Di **[00:00:49]**: We have talked about some of the XACML [eXtensible Access Control Markup Language] architecture components that have been implemented in the Orange County pilot – the JUICE [the Juvenile Information Content Exchange] pilot – and also in the CONNECT pilot.

[00:01:03]

There are some other components of the XACML architecture that, to the best of my knowledge, have not been implemented yet in the public sector. I'm thinking here of the policy administration point, the PAP, and the policy information point, the PIP. I'm wondering if we could talk for a little bit about what you might see on the horizon in terms of a business need for the PAP or the PIP and what kinds of information sharing might necessitate their implementation in the public sector. So John, would you like to start us off? What do you see as the business purpose behind the policy administration point?

John: I think the policy administration point is really the registry or a repository for where, as you author your different policies for your different exchanges, you store it there.

[00:02:00]

You can keep track of versions, whether this is a test version or this is the version that you're actually deploying into production. You can use it as kind of a control point. When someone wants to bring up a new service or a new exchange, you can have as one of the steps that the policy on that exchange is registered with the policy administration point. So I see it as a management tool, as a discovery service for what policies are already out there and what rules, which may apply to the particular exchange that you're working on.

I think really once you go beyond a couple of pilot projects, most of the toolsets out there offer a policy administration repository and the user interface for the administrators to use. I really think just about all organizations would want to have a policy administration point.

[00:03:00]

It's also used by the policy decision point to retrieve the particular policy, so in terms of actual transaction processing, it has the

interface for doing that. I think it's a key component. When you're first starting, you can probably just write a policy or two and handle it on your own without actually having an actual registry. With regards to PIP –

Di: Do you mind just stopping there for one second? Maybe we'll just drill down a little bit with some of our implementers. What do you see as a potential application of the policy administration point? Mike, I think you might have mentioned that you can kind of see that as becoming a need for CONNECT in the future?

Mike: Just to follow up on John's statement, I think the whole notion of having rules and how they apply, particularly when we're talking about multiple jurisdictions or multiple states, and we have this being driven by statute, when statutes change and when we have to tie it back to our auditing logs and everything else, being able to easily track through those rule changes, when they apply and what activity related to those changes can be really key.

[00:04:06]

We have to be very aware that it's not just there so that we have a log. We're going to have to be able to adequately and appropriately and easily use those logs, use the tracking of the administration rules. So I see it tying in definitely with the exchanges across the states and any potential issues that might come up with use or misuse.

Maury: I'm sure with the growing potential for more and more data that is going to be coming into CONNECT, along with the diversity not only across state lines but within our states about where the data custodian has the data, if we could build this common interface using an administration point like this, a common interface, a common approach towards training people how to apply that toward the service there, the access service, I think that would make administration across the board much easier for a very complex system.

[00:05:04]

Tony: And then there's two other aspects of that. One is the growth of this. We're finding that we've got a couple of policies in place with some data resource owners, and somebody comes up and says, "Gee, I'd like to register my resource with you, my data, but I'm not sure about the policy." Having those allows us to

streamline new rules. And the number of these rules is going to grow, as you were saying, without limit. So we need to be prepared to sustain that kind of motion, because if somebody approaches us to register it, and we say, "Gee, that's a six-month or a two-year process," we're not going to have the adoption that we need to have.

It needs to be easier to do the data exchange via the ***formal*** process than the back door, because the side doors will occur if you don't allow adoption.

The second part of that is that I think in the future, we would see that an entity might share its data with two or three consumers through separated identity providers.

[00:06:04]

And so by having some standardization about the policy, it makes it easier to make sure that Identity Provider One isn't using a different policy than Identity Provider Two, because your users will find the one that gets them where they want to be, so you need that regulatory consistency across.

Di: When you envision the architecture for the PAP – let's take CONNECT, for just one example – my understanding is that in CONNECT, you have a distributed policy decision point, one at the central portal, but also policy decision points next to each state's data source, right? Do you imagine that you will also have a policy administration point next to each data source so that each state can administer its own policies? Or do you see that being as more of a centralized function?

[00:07:07]

Maury: Certainly we would like that to be right alongside the data repository. Right now, we're not using the PAP. It's a XACML policy file that we've created. We have the ability of the chief point of contact within each state to manage that directly for all the data sources, because right now it's fairly limited to a few data sets, so it's not that difficult to deal with. But as we grow, certainly it would help tremendously, I think, for the administration moving forward and then for the versatility of what we can do with it. This is one of those areas we need to move toward.

Di: Tony, can you think of – I know I'm asking you to predict the future and it's hard.

[00:07:56]

But can you think of some best practices or some disciplines that we're going to need to adopt in order to ensure that we're really maximizing re-use of existing policies to speed up that kind of "time to market" for new exchanges? What do you imagine is going to work well in terms of –

Tony: I think if we can get all of the providers – and again, it's always that idea of "can we influence the market?" – I don't know – the commercial market, the vendors, if you will – to standardize across the way they deal with it. You mentioned having the policy piece associated with the data, and having the policy piece associated with the identity provider consumer site. I think that check and balance is a good architecture because it does allow for the multi-use regulatory. But at the same time, I think going forward we need to try and have the vendors standardize on the uses of the terms associated with their policy pieces.

[00:09:00]

Otherwise, this is going to be chaos, because you're going to define ABZ over here, and it's going to be ZAQ over here, but it's going to be the same, and I think that creates confusion for our users. So looking out toward the future, that would be something that would be an asset to the community if we can try and align and publish those standards, as I think you're working on.

Di: Mike, do you have anything to add to this idea, what some of our practices need to be to maintain standardization, which reduces complexity, which –

Mike: Whether we call it a federation or we call it a local implementation or a multi-jurisdictional implementation, I think the clarification of what it's going to mean to talk to each other across those identity providers is key. So I think the standards are going to be the language, the terminology, the commercial implementation, the following a standard is key, but I think it'll evolve.

[00:10:01]

Tony: I think you're already taking steps in terms of focusing on some of the successful projects and pilots and sharing that among the

community, because I think one of the best things is for that organization who's just starting to look at something that's been done, because then that leads to a natural standard. It's not necessarily a formally approved standard or a regulation, but people will do what they have seen others be successful at.

Di: Anything else about PAPs? You want to switch gears to PIPs?

John: A little PIP.

Di: A little PIP. *Great Expectations*, Charles Dickens, right?

John: You got it. That's right.

Di: Okay, so another component of the XACML architecture that has not really been implemented yet in the public sector is the policy information point. My understanding is that this capability can go out into the external world and fetch data about some environmental conditions that are relevant to making that access control decision.

[00:11:02]

John, maybe we could start off just by trying to imagine what a use case for a PIP implementation would be in our sector.

John: Well, let me take a – attorney/client privilege is something that the records that are being kept for a particular attorney on his client are confidential, and they're not accessible to other attorneys. I can see the case where you have a database with a lot of clients in it, and on that database there's a record that says, "This is the attorney of record for this client."

[00:11:54]

So what would happen is if I'm Attorney Jones, and I try to access someone who's not my client in that database, part of the policy is this attorney has to be on the record in the database as your client for me to grant access, so the PIP would have to actually access the record and read that particular attribute on the resource, and it's part of the matching logic between – so the requester, when it comes in, the policy says, "You have to be the attorney for this client," but the policy doesn't know whether that's true or not until it actually goes outside to a data source and reads and finds that. That's an example of what we call fine-grained authorization.

[00:12:51]

The other use case that I can think of is I may have a rule about me, the requester, the attorney, that before I can access this client record, I have to be verified that I am a good-standing member of the Bar Association in the state of California, for example. And so the PIP has this policy that says, "Is he in good standing with the State Bar?" Well, I didn't declare that I'm in good standing when I sent my request over, but the PIP can make a request using my Bar number up to another resource – external obviously to the policy – and check and say, "Oh yeah, he's not debarred right now." It comes back. It says, "He's in good standing."

Now my policy says, "Okay, and in good standing, therefore I will grant access." So those are two examples of where the policy information point is doing the job of getting information that isn't available to the policy decision point without going out somewhere and getting that particular information.

[00:14:04]

Tony: See, we're currently accumulating that information in a profile, but it's self-professed, right?

Di: I'm telling you that I am in good standing.

Tony: I'm telling you I'm in good standing, and we actually have an e-mail process that verifies it with whoever your sponsor was, because all of our users have to have a sponsor to be in the identity as a requestor. So someone in our workflow approved you to be there, so there's a random e-mail that goes back and says, "We're verifying that this profile looks correct to you as the sponsor," but ideally we ought to go back to the main source, as in the case of the Bar.

Di: In the CONNECT project, Maury, Mike, do you see maybe some future need to implement a policy information point?

Mike: We have some similar things relative to status of officers, whether or not they're working for – in Nebraska, we have a lot of rural agencies.

[00:15:04]

We have officers working for multiple agencies. We might have different roles for different agencies, an intell officer or contact point with the fusion center, those types of things. So in talking about it, I could almost see trying to get back to those kinds of roles. Similarly, and the whole notion of supervision with parole officers or probation officers and whether or not someone should be able to get at medical information might be restricted only to their current supervisor or their current probation officer. That seems to make sense. We haven't thought about that.

But then, you can go back to the natural – to some extended database or something to see who actually is providing coverage. That makes sense, I think.

Tony: In our project, we used that – in the public defender's office, for example, "Are you on juvenile or are you on criminal or adult?" So there's a flag, and that is set by the individual, but it's validated by that individual's supervisor. "Yes, you're still on juvenile."

[00:16:01]

Again, that's on our side, not on the database owner side, so it's again validating that granularity of access right.

Di: So there are relationships between requesters and the subject of the data resource that they seek. Could you think of any examples with environmental conditions? One that's frequently put out there as a use case is you would not normally have access to this data, but if we're in an emergency situation, like a tornado or a hurricane or an earthquake, then there might be larger access. Do you see any of those being implemented?

John: I've certainly read the research on the "break the glass," is what they call those policies.

Di: Break the glass.

John: The actual implementation of those, we haven't done any of those yet.

[00:17:00]

But it would basically represent here's Policy A, which is the norm; here's Policy B, when there's an emergency. Basically when an emergency is declared, when the requests come in, and

say, "Okay, and this is an emergency," is basically a status, an environmental variable, so it's now going to point you over to that policy instead of the normal policy you go to. That's how I see it would be implemented.

Tony: Yeah, we implemented a role we call the EOC role. It's the emergency operations center. In our county, the way we staff the EOC is, in the case of an event, certain people have a designated secondary job, if you will. My secondary job is I do this at the EOC, so you have your primary role, which might be HR [human resources], or as an entity, you might be in law enforcement or in criminal justice as an attorney or a doctor, but then you have your secondary role. If your secondary role is EOC, then once you're in that EOC role, you have access to a lot more in our identity manager than you do in your normal role, which is very channelized.

[00:18:03]

That's just been experimental, and we've only got like 45 people, but we were looking at that, because we had an event and everybody said, "Gee, if I only had access to that, and I know it's in your database, but I couldn't get to it, and I was operating up at the hill during an emergency. Why is it I couldn't get that information?" So we said, "Okay, we can give you a secondary role."

Maury: This dynamic authorization opens so many doors to allow new possibilities to maybe not just access data you didn't have before, but maybe take an extended view of data you would have access to. There are certain situations that could occur, and now there's – all the data that we deal with is usually very, very deep. The data custodian usually is the one that deals with it, because it's their business process, and it may not even make sense to other people.

[00:19:04]

But in these types of situations where the role extends or you have some unusual circumstance where you have to have other knowledge, this makes perfect sense to take advantage of this technology, that it truly extends your access in an unusual way.

Tony: To give those watching a specific example that follows that, in our EOC role, what happens is you can draw a circle geographically and you can say, "I need to know all the doctors and nurses within

that zone," and we ended up needing that. And normally healthcare would be the only one who could see the geographic location of the private residences of our staff within the county or where they're out of in the clinics. But when you're in that EOC role, you now move from the healthcare view to being able to see how that overlays. So when there was an event that closed down a community, we said, "Gee, that's okay, because we have this healthcare population already there to service it."

[00:20:02]

Di: Anything else about policy information point that you'd like to bring forward? Okay. Well, then let's spend the last little bit of this segment on obligations, because I know that obligations are near and dear to John Ruegg's heart.

So just as a little bit of an introduction, the idea of an obligation in the XACML architecture today is that the requestor or someone in this transaction is going to agree to do or not do something, as a condition of being granted access to the data they're requesting. But sometimes in our community, that act or that agreement not to act in some way extends beyond the time limit of that moment that the transaction is granted or denied.

[00:21:04]

So, first of all, John, maybe you could give us some real world use cases of where that obligation is actually delayed, just so that we understand what we're talking about.

John: Sure. First, I'd like to say that within the architecture, obligations are the way – the specification is the requirement of the PEP, the policy enforcement point. When we're looking at various exchanges, a real common obligation would be whenever you access a record at a resource, after the PEP has determined that you have the right to read, update, delete, or create a new record for that particular resource, that you audit that particular access.

[00:22:01]

So getting access is basically – that is the transaction. The obligation is, "Oh, you also have to create an audit record." So there's an example of the service provider or the data provider having one or more obligations.

Another obligation they may have is for that particular record, there's another party that needs to be notified that that record is being accessed, so maybe an obligation would be to send an e-mail to a particular organization, entity, investigator, doctor, patient, etcetera. So those are a couple of examples of obligations on the service provider side.

We also have commonly, in the way that the policies that are being written, obligations on the consumer. So for example, we may provide access to criminal history records to a research firm.

[00:23:04]

And they're interested in a particular population. We will have an obligation on that research firm that within an event, such as completion of the research or within one year or within six months, you will destroy this source information. That's an obligation that's really on the person who's getting a copy of the information. The PEP on the service provider, he can't execute that obligation, but what we're looking at is he can do a notification of the obligation to the consumer.

So now the consumer says, "Okay, well, I got this record" – and by the way, I have associated that it comes right back with a message, a directive that says, "This is how I have to handle this record, how long I can retain it, when I need to delete it, acceptable use of the record, etcetera."

[00:24:11]

So what we're working on trying to get down to the point where we could actually use XACML to package up an obligation to send back along with the record that you've requested, so that then they could store that obligation, and then their systems can read those and process those.

So those are the two kinds of, sides of obligations. In terms of Global, we're working on what the most common types of obligation verbs there are out there, and then you would substitute the nouns for which particular kind of resource, so no secondary dissemination, redact these particular elements. We require you to encrypt this record once you receive it. Do you need to notify, retain?

[00:25:04]

So these are common verbs that pop up when you read all these different policies, and then the nouns are the kinds of resources that you're sharing would be the objects of those verbs. So, "Redact this criminal history record for PII [personally identifiable information] before secondary dissemination," might be an example of an obligation. That's my little talk on obligations.

Di: Tony, are you looking at these kinds of obligation in any of the work you're doing in Orange County right now, where there's actually like a delayed enforcement?

Tony: We currently have two, if you will. In the case of healthcare information that you as a consumer have gone through the identity manager to get access to the record, it reminds you, after you've got the record, that your obligation is, if you change your business relationship, you're no longer part of an active BA [business associate] contract with that proprietor, that you must notify us of that change in status.

[00:26:11]

But it's really not as mature as it needs to be. I think we're anxiously looking for that sentence structure to come down so we can mature it. It hasn't been the focus, but it's been something we knew we needed to do.

Di: That's my understanding as well, Tony, is that some of these delayed obligations right now, they are managed primarily through some kind of a legal agreement with some additional reminders that you just described. But what I hear you saying is that everyone would sleep a little better at night if in fact there were some more technical enforcement of those, at the moment that, for instance, that person's relationship was changed.

Tony: The challenge in obligations for us is that things like data retention schedules are so loosely defined with the organization.

[00:27:03]

Yes, we are compliant; however, the enforcement of those schedules has always been on the data repository, not on the consumer.

And it hasn't been in a way that the data – how do I say this? – that it can scale outward.

If you've got all those boxes in your warehouse and you say, "Gee, every six years, I go clean out those boxes," that's a whole different retention process than if it's out in a disseminated community and you need to follow it up. Again, that's the challenge we're having is that some of those obligation areas have not been boxed in as nice and clean as other things we've talked about.

Di: Mike, not necessarily even from the CONNECT project's perspective, but maybe your day job as NCJIS [Nebraska Criminal Justice Information System], do you see any business needs for these kinds of obligation handlers?

[00:28:01]

Mike: In listening to John, I started to think about the intelligence community and fusion centers. Talking about obligations and people searching for data and – de-confliction is the term typically used on that data. You need to either notify somebody that the data has been accessed or that somebody wants access to the data, and you need to grant access to it. It seems like that whole query of the obligation could be set up to either notify someone to go ahead and provide access to the data, but notify them that the access has been made, to request access, to do things behind the scenes where the user wouldn't know it at all.

It seems like a straightforward scenario, but I also started thinking about the PIP because it seems like a lot of that data could have restrictions based upon who the requester is. You could actually make an outreach then to the PIP, try to make some decisions based upon the type of request, the type of data, the current status of the data, whether or not it should be redacted, so some of these things could actually come together.

[00:28:57]

But I started thinking about those types of really secure data and the different rules behind the data that people try to remember in the real world, let alone document and be able to automate.

Di: Maury, from your role in Alabama, do you see a need for some obligation enforcement?

Maury: Well, the possibilities – I must admit, this is probably the least familiar I am of these XACML technologies.

Di: For all of us, it is.

Maury: But where this could go is – it's got so many new ways to do business. In fact, I'm thinking the efficiencies we could build into our data sets now that we – depending on the type of data that you query and the results that come back, knowing that those can carry obligations or trigger notification, that tremendously is helpful because so often – I'm thinking in the role of the officer, depending on what they get, if they're pulling somebody over, and they have their standard protocol.

[00:30:03]

They're going to look up a tag or try to find out who's possibly driving the vehicle, but we've got so many layers now of data flowing in. Are they potentially wanted? Is there an outstanding warrant? Are they on some watch list? Or something to that effect. Immediately, I'm thinking the obligation, trigger the notice associated with the particular returned result, say, "Immediately call ATF [Bureau of Alcohol, Tobacco, Firearms, and Explosives], and here's the phone number," because this is the search result you returned based on the information that was in this file. There are those kind of notifications I believe we could take advantage of in much more dynamic ways again than the traditional ways we've sort of piece-mealed capabilities in the past.

I think this is again a different way of thinking, a different way of looking at our information systems and the data sets we have. It's so logical, and it's going to have so much benefit in the long run.

[00:31:04]

Di: So, no pressure, John, on your [Global] working group to go out and "make it so, Number One!"

John: Well, it does boil down to having again a common set of verbs and vocabulary that's common vocabulary for the resource so that you can communicate these things, and the systems can consistently know what to do with those particular directives.

Tony: What I'm hoping is that we get that as part of what we call the registration process. We're establishing this new data source. Here are all the things you need to know about that data source in

terms of its fields, its restrictions, its need for encryption, its obligations, its regulatory compliance because that's the time – what we found is people said, "Oh, well, just hook it up to that," without really having that whole context.

[00:32:00]

We're used to seeing it as just data, not as part of an overall intelligence where multiple data [sets] are going to come together.

Maury: This helps – we can't expect the typical user to know everything. We're inundating users now with information, and so if we can assist that end user with helping them understand use of data as it comes to them, instead of expecting them to know all at all times, then that's going to help them. It's going to a help the integrity of the data. It's going to help the privacy of the citizen, and so forth.

Tony: They want to do the right thing. We just need to provide them some hints on what the right thing is.

Di: Because it is very, very complex out there.

John: I think this comes back to the policy administration point, in a way. If you have a structure for writing an exchange that includes policy and contract agreements, you can go back to that section when you discover other uses or you have other obligations that, for this new community, need to be implemented.

[00:33:04]

You can then add that, and you're adding that to this external rules engine, as opposed to going back to a programmer and saying, "I want you to code me up a new rule because there's been a change in the regulations, or we've decided we're going to exchange some information with this new group out there." So having authorization and identity as manageable, exterior to the actual data resources functions, is going to really facilitate the maintenance and the speed, etcetera, of adopting these requirements.

Di: Do you see anything else on the horizon that you'd like to add? We covered PAPs, PIPs, obligations.

John: On the horizon is a lot of education and training and kind of building a small snowball and then getting it rolling down the hill.

I think this is years of work ahead of us. This is an evolution. So I guess, on the future horizon, I don't see any lightning jumps to this place we're talking about, but I do see us creating an architecture and a vocabulary to support that moving ahead.

Tony: I think the piece that I'm hoping comes in the future here is this toolkit that allows you to move from legacy to fully integrated because that's the challenge we're seeing is we have a lot of legacy systems, and getting each of those to now integrate is a challenge.

[00:35:01]

So we've been working on this checklist, this readiness assessment, for each of those legacy systems, because we can't expect to go back and change it. It needs to be that we put a wrapper around it or we come up with some other standard. But again, that's part of this registration process, I think. But this is a difference in that we've got to have the registration process, but we've also got to have a methodology by which we incorporate those things.

John: I think that's a really good thing to share that your legacy systems have all of this logic locked into them. Between the legacy system and your requestors, you add this architecture. It's when that particular data exchange is defined, you know what pieces of information are going to come out, and you now know what policies you want to apply to the information. You're kind of intercepting it and actually applying your controls, external to the internal.

[00:36:01]

Even if the internal application has already done some authorization, you can extend the rules outside of the legacy application, as opposed to trying to say to the application, "Strip out all your authorization logic. We're going to put everything outside." I think that's a very key design approach, to how to address the legacy world.

Di: Can you imagine that there will be a toolkit, as you put it, Tony, do you imagine that toolkit would be reusable for like legacy systems, right? Is there going to be a small number of those, like maybe a half a dozen, for handling different kinds of legacy applications, like a COBOL or a SQL?

Tony: So far, we've come back with seven. We expect it to be less than twenty.

Di: Okay, good.

Tony: Right? So it's manageable. But ours is still every intense in terms of the conversations.

[00:37:05]

We're still having to get a coder or a developer to talk to the legacy person, and I'm anticipating that we'll eventually get it where it's more of an automated tool above that. But it needs to be streamlined. It needs to be focused, much like we've done the architecture, is this methodology of integration.

Di: Everyone would agree with you that the idea of ripping and replacing legacy applications – that's just not possible in our current budgetary environment, right?

Tony: Well, to give an example, again, a hard-code example. We found that 70 percent of the legacy applications we wanted to integrate with had internal databases, typically a SQL table, and so we worked with our vendor to create a connector type, which is a table interface.

So now, it's a lot simpler conversation to go, "Okay, do you support SAML [Security Assertion Markup Language]? Do you support this or this? And oh, we support SQL." "Great! Let's just get to your table structure, and then basically we have the table connector."

[00:38:04]

But that's solving more of the data level. It's not solving all the obligations, the rights, the encryption, etcetera, which also we need to work on. So again, it's crawl, walk, run, like you said, the snowball. We're starting small, but I think that's going to be a growing area in the future.

[00:38:20]