we know that, we're just waiting to see that as a driver. The complexity of your identity management system is independent of what software and transport you use for shipping your identity assertions elsewhere. And so you may have a very—it's all like campus network. When I was selling the internet in 85, I didn't sell the internet, I sold my campus on building a campus network so those 3270s, we all remember those 3270s, could get down to the CDC that I was running off campus. And I said under my breath, "but you're really buying into the internet but you don't know that." My guess is now that 90% of the traffic on a campus network leaves the network. My guess is going forward that maybe 50% of the identity assertions that traverse the campus network now are going to be headed off campus in the future, because that's the way we face, we don't face on campus in our research community we face the rest of the world. So um you may spend a lot of money on business processes to create good attributes for your identity system, to {indiscernible word} students systems and your payroll system into your directories and infrastructure. Oracle will charge you a whole lot of money for that, but oracle gives away SAML for free or as cost to free as Oracle will ever get. So it's that external facing stuff that's just lost in the noise of the identity management piece. What was hard initially was getting campuses to sign the contract. Those first contracts take 6-9 months, now we go to a new campus and we go—well here's 250 universities that signed it, what’s so special about University of { indiscernible name}. Well we get to hear what’s special about them, but they sign it a lot faster then they use to.

Audience member: At the beginning you said that you spent a year trying to make sure you was using minimum, have you formally proven that what you have done is minimum?

Klingenstein: Not with a formal proof, and my background is a mess so I have longing for formal proofs. That said, we’re just watching invention after invention come back to SAML as the minimum set and that’s our confidence that we’re doing. But we went back and we said, “Oh God can I use a persistent identifier or different locations” and we came up with the attack vectors for privacy and other things. And I’m not a technical person, I’m marketing, I close other people product development. You know I kept on yelling at the product development people saying, “Awww that’s code that’s code” and they go “No”. This is going to be complicated enough that lets assure ourselves that nothing simply will do. So all I can say is we spent a year wandering in the waste land to convince ourselves, and we got convince and now we’re seeing other people wander through the same waste land and come back with SAML.

Audience member: Mr. Ken thanks a lot! I think that was a really good overview of this and I’ve been thinking a lot about what we should be doing to develop a national cyber infrastructure. First of all just as a scientist out there, then as the head of the OCI, and now as the head of the directorate that probably serve the most scientist who use the national cyber infrastructure and it’s still very very hard and I think InCommon and federated identity management would make a huge step forward in terms of making it an integration between the campuses where scientist really work and the national cyber infrastructure however we define that. So do you have any suggestions on what we can do as an agency to make this entire thing move forward so that, for example InCommon is not just agreed to but actually implemented on campus, especially now that we have things like not only terragrid but and following on exceed is going to use this as sort of their way of campus bridging to connect into the campuses and so on. So, at NSF we have a lot of influence but there’s certain things that just aren’t still getting done and so what would you recommend that we actually do to make this actually happen?

Klingenstein: Couple of ideas. First of all I want to move the conversation from identity to access—identity and access control because that’s really the full story. And then in that space I think—so the engagement we’re doing with iPlant and LIGO is just informative for everybody and somehow that word needs to go out. If you look at the surfnet video, surfnet is doing a service offering for all of escience, so that if you’re an escience group you immediately get a collaboration platform with whatever applications you like—Scarlett, if your wiki needs mathematical notation we’ll go to t- wiki, if it doesn’t need mathematical notation we’ll go to this wiki, etc. {indiscernible phrase} for the rest of the community there only providing the thin identity and access control. I don’t think—at one point I had the dream that math-physics would set up a collaboration platform, but you guys don’t want to run infrastructure that way and it doesn’t really make sense. So I suspect that you can take assuming, we have success you can take and highlight the successes. Physicist talks to physicists, I guess they do especially if they're in the same domain and get some cross pollination. We're talking about running an international VO camp next summer, where in fact will get some of these international VOs that need to deal with lots of different federations with the federated operators all in the room. If we can get 3 or 4 spectacular successes or even good success, I'm hopeful that word of mouth will convey this. One other place is writing {indiscernible name} solicitations and this has been a long time concern of mine, um I look at jisk and again we've talked about this in the past. Jisk writes really tight solicitations, this is the infrastructure here is where you innovate. I've seen NSF try to do that but sometimes it doesn't work. But I would argue that some solicitations that say you know here's bedrock, be really creative above that, some—you know, I'd like to see it that would work, so something about the solicitations themselves might be beneficial.

Audience member: I work in elections and voting and one of the really key important things there is that any Tom, Dick, Harry, Jane etc., is able to understand if they want to what actually goes on. I'm wondering whether that applies in this case as well, is it important that the user of the federated identity system is able to understand that in fact all of their identity information isn't going with them when they go to different sites. I'm not just speaking technical whether it is or isn't, I'm asking from social policy perspective, do we care whether they understand and do users read any of the release dialogs that you say we're going to pass this attribute on but not that attribute, the results of people installing apps are that nobody reads those { indiscernible phrase}.

Klingenstein: So there is a wonderful document that we've been working on called "Putting the Informed into Informed Consent" and it happened after some work that the UK consumer and privacy office did. Gorgeous stuff! All 1 point click through suck alright and they need release. There are good solid templates from an English speaking country about how you can write informative click through and give the people the opportunity to consent or not consent. We don't have enough deployment here to know what the track records going to be but we do have deployment histories in several European countries and people do know what's happening. {Indiscernible phrase} Last things is, I count on Facebook to continue to give us teachable moments, and as they give us teachable moments we need to be out there going it doesn't have to be this way.

Keith Marzullo: Ok that was the last question, why don't we wrap up. Thank you very much.