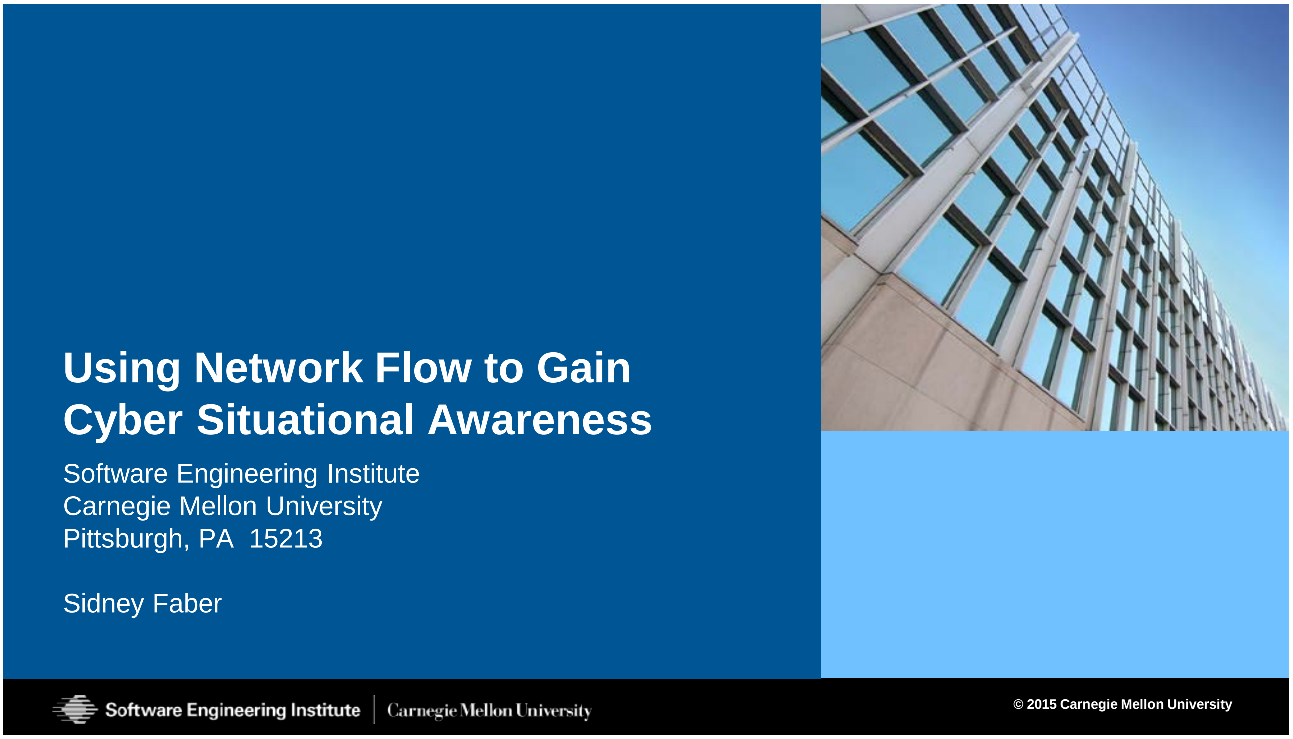
Using Network Flow to Gain Cyber Situational Awareness

Table of Contents



Using Network Flow to Gain Cyber Situational Awareness

Software Engineering Institute
Carnegie Mellon University
Pittsburgh, PA 15213

Sidney Faber

Software Engineering Institute | Carnegie Mellon University

**003 Shane McGraw: And hello, welcome to the campus of Carnegie Mellon University in Pittsburgh, Pennsylvania. We welcome you to the Software Engineering Institute's webinar series. Our presentation today is using network flow to gain cyber situational awareness. Depending on your location, we wish you a good morning, a good afternoon, or good evening.

My name is Shane McGraw. I will be your moderator for the presentation. And I would like to thank you for attending. We want to make today as interactive as possible. So, we will address as many questions as we can throughout the presentation, and again, at the end of the presentation. And you can submit your questions at any time to our events staff by using the questions tab on your control panel.

We also ask a few polling questions throughout the presentation. And they will appear as a pop-up window on your screen. The first question we'd like to ask is how did you hear about today's event.

Another three tabs I'd like to point out are the files, Twitter, and survey tabs. The files tab has a PDF copy of the presentation slides there now along with other network flow related work from the SEI. For those of your using Twitter, be sure to follow @SEInews and use the hashtag SEI webinar. Once again, that was @SEInews and use the hashtag SEI webinar. And lastly, the survey tab we'd like you to fill out upon exiting the event as your feedback is always greatly appreciated.

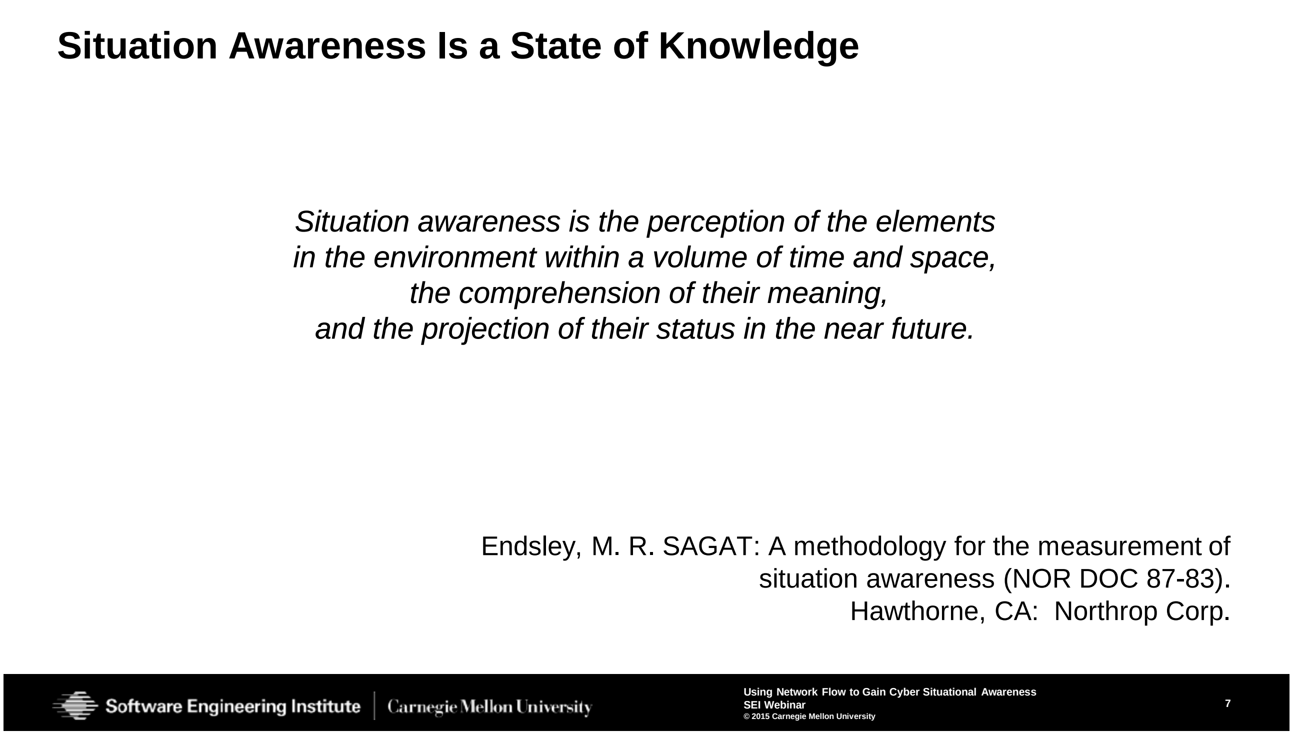
And now, I'd like to introduce our presenter for today. Sid Faber is a senior network analyst on the situational awareness team that is part of the Software Engineering Institute's CERT division. He works with both federal and civilian government agencies to secure, operate, and defend the nation's critical networks. And he runs the annual CERT FloCon conference, which is an open forum for sharing large scale network analytics. Faber also serves as an adjunct faculty member at the Carnegie Mellon University Heinz College of Information Systems and Management.

And now, I'd like to turn it over to Sid Faber. Sid, welcome, all yours.

Sid Faber: Thank you, Shane, and thank you to everybody for joining us today. As Shane mentioned, we'd like your feedback. And in order to kind of shape today's presentation, we've added a few polling questions at the beginning to figure out kind of where you're at and how familiar you are with situational awareness and also to understand what your background is and where your interests lie today.

I think probably the most important thing to begin with is a definition of what situational awareness actually is. So, let's start there. In our experience at CERT, we've been looking at network situational awareness for almost a decade now. And we've come to center around a definition on situational awareness that was proposed by Dr. Mica Endsley.

Situation Awareness Is a State of Knowledge



Situation Awareness Is a State of Knowledge

Situation awareness is the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future.

Endsley, M. R. SAGAT: A methodology for the measurement of situation awareness (NOR DOC 87-83). Hawthorne, CA: Northrop Corp.

Software Engineering Institute | Carnegie Mellon University

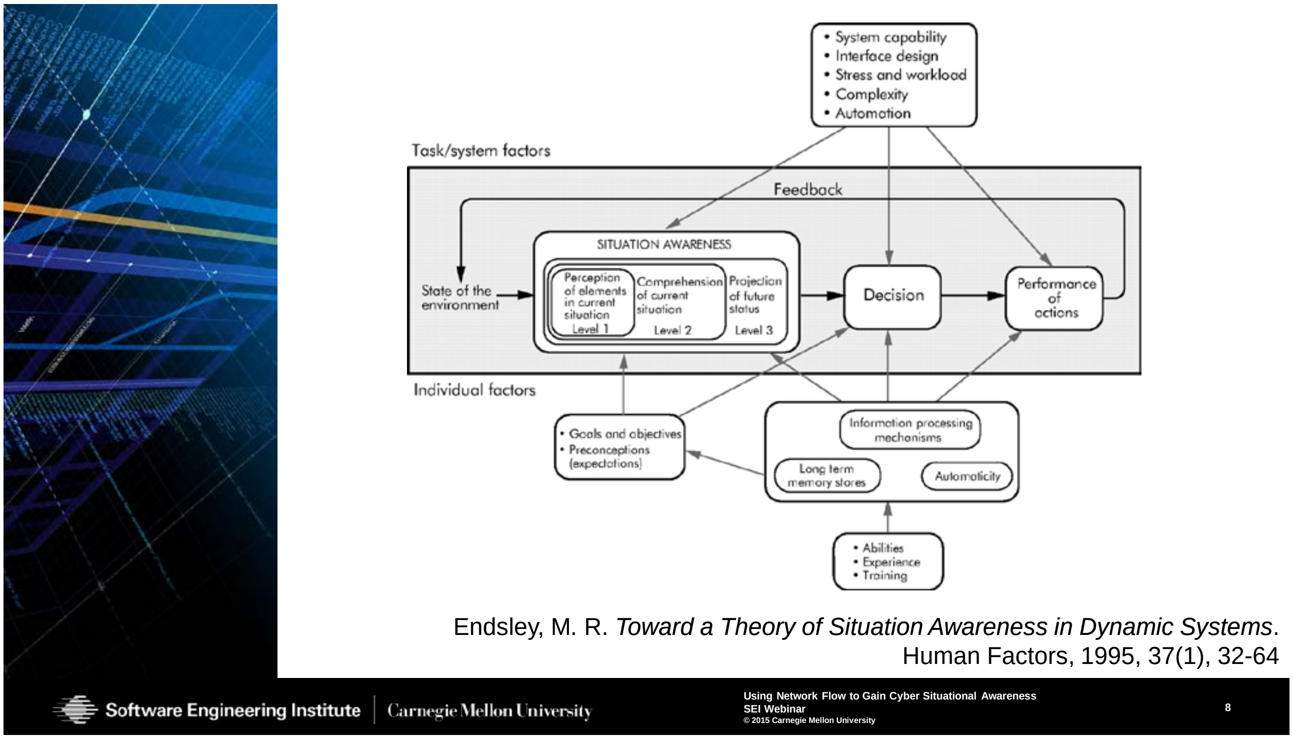
Using Network Flow to Gain Cyber Situational Awareness
SEI Webinar
© 2015 Carnegie Mellon University

7

**007 So, our first introduction into situational awareness comes from her definition. And that's that situational awareness is the perception of elements in the environment within a volume of time and space. The comprehension of their meaning and then their projection into the near future. So, as you can see, there are three kind of focal areas here. Dr. Endsley defines it as three different levels of situational awareness, perception, comprehension, and then Projection, and we'll follow these themes throughout.

There are other definitions that are out there and that largely circle around this one. But we find this one probably most relevant to our area.

Factors



Endsley, M. R. *Toward a Theory of Situation Awareness in Dynamic Systems*. Human Factors, 1995, 37(1), 32-64

Using Network Flow to Gain Cyber Situational Awareness
SEI Webinar
© 2015 Carnegie Mellon University

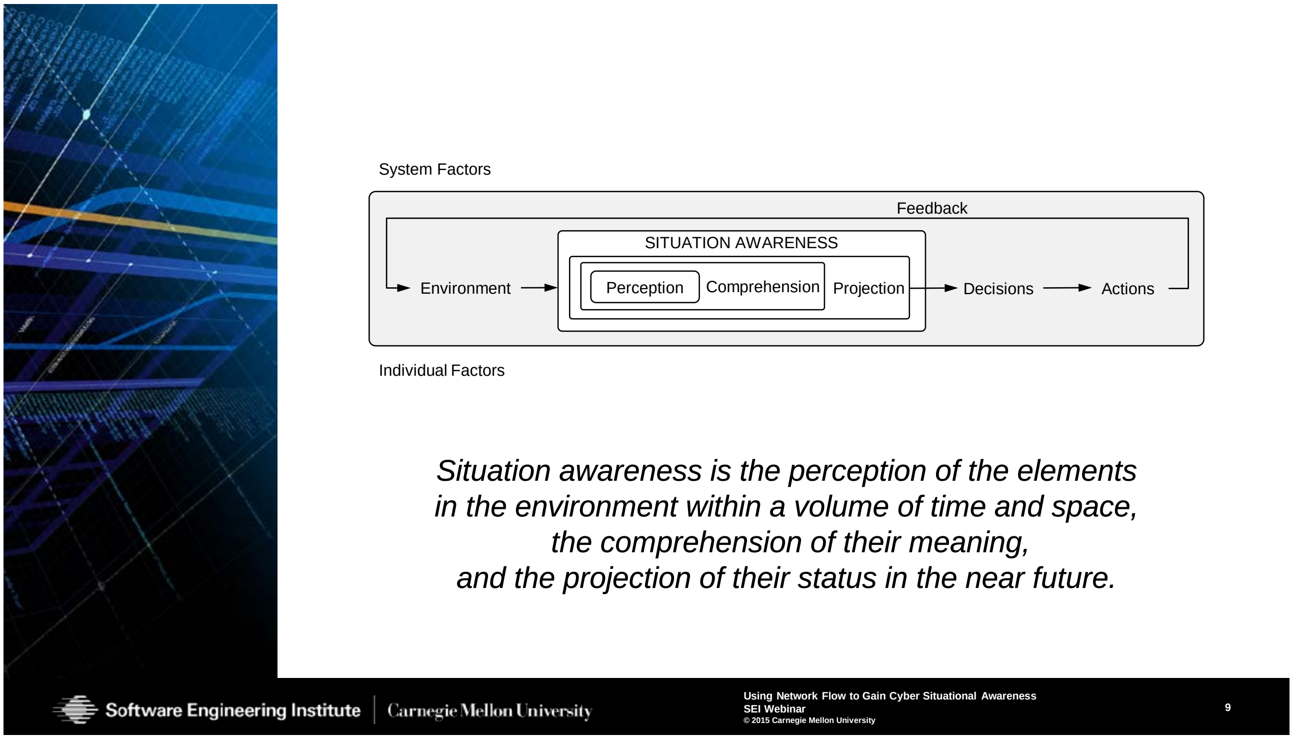
8

**008 Along with that Dr. Endsley proposed a formal format, this model, for gaining situational awareness. This model was first presented in 1995. And that's one of the attachments, I believe Shane, that's with our webinar today.

Shane McGraw: That is indeed there now, yes.

Sid Faber: So, you should be able to download this. It's kind of a seminal document I feel on what is situational awareness. And it proposes and describes this model. This model has a lot going on. We won't be able to go into everything today.

Factors



Situation awareness is the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future.

Software Engineering Institute | Carnegie Mellon University

Using Network Flow to Gain Cyber Situational Awareness
SEI Webinar
© 2015 Carnegie Mellon University

9

**009 So, I'm going to present to you a somewhat simplified version of that so we can focus in on the things that are of most interest to us.

As you can see, there are those three different levels of situational awareness in our model, that level one perception, level two comprehension, and then level three projection, and how they all fit. And perception leads to comprehension or understanding. And that understanding leads us into projection or predicting something in the near future about what's going on. That's situational awareness.

Look, however, on the right hand side how situational awareness actually feeds that decision making process. Decision making is in support of taking actions. I think that's critical to keep in mind that we're looking to make some sort of action, take a decision that makes an action. That action should be aligned with our goals. That's successful situational awareness, when we achieve our goals by taking action.

Also, if you take a look on the left hand side of the model, you'll see that there's the environment. And the environment provides input into the model. So, with that, we need to sense the environment, understand the environment, so that we can actually perceive what's going on around us.

How does that actually work? Well, let's take a very concrete example. Shane, let me ask you, it's a November morning here in Pittsburgh. It's kind of chilly out today. Did you wear a jacket to work this morning?

Shane McGraw: I did wear a jacket today.

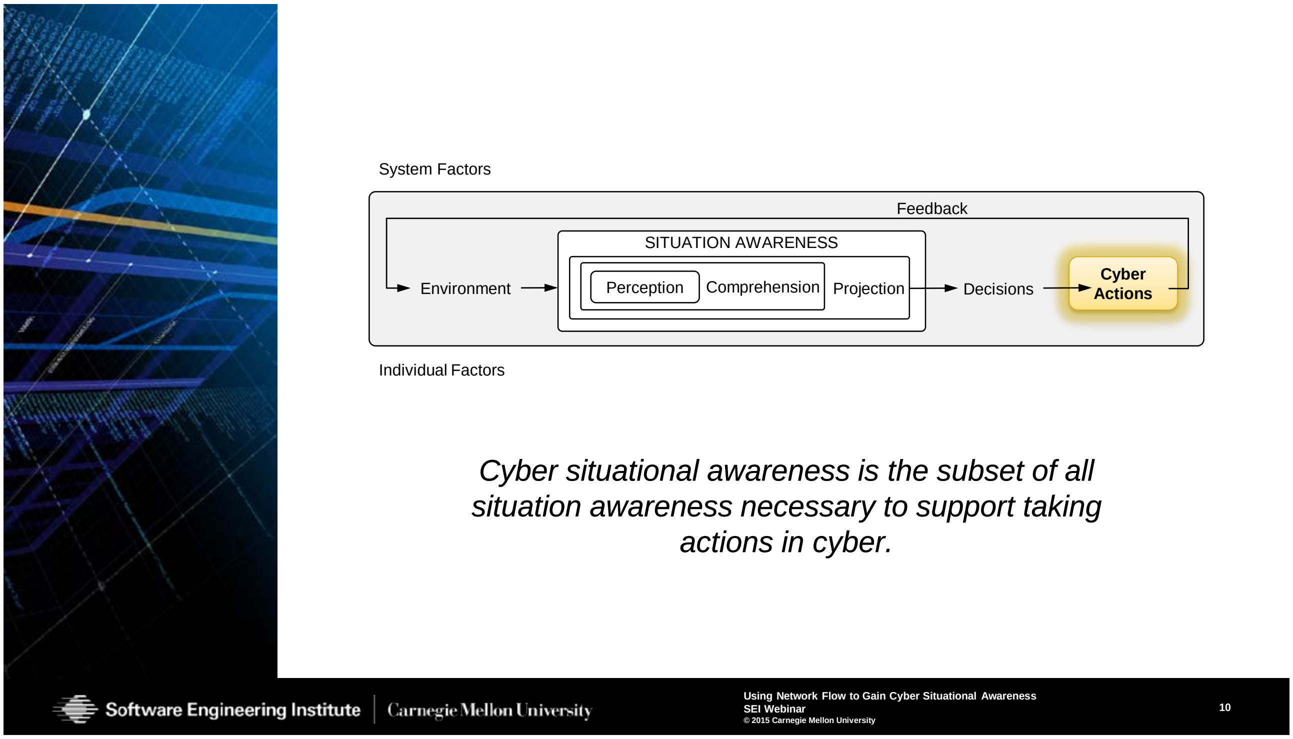
Sid Faber: How did you decide whether or not to wear a jacket into work?

Shane McGraw: Weather.com.

Sid Faber: Okay. So, very straightforward, go to weather.com, take a look at what the weather is outside today. And based on that, decide what action you want to take. We can fold that right into our situational awareness model where Shane knows that his decision is whether or not I'm going to decide to wear a jacket. On the left hand side, the data input is that weather.com data. Based on weather.com, I understand what's going on in the environment, and I understand whether I'm going to be uncomfortable later on during the day and decide whether or not to wear a jacket. And that resulted in finally an action that you would wear a jacket to work because it's a chilly day today. That's how situational awareness works in the general case.

How does that work in the cyber case, though? Cyber actually fits really nicely into this model.

Factors



System Factors

Feedback

SITUATION AWARENESS

Environment → Perception | Comprehension | Projection → Decisions → Cyber Actions

Individual Factors

Cyber situational awareness is the subset of all situation awareness necessary to support taking actions in cyber.

Software Engineering Institute | Carnegie Mellon University

Using Network Flow to Gain Cyber Situational Awareness
SEI Webinar
© 2015 Carnegie Mellon University

10

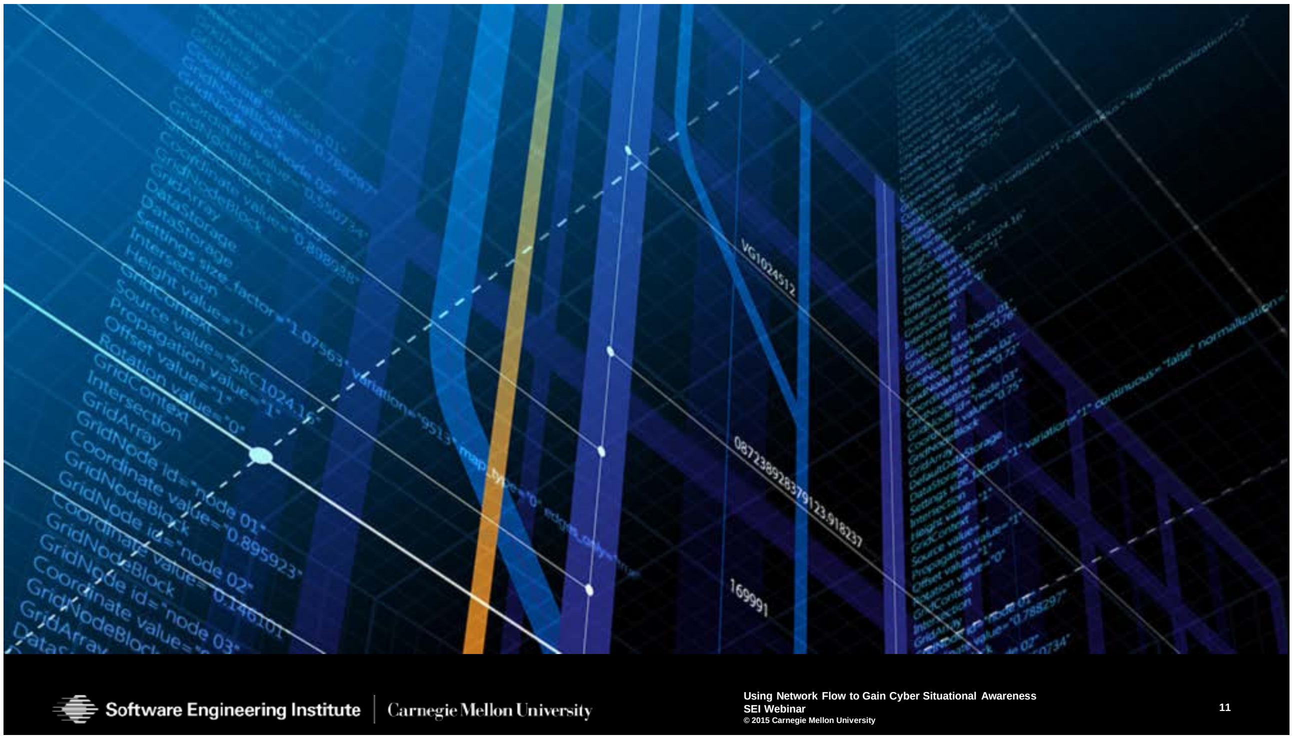
**010 All we have to go is constrain ourselves on the right hand side. Those actions that we want to take are actions in cyberspace. So, when we look at cyber situational awareness, it's that subset of all situational awareness that just simply looks to take actions to achieve our goals in cyber.

So, we simply constrain our problem into this cyber realm. And then we've defined what it is that we're most interested in. once we constrain ourselves and understand that this is where we're at, now we can look at some of the work, some of the rich body of work that's in the situational awareness realm, and apply it to our case.

So, for instance, Dr. Endsley in her work talks about things like mental models and creating models of the environment so that we can quickly move from data about the environment into actually understanding and making decisions based on what I understand about the environment. That applies directly to cyber. We can model the cyber environment and quickly move from data about cyber into that model, that understanding. And then based on that understanding, move into actions.

There is also constraints. Dr. Endsley proposes a lot of constraints, particularly on working memory and attention span. And we can see how those apply. We can get lots and lots of data about cyber. Sometimes, we get overwhelmed with data. And that directly applies to in the non-cyber case where I have a limited amount of working memory, a limited amount of input that I can receive about the environment. So, I constrain myself on what inputs I take so that I can better move from that data to the decision making. So, again, this model helps us define where we're at and helps us move toward a better understanding of cyber situational awareness.

Situational Awareness



Software Engineering Institute | Carnegie Mellon University

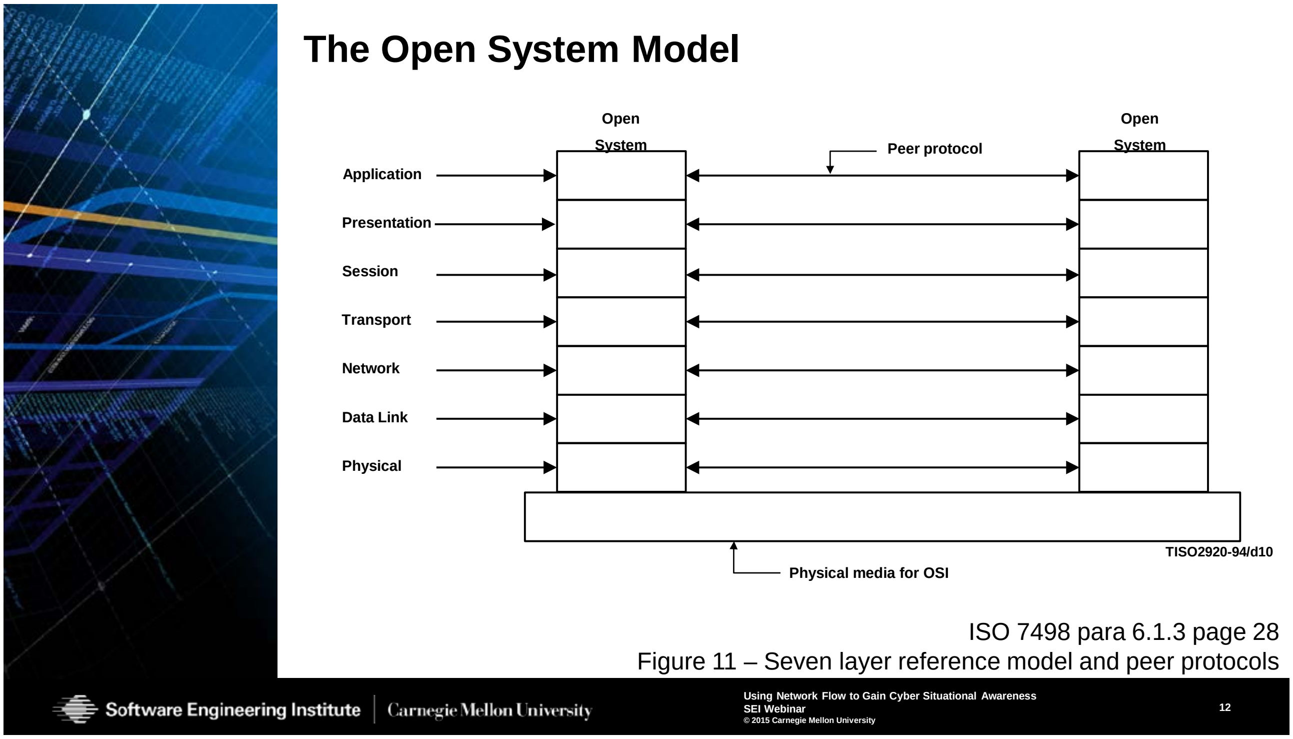
Using Network Flow to Gain Cyber Situational Awareness
SEI Webinar
© 2015 Carnegie Mellon University

11

**011 So, that's where we're at. So, the big question now is what about flow. How does flow and network flow fit into this realm of cyber situational awareness? I haven't talked much about flow yet. So, I'll ask you to kind of bear with me or believe here when I say that the cyber environment was formed through the network. Without the network, we don't have much of a cyber environment. We have individual computers, individual machines that work in isolation. But if we really want to look at cyber, we need to look at the network that creates that cyber.

So, the network becomes this critical foundation. So, if you believe that this network is foundational to cyber, then we want to gain some sort of understanding of that network. And one of the best ways to do that is using flow.

The Open System Model



The Open System Model

ISO 7498 para 6.1.3 page 28
Figure 11 – Seven layer reference model and peer protocols

Software Engineering Institute | Carnegie Mellon University

Using Network Flow to Gain Cyber Situational Awareness
SEI Webinar
© 2015 Carnegie Mellon University

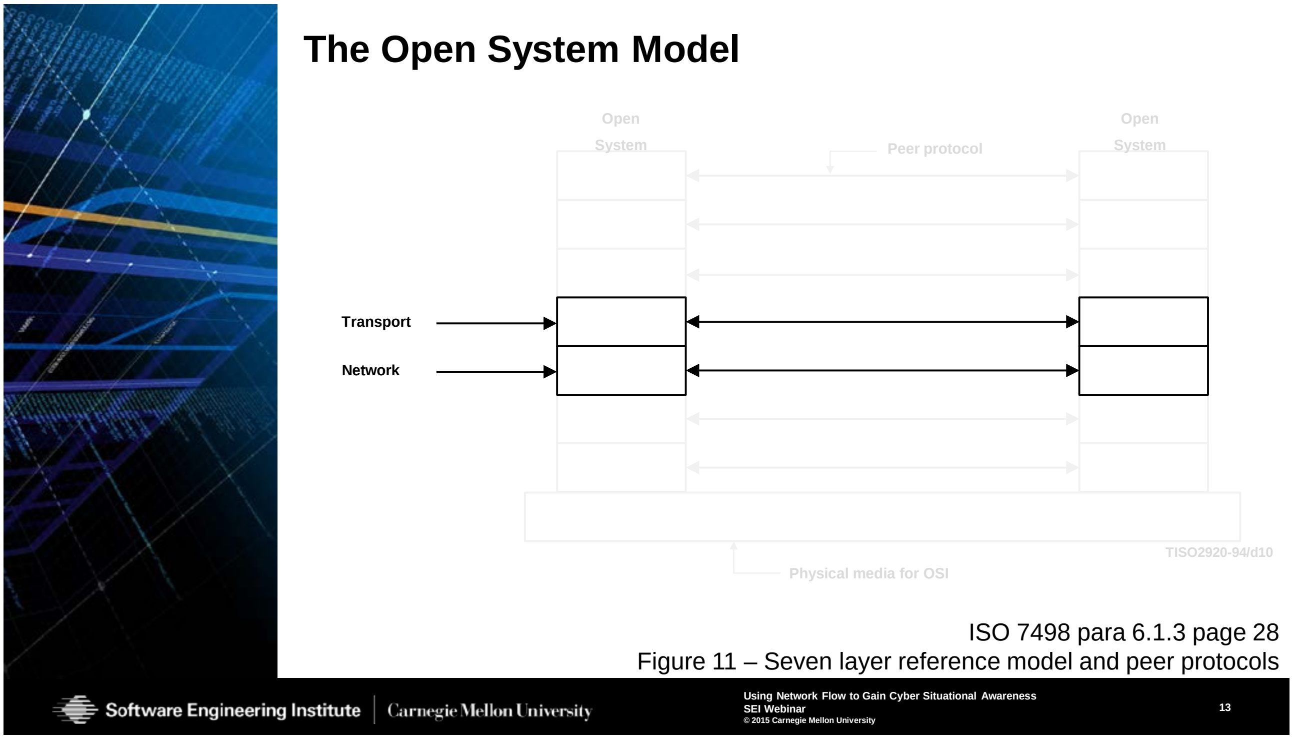
12

**012 Let's take a closer look at how that works. I think everybody's probably familiar with the seven-layer model, the OSI model. This is a model that you find in our textbooks, you find it in our citations throughout. I'd encourage you to take another look at the standard. Often, we only look at the textbooks and don't go back to the original standard. I believe this was put out probably in the mid to late '80s.

It's ISO 7498. It's a really, really interesting standard that talks about how to create communications. Although, we often criticize the seven-layer model because communications don't clearly fit into that model anymore, the fundamentals, the open system interchange that that model is based on is very foundational and really fits where we're at quite well.

So, let's start with this foundation about the seven-layer model--

The Open System Model



The Open System Model

ISO 7498 para 6.1.3 page 28
Figure 11 – Seven layer reference model and peer protocols

Software Engineering Institute | Carnegie Mellon University

Using Network Flow to Gain Cyber Situational Awareness
SEI Webinar
© 2015 Carnegie Mellon University

13

**013 And then consider that flow with flow, we're actually only looking at the network and the transport layer. This is a standard. This is the way communications work. So, if we agree that cyber comes through networking, and networking comes through this open system model, then this transport, this network layer exists in cyber. And we can leverage that. That's where flow fits. So, flow becomes this really nice kind of foundational layer that exist wherever the network exists, where the cyber domain exists. And it's necessary for that bit of communications.

Traditionally, we talk about a TCP flow, a UDP flow. We talk about IP addresses communicating back and forth. And we'll look at a few examples of that. But I want you to also think more generally that this flow is how networks communicate.

Shane McGraw: Sid, I want to chime in real quick with our polling results while you take a drink there. So, the first question we asked our audience-- and again, we ask these questions folks so you can get an idea, or Sid can have an idea who's in the audience, kind of tailor some of this examples and just to know who he's talking to to best fit your experience here. So, the first question we asked was how familiar are you with the concept of cyber situational awareness. The leading number was sixty-three percent was it's important part of my work, but I would like to better understand it. Then seventeen percent was familiar with situational awareness but have never applied to cyber. Fifteen percent was sounds interesting but haven't put much thought into it. Four percent studied and applied the concept extensively.

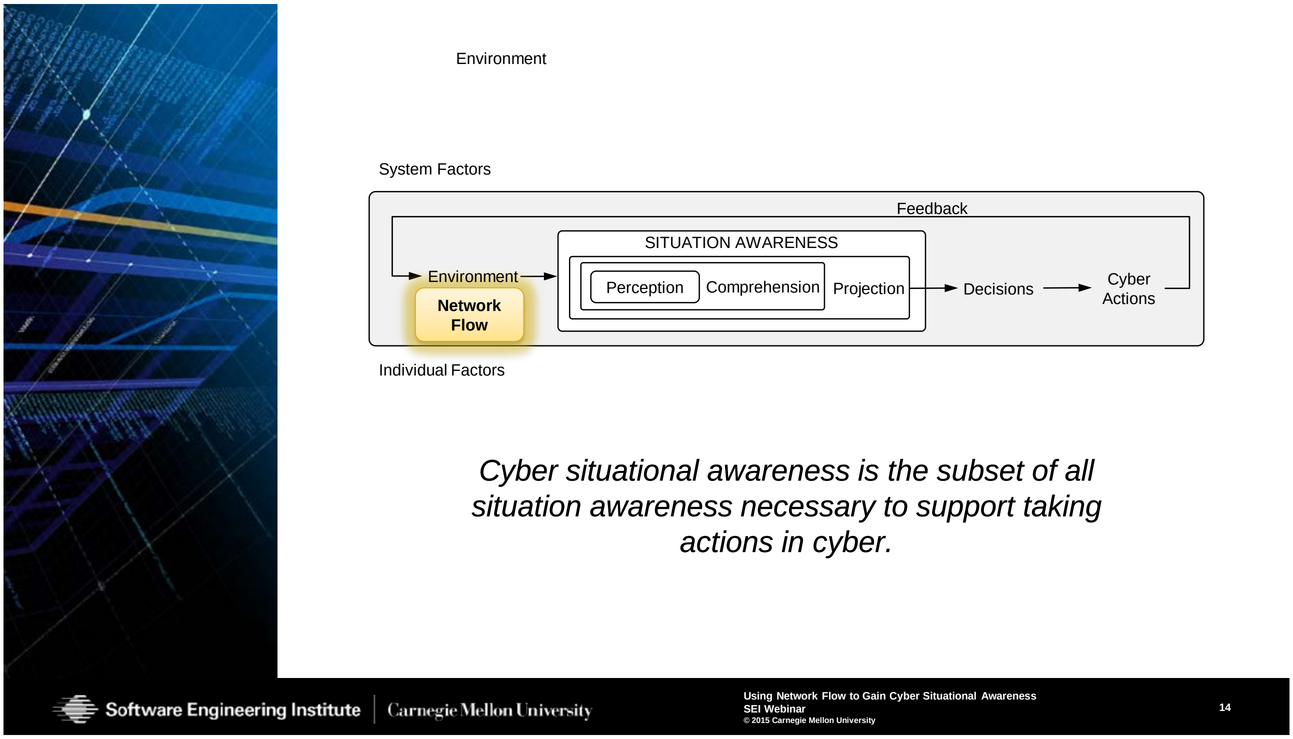
And then the second polling question we followed up with exactly who's in the audience by role or title. We had eight percent executive management, thirty-seven percent technical management, and forty-seven percent technical or operational staff. So, hopefully, that will help you tailor some examples and feedback during the talk.

Sid Faber: Sure. That's great. And again, as we go along, I encourage anybody to feel free to drop in some comments to make sure I tailor this to what your needs are. Great mix of folks out there today.

So, again, looking at that open system model, looking at where flow fits, we're talking about how these two machines talk back and forth, how networks as a whole talk back and forth. And we can sense a lot of that because the foundation on that and all this networking is this flow model. It's foundation to that cyber environment.

So, then when we go back to our model for the situational awareness--

Environment



Environment

System Factors

Feedback

SITUATION AWARENESS

Environment → Perception | Comprehension | Projection → Decisions → Cyber Actions

Network Flow

Individual Factors

Cyber situational awareness is the subset of all situation awareness necessary to support taking actions in cyber.

Software Engineering Institute | Carnegie Mellon University

Using Network Flow to Gain Cyber Situational Awareness
SEI Webinar
© 2015 Carnegie Mellon University

14

**014 Cyber network-- cyber takes its inputs from network flow. So, we can use this as one of those key ways to gather data about the environment to put sensor data into our situational awareness model to make sense out of what is actually going on in the cyber network.

So, flow gives us the network, gives us understanding about what that cyber domain is, what's going on there. And we have to do some work with the data to understand how we can make decisions based on it.

But yet, there's a challenge there. And this is, I believe, one of those fundamental challenges that makes cyber such a hard and difficult discipline to work with in situational awareness. Shane, let me ask you another question. Did you turn on your firewall today?

Shane McGraw: No, I think it's just automatically tuned on.

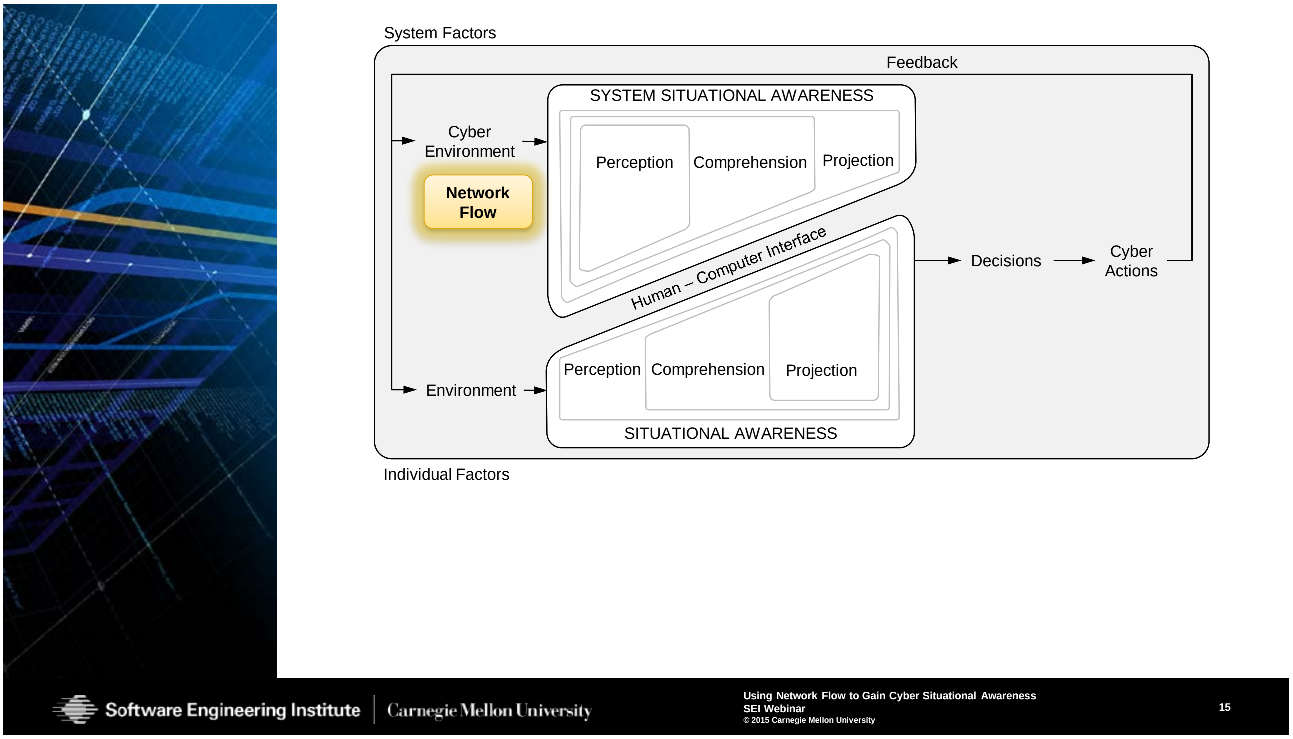
Sid Faber: Yeah. This is an odd question. It's uncomfortable. And I'd encourage you to think, "What did you mean by that?" Fundamentally, we can clearly understand "should I have worn a jacket today." Well, yeah I know what it means to be cold. I know what it means to put a jacket on. What does it mean to put a firewall on? What does it mean to turn my network on? How do I sense cyber?

When we look at other domains and other areas, a lot of times, there's that physical connection. There's the sense, sight, taste, smell, sound, that help us better understand. We're born into it. We learn all these concepts. And we build these things that we read about in the situational awareness literature. We build mental models. We build an understanding that helps us to make a decision.

In cyber, we don't have that native understanding of what actually is the cyber domain. So, when I ask you a question like did you turn on your firewall, it's not like you'll get cold if you don't have your firewall on. It is something completely different. And that is a huge challenge in cyber that we really need to address. What does it mean to actually be aware of cyber?

Because that's such a challenge, I actually would like to go back and revisit--

Factors



System Factors

Feedback

SYSTEM SITUATIONAL AWARENESS

Cyber Environment

Network Flow

Perception

Comprehension

Projection

Human – Computer Interface

Decisions

Cyber Actions

Environment

Perception

Comprehension

Projection

SITUATIONAL AWARENESS

Individual Factors

Software Engineering Institute | Carnegie Mellon University

Using Network Flow to Gain Cyber Situational Awareness
SEI Webinar
© 2015 Carnegie Mellon University

15

**015 That model that we proposed for situational awareness and split it in half to explicitly recognize that there is a big challenge here. When we sense data with network, when we sense our data in cyber, we're doing in a computer system with the computers that make up the cyber domain. They present us with cyber data. We can leave that data in that computer domain and let the computer do work with it. But at some point, I have to present that to a human. And a human has to understand what is going on there so that they can make decisions. That interface between the system domain and the computer domain is critical for situational awareness. And it's not really flat. It depends on how much data you want to present, how well the systems have baked data and so on. And so, that's why I put that as a kind of slanted line. And again, we'll work through some of this. But recognize that it is critical to move from data that's on the system side of the equation into data that's on the human side because, at the end of the day, the human has to make a decision based on the data.

Shane McGraw: Can we get two audience questions in quickly here, Sid?

Sid Faber: Absolutely.

Shane McGraw: Both from Tanesha asking, "From a strategical perspective, how does one go about to ensure cyber situational awareness from the top level down echelons," is question one. And question two is, "How should one apply cyber situational awareness into network redundancy?"

Sid Faber: Great questions, thank you. So, first of all, how can I ensure cyber situational awareness from the top down?

Shane McGraw: From the top level down.

Sid Faber: I think probably the most important key that's often overlooked is to start from the right hand side, start from that decision making process. Usually, at the top of the food chain, at the top of the organization structure, you have somebody that needs to make decisions. So, from those decisions, then we drive backwards. What are the things that I need to make those decisions? We can talk about that a little bit more later on, but certainly, in cyber you might have decisions like do I need to dynamically add compute capability. So, we see a lot of this in cloud environments. Should I add some cloud capability now based on what I have going on? That might be a very executive decision making process, but that should be fed by situational awareness. I understand what's going on in my cyber environment. So, I need to build situational awareness.

That maybe very implicit. It might not be so explicit. But still, you understand that perceive, comprehend, project allows us to support that decision making process.

So, and the second part of that question.

Shane McGraw: How should one apply cyber situational awareness into network redundancy?

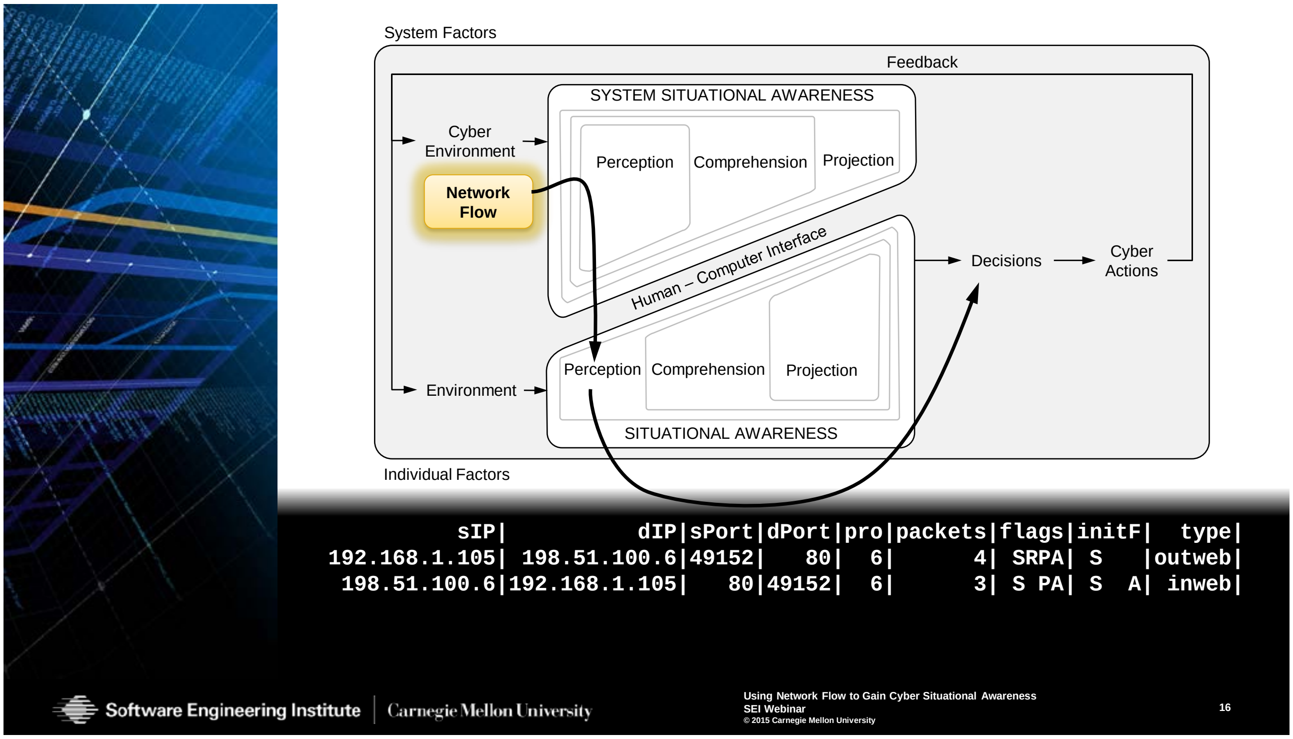
Sid Faber: How should one apply cyber situational awareness into network redundancy? I think the best way to do that is to look, again, starting at your decision process and then work backwards to understand how your network is put together in that when it fails, how am I ready to deal with that failure. So, I may have a decision that is should I add a redundant link or should I add redundant capacity, and I work my way back through that decision making process to understand how I would perceive that, how I would be able to act on that. and if you find that your actions can't be quick enough, maybe you need to change your continuity plan.

So, an example might be should I decide whether I should be hot/hot or hot/cold. Should I have redundant capability always available, or should I be able to spin up redundant capability when necessary? And again, I work backwards from what are my requirements. When am I able to make my decisions? And what am I able to sense on the network? I put that together and hopefully can arrive at an answer.

Shane McGraw: Terrific.

Sid Faber: Thank you. So, we'll take a look--

Raw Data



```
          sIP|          dIP|sPort|dPort|pro|packets|flags|initF|  type|
192.168.1.105| 198.51.100.6|49152|   80|  6|      4| SRPA| S   |outweb|
 198.51.100.6|192.168.1.105|   80|49152|  6|      3| S PA| S  A| inweb|
```

Software Engineering Institute | Carnegie Mellon University

Using Network Flow to Gain Cyber Situational Awareness
SEI Webinar
© 2015 Carnegie Mellon University

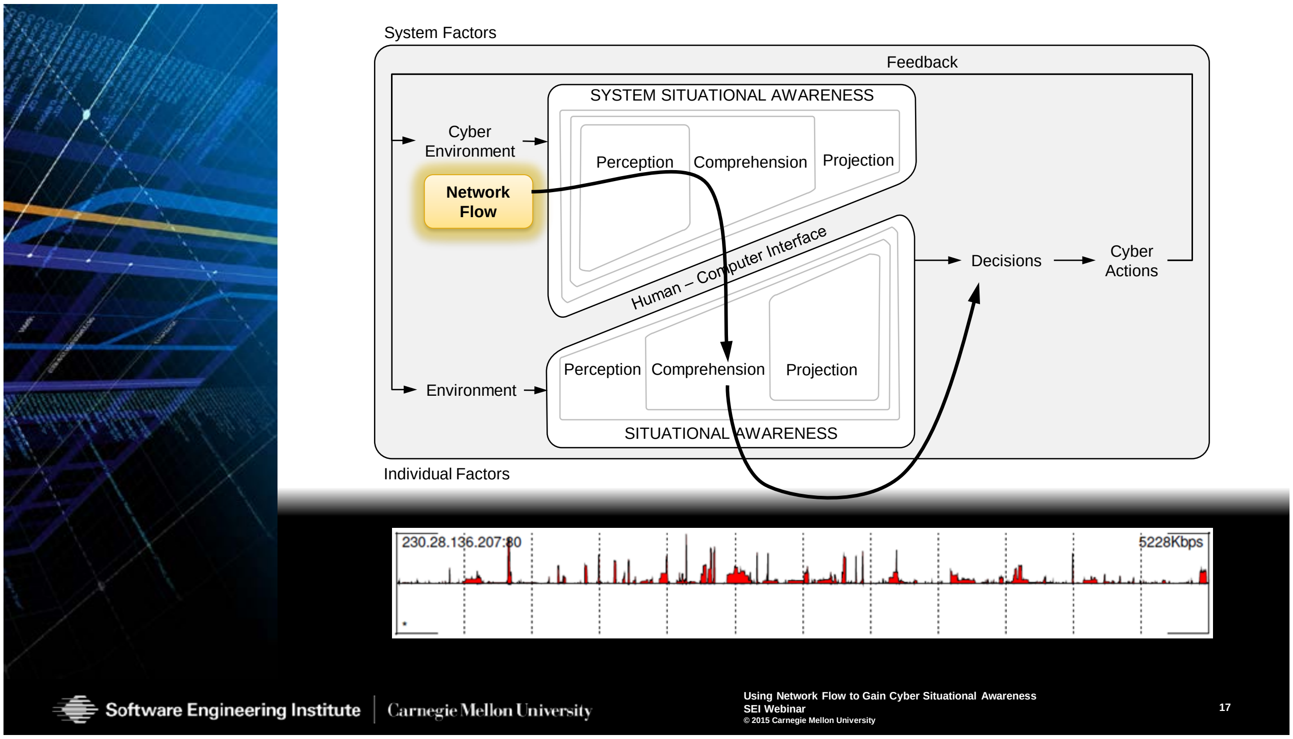
16

**016 Again, I want to dig into this just a little bit looking at our model here to see how this could actually work with very raw network flow data. So, again, this network flow data is telling me that I have two hosts that are communicating on my network. I have an example of a very low level piece of data here. Here you see two IP addresses, 1192.168.1.105. And then the other one is 192.151.100.6. There's a destination port 80. That makes you think this is probably web traffic. There's protocol 6 which says it's probably TCP, couple of packets, some flags. That sounds kind of boring, doesn't it?

There's not a lot of information there. And that's because the interface that you're being presented to the data is very low level, just very much at the beginnings of perception. I'm giving you the raw data. And then you have to make sense out of it. You have to boil this, bake it up, and then turn it into some sort of decisions.

And yet, I also haven't given you any decisions to make based on this data. So, there's not a lot there. A lot of times, this is where we begin with the raw data but recognize the system presenting us with such raw data really doesn't do much for us.

Flow Records



System Factors
Feedback
SYSTEM SITUATIONAL AWARENESS
Cyber Environment
Perception
Comprehension
Projection
Network Flow
Human – Computer Interface
Decisions
Cyber Actions
Environment
Perception
Comprehension
Projection
SITUATIONAL AWARENESS
Individual Factors

Software Engineering Institute | Carnegie Mellon University

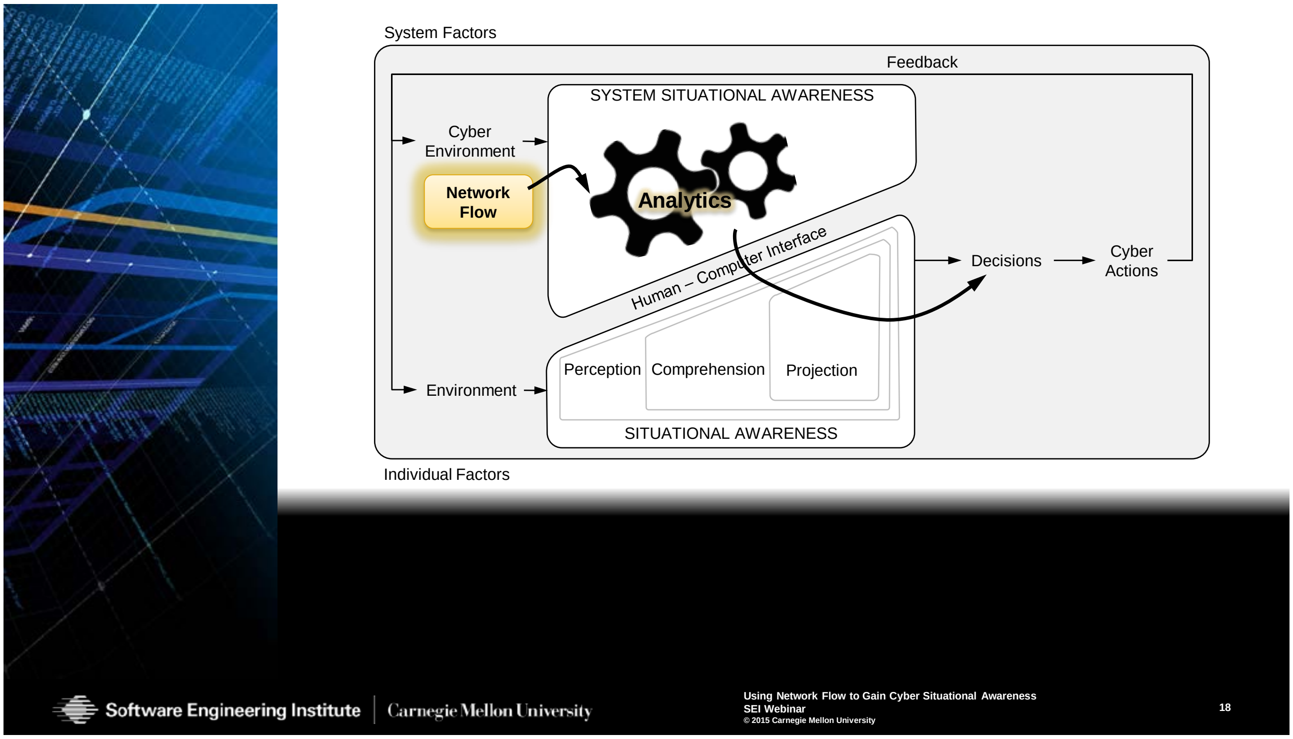
Using Network Flow to Gain Cyber Situational Awareness
SEI Webinar
© 2015 Carnegie Mellon University

17

**017 So, let's take a look at a little bit more complex example, if you will, again based on flow data. In this case, it's the same fundamental underlying sensor inputs that's provided into the situational awareness model. But here you see just a straightforward time series plot. In this case, we're just looking at byte volumes for traffic over a time period. But you can see how this is a little bit denser information. So, that human-computer interface gives me a little bit more toward the comprehension version of situational awareness, so I'm moving from the level one perception more into level two comprehension. Arguably, there's not a lot of comprehension going on here. There's not trends or anything on this plot. But still, it's a lot better than what we started with with the raw data.

I think this kind of helps demonstrate one of the fundamental concepts where we understand that split between the machine and the human where we can expect the machine to continue to give us more and more rich information. In that plot that we just saw, maybe if we had some trend lines, maybe if we had last week's data compared with this week's data, we might be able to better understand where we're at. And then you can go one step further and see how the system in its maturity could actually say this is what I expect to happen and potentially even alert on the case when it doesn't meet with what I expect to happen.

Flow



Using Network Flow to Gain Cyber Situational Awareness
SEI Webinar
© 2015 Carnegie Mellon University

18

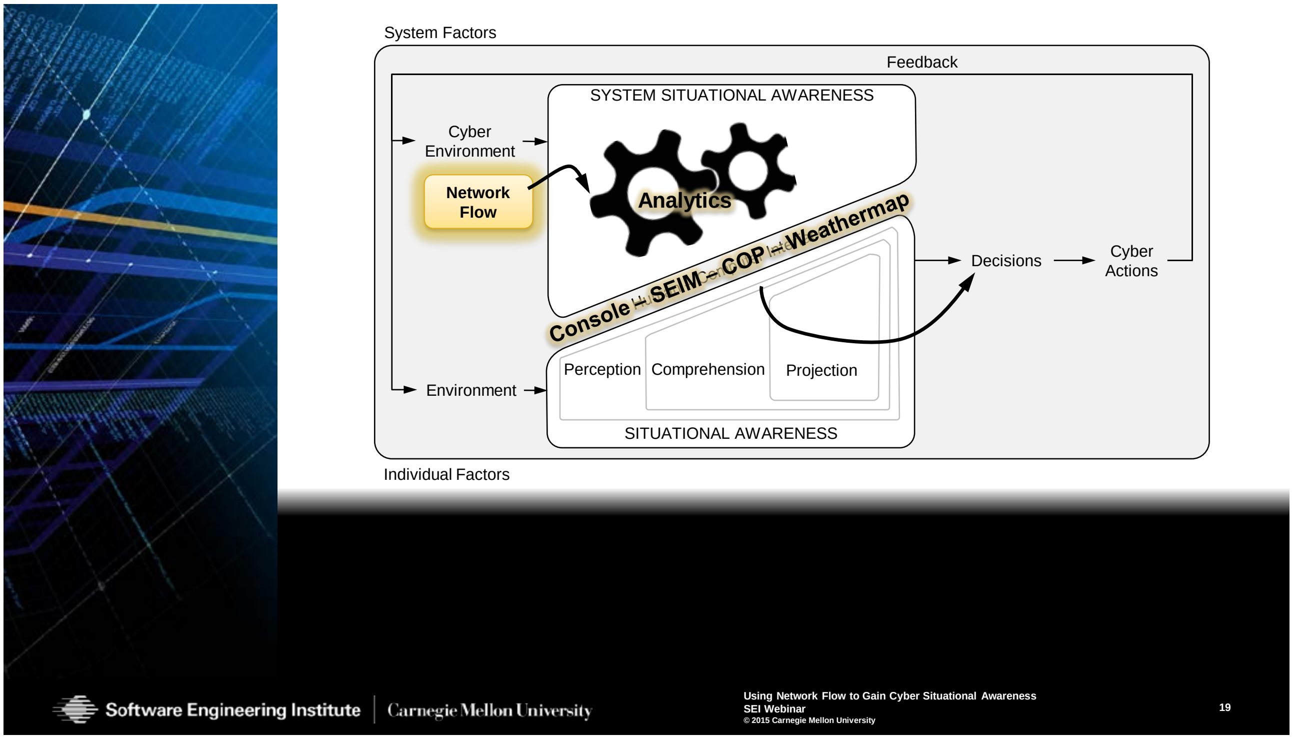
**018 There's a lot that you can do to feed into this raw model with this raw data and then bake the data to turn it into some useful knowledge. In a more general way, then, I'd like to propose that in that upper side, in the upper half, the computer version of the model, that's where those analytics lie. There's a lot of attention that has been paid lately to system analytics and flow analytics and network analytics. That's where they lie. We leave that cyber data in the cyber environment and let the machine work the data and then present to the human a more richer version, something that's more actionable, something that's closer to my decision making process.

Successful analytics are going to get

us much quicker into that decision making process. And that's my goal. I want to make decisions and take actions to achieve what my actual goals are.

The raw information is very cumbersome. We need those analytics. And that, I think, is where we find a lot of work. But there's another important part that we mentioned is that interface.

Console



System Factors

Feedback

SYSTEM SITUATIONAL AWARENESS

Cyber Environment

Network Flow

Analytics

Console – SEIM – COP – Weathermap

Decisions

Cyber Actions

Environment

Perception

Comprehension

Projection

SITUATIONAL AWARENESS

Individual Factors

Software Engineering Institute | Carnegie Mellon University

Using Network Flow to Gain Cyber Situational Awareness
SEI Webinar
© 2015 Carnegie Mellon University

19

**019 So, we see a number of different interfaces in between the computer and the human. And they've come across with a number of different names. So, I'm going to present a couple of those to you. See if some of those strike a chord with you, if you recognize the importance of that interface. Sometimes, it's just called the console.

When I think console, I think of often something like an SSH session, a raw session where you're just typing and you get very raw data. I think of that often as the left hand side of the equation, the left hand side of situational awareness presenting very low level data to an individual analyst. A lot of times, again, important to note that it's for an individual to serve an individual need to understand what's at the data. There's a lot of further comprehension and projection that happens in the individual side.

Another very common visualization tool is the SEIM, or security event and information manager. I think that's probably a lot more dense information. And how does that relate to flow? In some cases, we see organizations that will push flow data into the SEIM. That tends to be very cumbersome because flow data is large. So more often, organizations will do some analytics on the flow data, some roll up, some volume, some inventory or something along those lines, and then present that rolled up data into the SEIM.

The SEIM then gets it in front of an individual or potentially a team of individuals. And then they can work the data from there. They may understand what the disposition is of their network and be able to project how to take actions based on that.

In the military context, a lot of times we see questions about the COP, the common operational picture. Shane, I don't know if you're familiar with the classic, some of the classic movies where you see the battlefield. And there's things laid out on the battlefield. Maybe they're ships laid out on an ocean or soldiers laid out. And what do you think of when you see that? why would you see that, that a commander has that?

Shane McGraw: Strategy, what the next step is to go on the attack or retreat.

Sid Faber: Absolutely. Absolutely, strategy, what's my next step. It supports that decision making process.

We're looking for one those in cyber. Unfortunately, all the ones that we have from the military context sit in a geographic terrain. We don't have that in cyber. As I mentioned before, the parallels are very challenging there. So, it's very challenging to get that common operational picture. I want something like that battlefield that shows the disposition of my forces, the disposition of the threat, and I can understand and take action and figure out how to maneuver. It's just very challenging to present that in a cohesive way to cyber. Fundamentally, I believe it's because we just don't sense cyber the way that we sense the other physical environments.

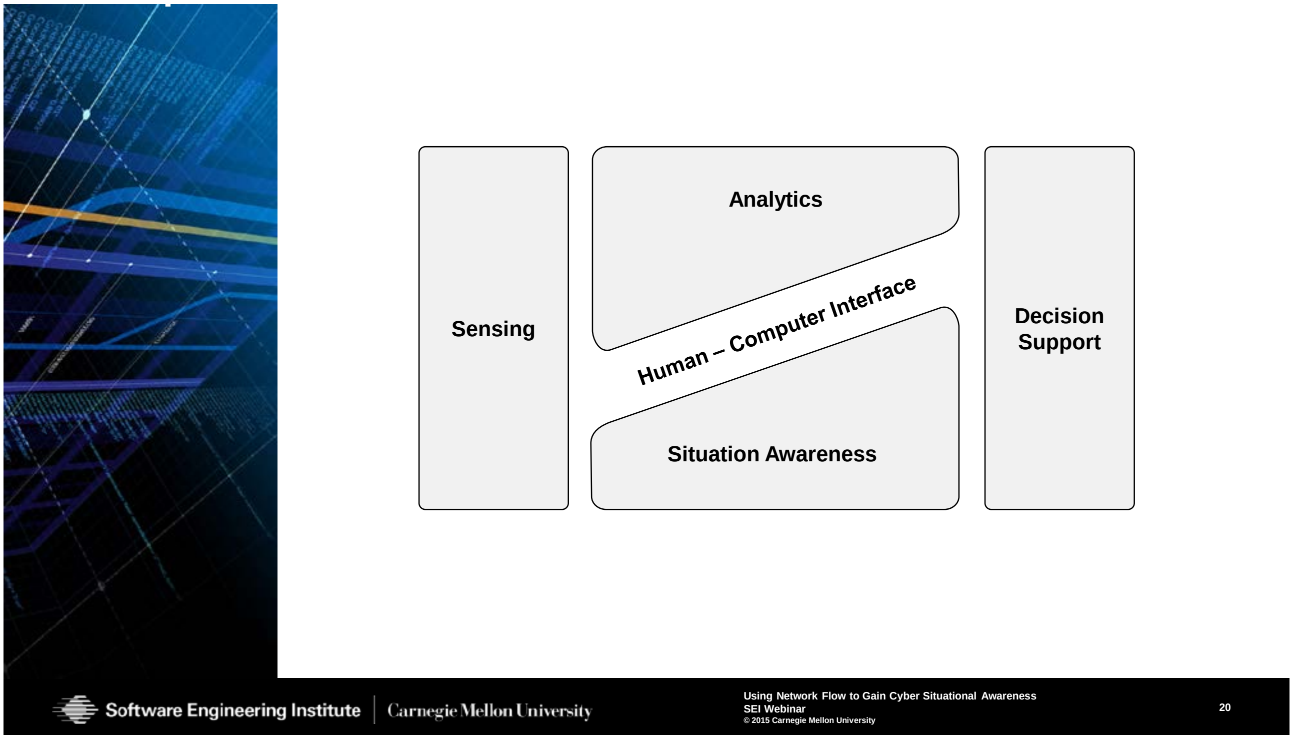
Still, there's a lot of work that can be done there. Sometimes, just presenting a list of alerts or list of open tickets or a list of systems that are up and down goes a long ways toward presenting that situational awareness picture, those things that are most important to help a leader make decisions or take actions within that cyber domain.

Another good analogy is the weather map. So, we started by talking about should I wear a jacket today. That's perhaps kind of a trivial decision. But think about other decisions that can be supported by a weather map. We're looking at a holiday weekend here. If you're watching this later, this is actually before a holiday weekend in the United States. So, there's a lot of people that might be travelling and want to pack. How do I decide what to pack? Fundamentally, that comes from the exact same data that Shane used to figure out whether to wear a jacket. I want to sense the environment.

There are slightly different analytics, different analysis that happens to that data because I'm going to look at what the weather is going to be like for another couple of days, longer time period, or maybe in a different location. But still, that analogy holds pretty well that I want to compute and work with the data and use that data to support decision making based on what's going to happen in the near future.

So, that leads us to what I'll propose to you--

HCI Space



Sensing

Analytics

Human – Computer Interface

Situation Awareness

Decision Support

Software Engineering Institute | Carnegie Mellon University

Using Network Flow to Gain Cyber Situational Awareness
SEI Webinar
© 2015 Carnegie Mellon University

20

**020 Is a more general kind of model that simply has five very basic areas that we can talk about for situational awareness. So, we look at sensing, how we get data into this situational awareness model. We look at analytics, so how we compute or how we work on the data that we have available to us. That critical human-computer interface of how we present the data that's derived in the computer systems how we present that to a user. Then the actual situational awareness, the human aspect of situational awareness. And then how that all supports our decision making process.

Shane McGraw: So, Sid, I'm going to jump in here with a preemptive audience question. That's a lot of work, obviously, to cover that and monitor that. are there tools that you can talk about that help with that. Or how would a company attack that to monitor that?

Sid Faber: Yeah, that's a great question. There's a lot of work here in this space potentially. And how would you attack that? and as we mentioned in an earlier question from you was how do we get after this, how do we help people achieve situational awareness?

There are a lot of tools out there. We'll start by focusing on that human-computer interface. You probably have one of these in your organization today. Open source tool might be Splunk that pulls in operational information about your cyber environment and presents it to you. You can create rules and roll up rules in Splunk that will tell you when things are very important to you, that have happened that are very important to you. Maybe a system goes down or a threshold is exceeded.

Another common one we see is ArcSight, commercial offering, where you can take in data from a number of different sources. And you can have a whole ecosystem of data sources from the cyber environment presented into this one interface, the ArcSight console. And you can create rules in that console to actually move you more towards the right hand side of situational awareness, move you closer and closer into that projecting what's going to happen based on what you see currently occurring in the environment. Again, that supports that decision making process. These consoles get tuned toward the decision making process.

There are a number of geographical displays that happen in that human-computer interface as well. A lot of times it is useful to take information that we find in the cyber environment and lay it over top of a geography. So, I can see where things are going on in the world. It becomes very interesting and useful, particularly when you want to bring in information that's happening in other domains. So, for instance, one of the common things might be I use that human computer interface to overlay physical equipment on a map. And then I look at the weather on the map. So, if I have a storm coming in, I can see a storm is going to potentially hit where certain physical network components exist. And then I can understand how to take action based on that. So, geography plays a role in that interface as well.

As far as the sensing side, there's a number of sensing tools. And particularly speaking about flow, as flow becomes this valuable way to get information from the cyber domain, there are a number of flow tools. Flow actually is native in a lot of your networking equipment. So, it may just be a configuration option that you can turn on on an existing router. And the router will generate flow information, information about those connections that it sees crossing it and be able to send that out to a repository, a database somewhere that can collect that. And again, your SEIM may be able to collect that data.

CERT actually produces a number of tools as well. And I encourage you to take a look at those. If you're interested at all in exploring flow analysis, we split those into a few different categories. On the pure sensing side, we have a tool named YAF, yet another flow meter, that will sense the network data. You send it packets, send it a stream of network data, and it will turn that data into a flow record and then send that flow record off to a repository. And it can gather more rich information, if you're able to accept it, about some metadata about the communications as well.

There's the SiLK tool set, which is the receiving end of the flow data that allows you to do analysis. And this is very low level analysis close to the left hand side is situational awareness off of that raw flow data so that you can bake and work with that raw flow data and create routines that will then generate you more and more rich information off of that. We also have Pipeline as a streaming flow data analysis engine, very lightweight analysis, but yet quickly moves that data into the human interface and allows you to alert very quickly on what's going on in the environment.

Our tools have their parallels, other parallels in the open source community. Argus is another very common flow tool. And there are other sampled flow engines that you can work with as well.

Analytics become a huge challenge there. That's another tool that's on the cyber side of the equation. That's one where I see that there's a lot of work going on in that space today. And there's a lot of maturing that goes on in that space. Just like we used the weather analogy and how the analytics change based on what the question is you want to answer, so too in the cyber analytics realm. Those analytics are very dependent on what the decision is that you want to make. So, if you map out your decision space, you should be able to find the analytics that will support that decision making process and also may drive what sensor data you want to achieve as well.

So, long-winded answer to a short question. But yes, there are tools out there. And certainly follow up with us afterwards. Feel free to drop us an email after the presentation as well if you're looking for a tool in a particular niche area.

Shane McGraw: So, that's going to lead us to our fourth and final polling question for today, which we will launch now. And that question is what would you like to hear more about. Is it sensing, analytics, situation awareness, the human-computer interface, and decision support? So, we'll leave that up for a couple seconds. And we'll head back to our question console and get an audience question for Sid while we're voting on that. From Robert asking, "Please comment on the different dimensions, and encapsulation attributed to the perception, comprehension, and projection." Does that make sense to you or do you need or detail to it?

Sid Faber: Yeah, so could you read that one more time for me please.

Shane McGraw: Please comment on the different dimensions and encapsulation attributed to the perception, comprehension, and projection.

Sid Faber: Okay. Great, thank you. That's a great question. So, start with the concept of encapsulation. Why does that model show that perception is encapsulated inside comprehension and comprehension is encapsulated inside projection? And again, Dr. Endsley states this much better than I will. But I'll summarize by saying you simply can't have comprehension without perception. And you can't have projection without comprehension. So, one depends on the other. So, you begin-- you must begin with perception of the environment. And you have to do something to sense the environment, to measure the environment, to see what's going on in that cyber domain before you can actually understand what's going on. So, I think that speaks to the encapsulation part.

For the features of what's actually in perception, comprehension, and projection, I think starting again, working from the left to the right. Perception is sensing the environment. There are a number of cyber sensors out there. I think perhaps a weather analogy fits really well. So, let's step over into that weather domain and bring it back to cyber. In the weather domain, I can sense a number of different things. I can sense the temperature. I can sense the pressure. I can sense the pressure at a hundred feet in altitude or even a thousand feet in altitude. I can sense the humidity and so many other things if you start to think about it. And then I have all that where I'm actually sensing it. Do I sense it here? Do I sense it two feet over closer to Shane? Do I sense it--where do I sense it, and where do I need to sense it?

That applies directly to our cyber environment. Sometimes, we get caught up in the idea that I need to do too much sensing. More data is going to get me better answers. Going back to the weather analogy, that's not necessarily the case. If I want to figure out whether or not to wear a jacket tomorrow morning, I do not need to know what the temperature is at a hundred feet in altitude, nor do I need to know the cloud level. All I need to know is what's the temperature outside my door. So, we want to make sure that in that perception that we gather the data. But we also use this entire situational awareness process to gather the right data, gather data that actually helps support the decision making process.

In the comprehension part, I think we touched on that briefly a bit earlier. But also recognizing here we're converging from both sides of the problem. On one hand, I need to have data in order to do comprehension, in order to do that analysis. On the other hand, I need a decision. What is the decision that I want to support? So, once I understand both of these parts, then I can bring them together to create or define what the analytics are that I either want to purchase, or I need to create, or I need to tune to support my environment.

Projection is a very interesting case. It's a challenge potentially for today's systems. We don't see a lot of projections. A lot of times the system doesn't project, but the human does. Let me give you an example, a very simplistic but concrete example, of a successful projection and a projection based on flow. It's not uncommon for us to see some sort of denial of service attack. Denial of service attack is relatively easy to find with flow data because again it's an attack using the network. So, I recognize--no, I've sensed the network enough to see that there's a denial of service attack going on. I understand. I've got the analytics that help me understand that there's a denial of service attack.

Now, I can actually take an

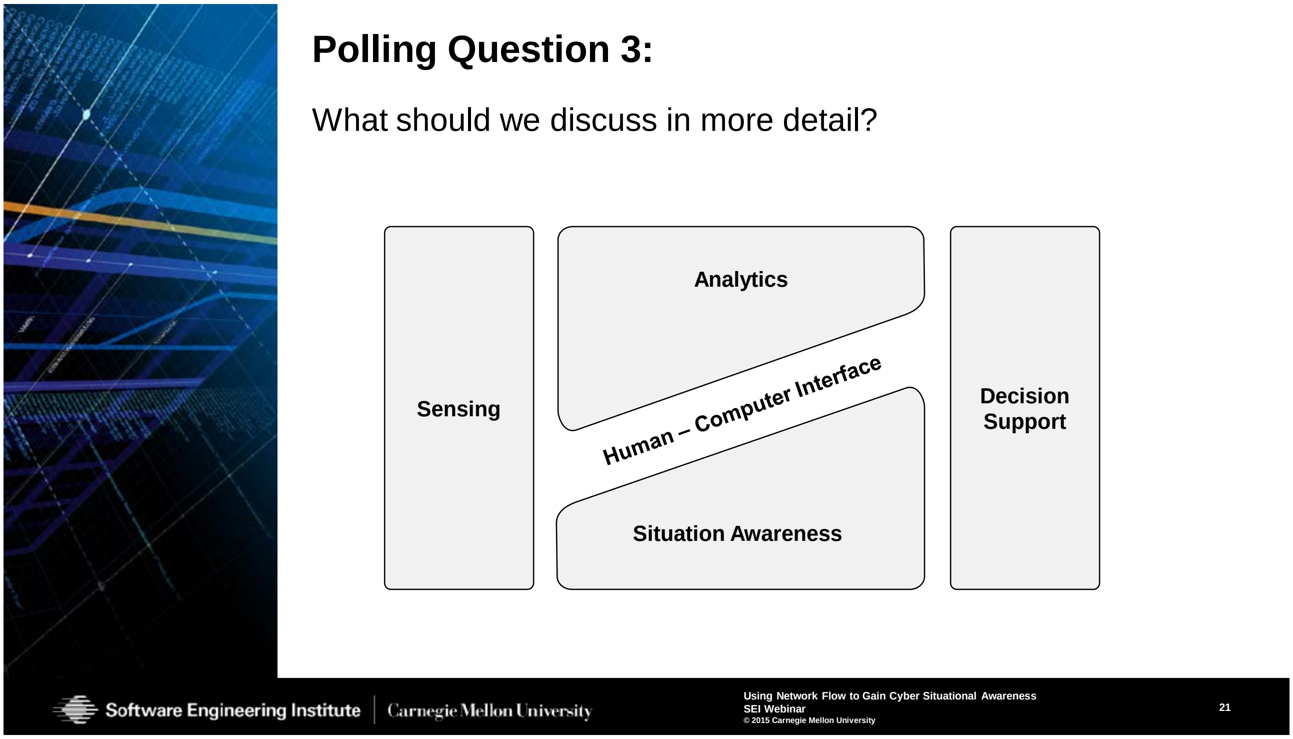
automated action. I may choose to block one or two of those sources automatically. And I can configure my system to block them. That's a case where the system actually is recognizing that not only am I under attack, but also that I need to take action. And I have a pre-approved action that I can take. And it takes that action.

One thing I think though that is critical in that whole, overall process is to remember that now that I've taken that action, I still have to get that into the human side of the equation. So, not only do you allow the system to take action, but you also have to let the system notify you. So, that human-computer interface still plays a part in that whole cycle.

Shane McGraw: Okay. So, we'll get to that poll results real quick, Sid. Forty-five percent would like to hear more about analytics, twelve percent sensing, twelve percent situational awareness, fourteen percent human-computer interface and decision support, so analytics by a pretty wide margin.

Sid Faber: Okay. All right, great, thank you. So, let's dig into that analytics part here. Let me flip over to my notes.

Polling Question 3:



Polling Question 3:

What should we discuss in more detail?

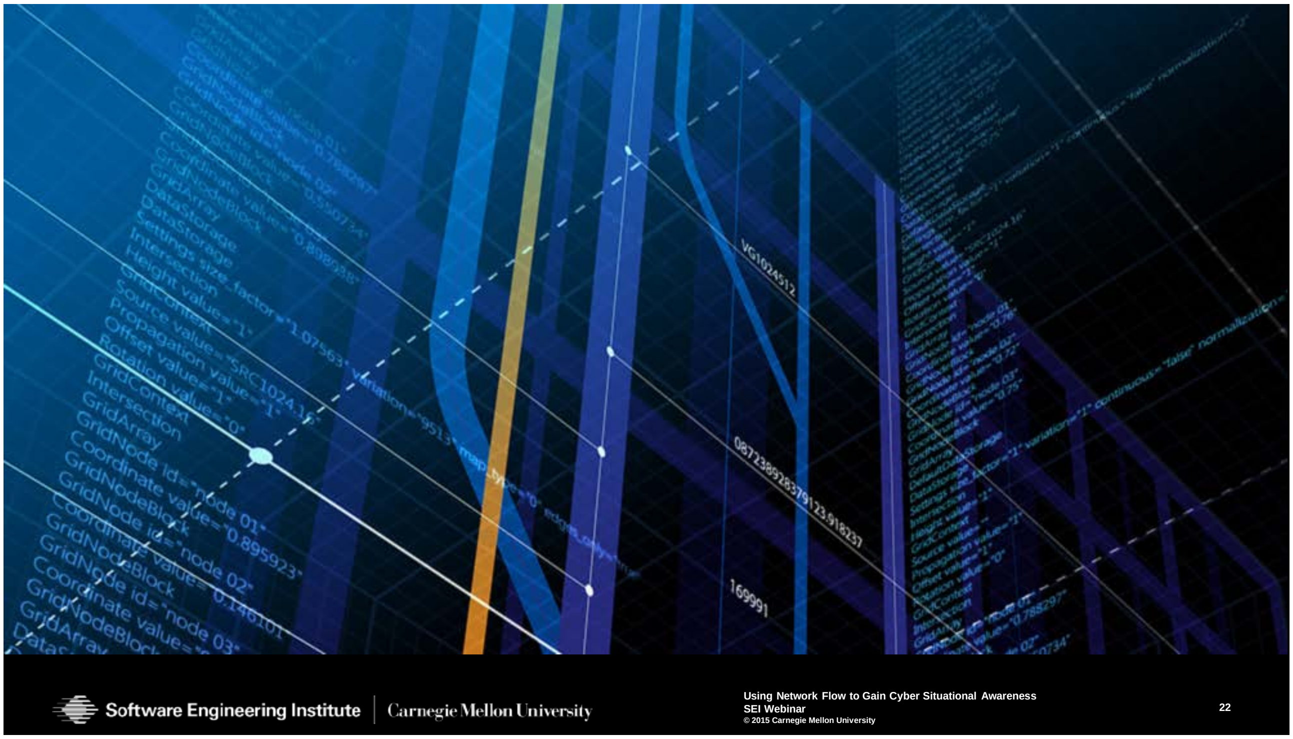
Analytics
Sensing
Human – Computer Interface
Decision Support
Situation Awareness
Software Engineering Institute | Carnegie Mellon University

Using Network Flow to Gain Cyber Situational Awareness
SEI Webinar
© 2015 Carnegie Mellon University

21

**021 This is definitely--

Situational Awareness

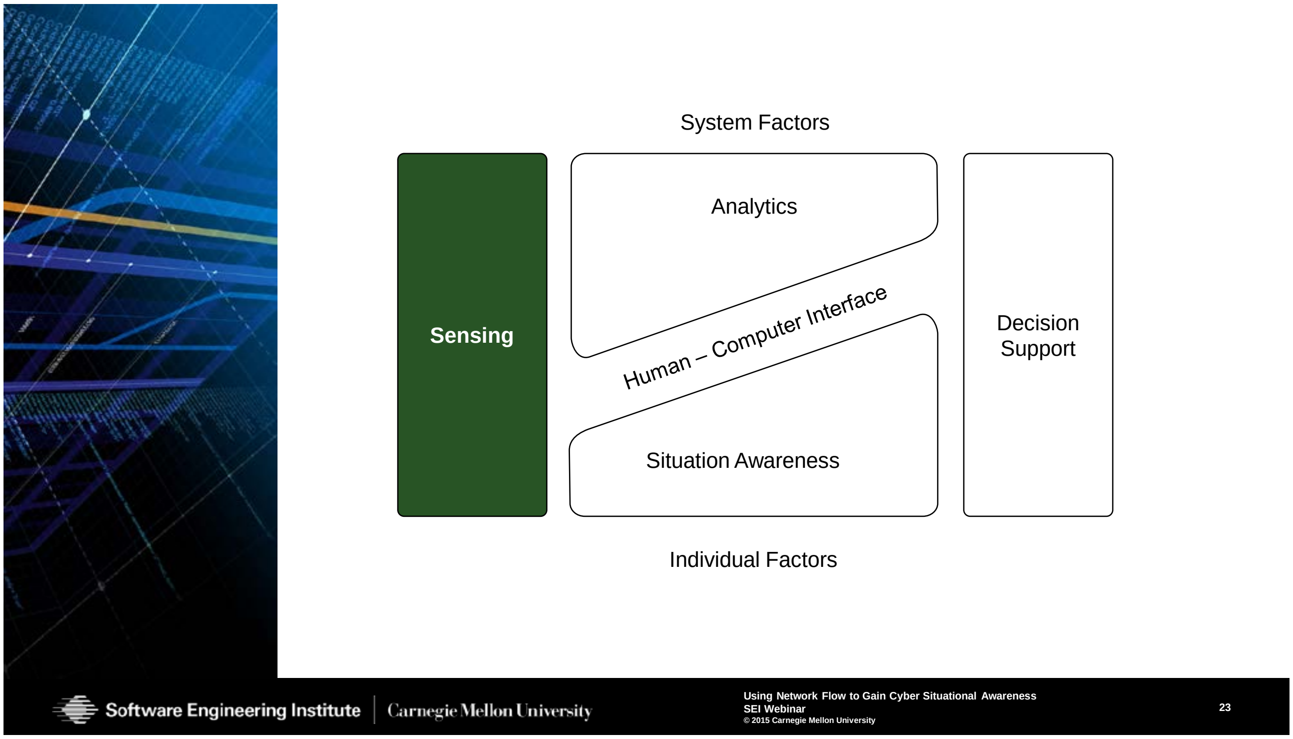


Using Network Flow to Gain Cyber Situational Awareness
SEI Webinar
© 2015 Carnegie Mellon University

22

**022 I think you're spot on.

Sensing



System Factors

Sensing

Analytics

Human – Computer Interface

Situation Awareness

Decision Support

Individual Factors

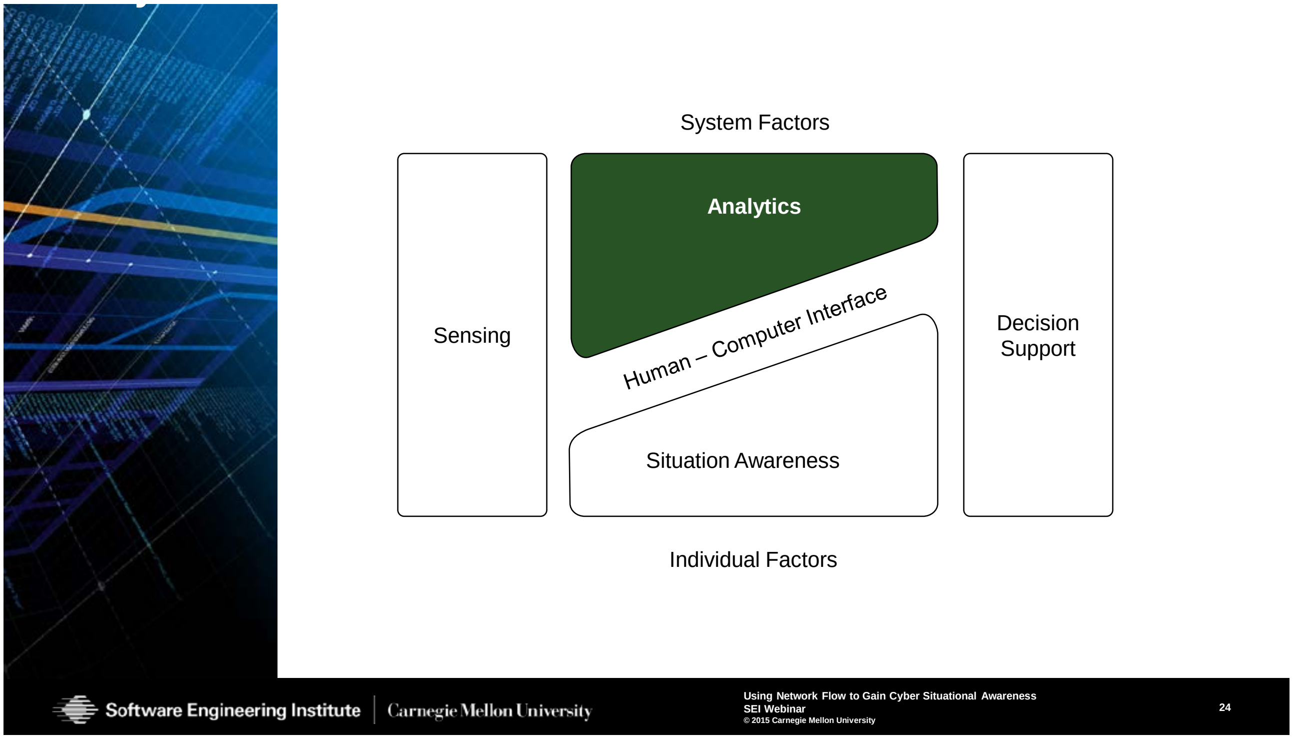
Software Engineering Institute | Carnegie Mellon University

Using Network Flow to Gain Cyber Situational Awareness
SEI Webinar
© 2015 Carnegie Mellon University

23

**023 This is a key component and where we see a lot of work going on in this space.

Analytics



System Factors

Analytics

Sensing

Human – Computer Interface

Decision Support

Situation Awareness

Individual Factors

Software Engineering Institute | Carnegie Mellon University

Using Network Flow to Gain Cyber Situational Awareness
SEI Webinar
© 2015 Carnegie Mellon University

24

**024 And there's a lot to be done here. So, let me start by saying I brought a few books with me today. And these are-- so, if you'll notice, for instance, the title of this first book that I have, this is a classic book. It's called "Flows and Networks". This book was actually written in the mid '60s. And it has nothing to do with computer networks. That's why it's kind of so neat is that there's a lot of rich history in network analysis, flow analysis, that doesn't relate to computer systems. I'd encourage you, if you're interested in analytics and have any sort of academic banner, you're interested in math, to look into this discipline.

This book in particular was written by two gentlemen, Ford and Fulkerson.

And they started this work in the late '50s. The book was published in 11962. And here's the problem that they were getting after. They were to determine the maximum rate at which a material could be moved through the Eastern European rail network to support a hypothetical Soviet invasion. These gentlemen were studying the traffic problem of railroad traffics. There's a lot of rich analytics in that discipline that we can directly apply to our network discipline. And I think if we understand, again, how these analytics works, it will go a long ways.

Let me put that in perhaps a slightly broader context. When we bring flow into the equation and we're sensing that network data, most often today what we do is we count. We count how many bytes we saw. We count how many flows we saw. We count how many packets we saw. And then we present that to the user. That's a pretty rudimentary measurement. It's very useful in some cases. But there's not a lot of analytics that happen on that.

Some of our work here at CERT that you should see publications coming out recently and in there near future have to do with getting richer analytics, getting richer metrics out of that data. So, for instance, persistence, I can measure persistence. How persistent are people visiting my website? That's very achievable using flow data.

Another interesting one is popularity. How popular is a public service that I've provided? And what does it mean when that popularity changes? As I mentioned, we're at the time that this is being recorded, being given, we're working our way into a holiday weekend in the United States, which tends to be a very heavy shopping season. Popularity is a huge issue with retail during this time period. So, in the network side of things, in the situational awareness side of things, it's very important to understand popularity. Again, this is a metric, a measurement that's been studied in other disciplines. We can apply it our discipline using flow data with some minor adjustments.

Another one is conformity. So, you've heard a lot of talk, I'm sure, about finding aberrations in flow, finding abnormal patterns. One abnormal pattern is when things don't conform. So, there's a rich body of work in conformity measures, and how I create what my typical behavior is, and how well today's behavior conforms to that traditional nature. So, that becomes a really interesting metric, a very valuable metric as well. So, one part of analytics is creating those new metrics.

Another part of that analytics platform that I'd like you to consider is that it's not a flat space. There are actually different flavors of analytics that we should consider. I like to break them up into three different planes. And you'll see how the systems that we use to analyze our flow data, our network data, depend largely on these different types of analytics, these different problems.

So, first of all, there's a data plane analytic. So, in the data plane analytic, I'm looking at that data that actually is transmitting across that flow, the data that goes between two endpoints on my network. These analytics are very lightweight because they're very fast. So, data comes and goes very fast. They don't have to be very rich analytics, or very deep, or very stateful. This is where a lot of our counts and our quick measurements, and our quick metrics come into play. Very high volume data, streaming analytics fit very well into understanding the data plane, understanding what that data is that's communicating back and forth.

A second but very different problem, I believe, is that control plane analysis. Control plane is the opposite extreme. In control plane, we're looking at how things have been configured. In particular, how do I establish that communication? How do figure out how to find my other endpoint? And we see a lot of, for instance, DNS analytics and routing analytics that fit this into this realm. These analytics are very model intensive. It requires you to create a very complex model, whereas the streaming analytics were very lightweight model, create a very complex model, which is very compute intensive. But also it doesn't change nearly as much as that data stream does. So, you would tune a system very different to do control plan analytics. And it would be a very different set of problems that you go after.

Again, we're trying to constrain our situational awareness problem based on that decision that we're trying to support. So, that all weighs into it. If I am trying to answer a data plane decision, then I want that very fast, very high volume type of platform to do analytics on. If I'm talking about a control plane problem, then I want this very complicated, very rich model type of platform that doesn't change very much. And I can probably even do batch analytics there.

Somewhere in between there, there's this endpoint analytics where I want to study the endpoints of my communication. Examples of that are things like reputation analysis is very common nowadays. Another one is system inventory. It kind of fits in between those two different realms, but it's a unique analytic on its own. And you'll see that particularly inventory analytics are very much in demand. And not terribly complicated, especially when you have some good, rich flow data.

There's also one of the attachments to our presentation. You'll see there's a network profiling using flow. A technical report was put out about two to three years ago. And it's a great way, I think, to start looking at flow data and how I can use flow data to get after that endpoint analytics problem, that inventory analytics problem. Hopefully, that starts you thinking down the process. If you have more questions along that, please feel free to type them and send them our way.

Shane McGraw: Great.

Sid Faber: Do we have any more questions?

Shane McGraw: We've got lots of questions in the queue. So, we have about eleven minutes yet. So, are we ready to dive into the final question thing? Or are there any other fields you want to cover before we start that will--

Sid Faber: So, let's go into the final questions, and that probably will tease out what we wanted to talk about here.

Shane McGraw: And folks, before I want to do that, I just wanted to remind everybody in that downloads section within your webinar console, you will see a postcard for the FloCon conference. I mentioned that at the beginning. That is coming up this January in Daytona Beach. If you like what you heard today, that conference will take it to another level there, really a deeper dive. And if you like what you heard, and Sid was talking about, it's a place you need to be. If it's not your area of expertise, please share it with network people in your organization or your communities. But the conference, I believe, is in its twelfth year. It has really grown into a community of users. So, it's a great conference. Actually, the early bird rate for the conference is going to expire November 30th. So, you can have a savings if you go to www.cert.org/flocon. And you'll see all the conference detail. The program is up now, a great line-up of speakers, great keynotes. So, take a look at that upon exiting the webinar.

So, let's get in some questions for Sid. From Robert asking, "Is there an established cybersecurity-as-a-service market which might provide the firewall and situational analysis tools for firms? Such a service firm could also benefit from the aggregate flow analysis of many firms." Let me know if you need a repeat.

Sid Faber: Yeah, if you could say that one more time please? Thank you.

Shane McGraw: Sure. Is there an established cybersecurity-as-a-service market which might provide the firewall and situational analysis tools for firms? Such a service firm could also benefit from the aggregate flow analysis of many firms.

Sid Faber: So, is there a cybersecurity-as-a-service? Obviously, yeah. So, managed security service providers. But do managed security service providers provide situational awareness? It may not be explicit, but it definitely is implicit. And also, the fact that a managed security provider being able to understand the broader picture using telemetry, using information, using inputs from a much broader audience definitely exists. And you see that.

Recognize, however, again going back to that weather analogy, is that understand those decisions that you want to support. And that drives a lot of what you need out of your security service provider. If, for instance, you are in a particular vertical-- again, I mentioned retail. So, if you're in the vertical of retail, and you're providing retail services, your managed security service provider may have a broader reach into the retail market and understand a bit more about threats in retail, understand more of what's going on in retail and be able to apply that specifically to you. So, that becomes an advantage to you. They're focusing not just on cyber as a whole, but also on your area and what matters to you. They're closer to the decisions that you need to make.

Also, recognize that sometimes it doesn't help to bring in a lot of data from other areas because sometimes there's something that's just simply unique to you. If somebody has-- if there is something that is specific to you, if there's a particular architecture that you have that doesn't parallel elsewhere, you need to make decisions on how to use that architecture, leverage that architecture, it may not translate to other environments. Definitely, if you map out that decision cycle and then levy that as a requirement on a service provider, I think you'll find that they will be able to provide situational awareness to you.

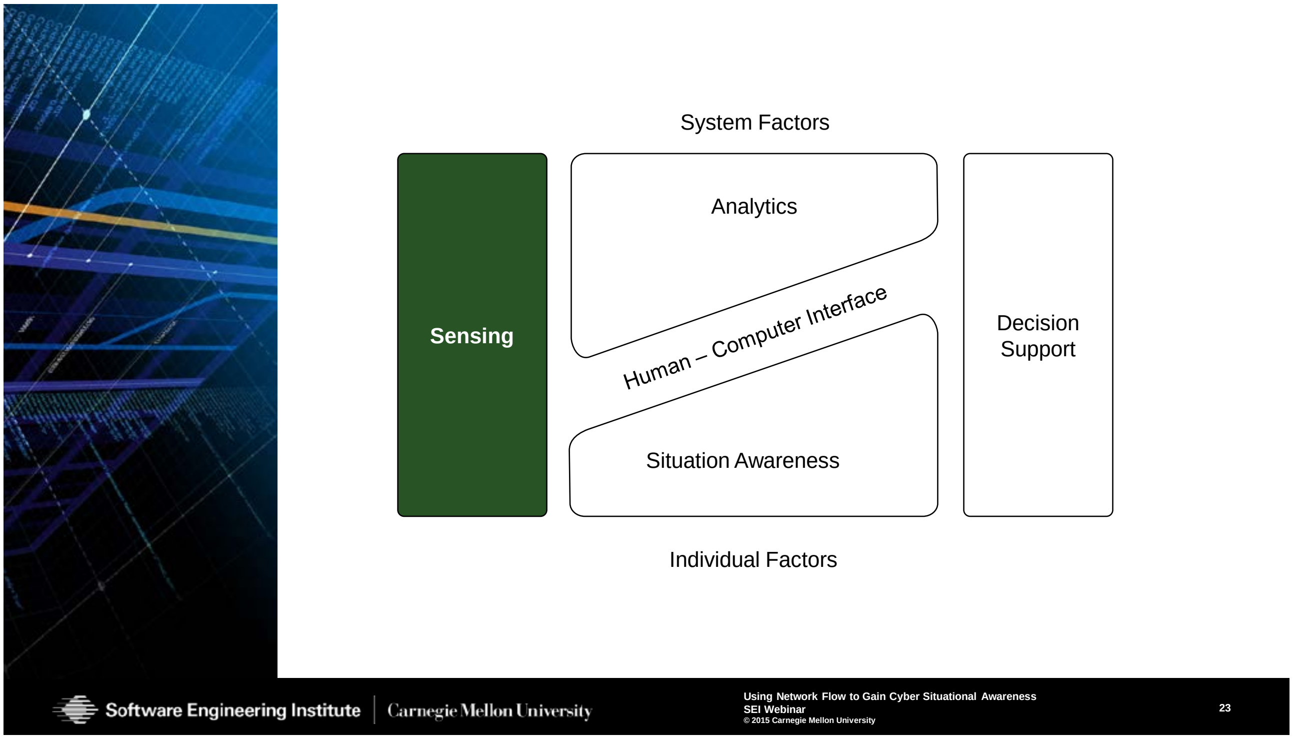
Shane McGraw: Okay, question for me. Someone asked me to repeat the URL for the FloCon conference. So, once again, it's www.cert.org. And it's FloCon, F-L-O-C-O-N. That will take you to the website for the conference.

Next question from William asking, "Can you recommend some good network visualization or GUI tools? Not everyone loves console text. And I've found iSiLK to be a bit less than friendly.

Sid Faber: So, I'd have to agree with you there. Can I recommend some graphical tools to work with? One of the tools I think that I've seen that's made a lot of headway in this area that's an open source tool, it's free for you to implement out of the box and download from the Internet, is the ELK stack, Elasticsearch Logstash and Kibana. It becomes a really nice toolset. And it really fits, I think, what we're talking about today. So, you can choose to implement. And there's different-- there's a lot of help online about how to-- even specific to flow data, how to get flow into this ELK stack, the Elasticsearch Logstash and Kibana stack.

One of the things that really appeals to me about that solution for visualization is that it supports this whole overall architecture--

Sensing



System Factors

Sensing

Analytics

Human – Computer Interface

Situation Awareness

Decision Support

Individual Factors

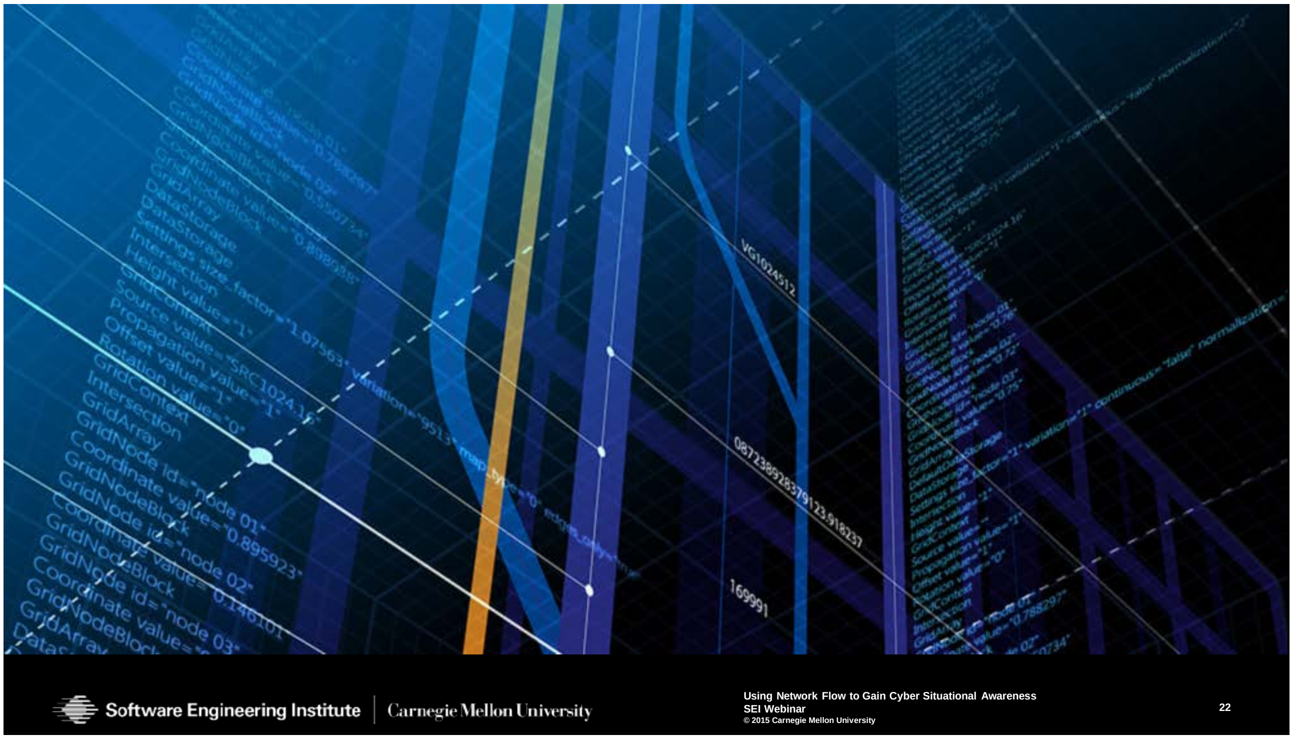
Software Engineering Institute | Carnegie Mellon University

Using Network Flow to Gain Cyber Situational Awareness
SEI Webinar
© 2015 Carnegie Mellon University

23

**023 That we're going after today.
It presents to you--

Situational Awareness



Using Network Flow to Gain Cyber Situational Awareness
SEI Webinar
© 2015 Carnegie Mellon University

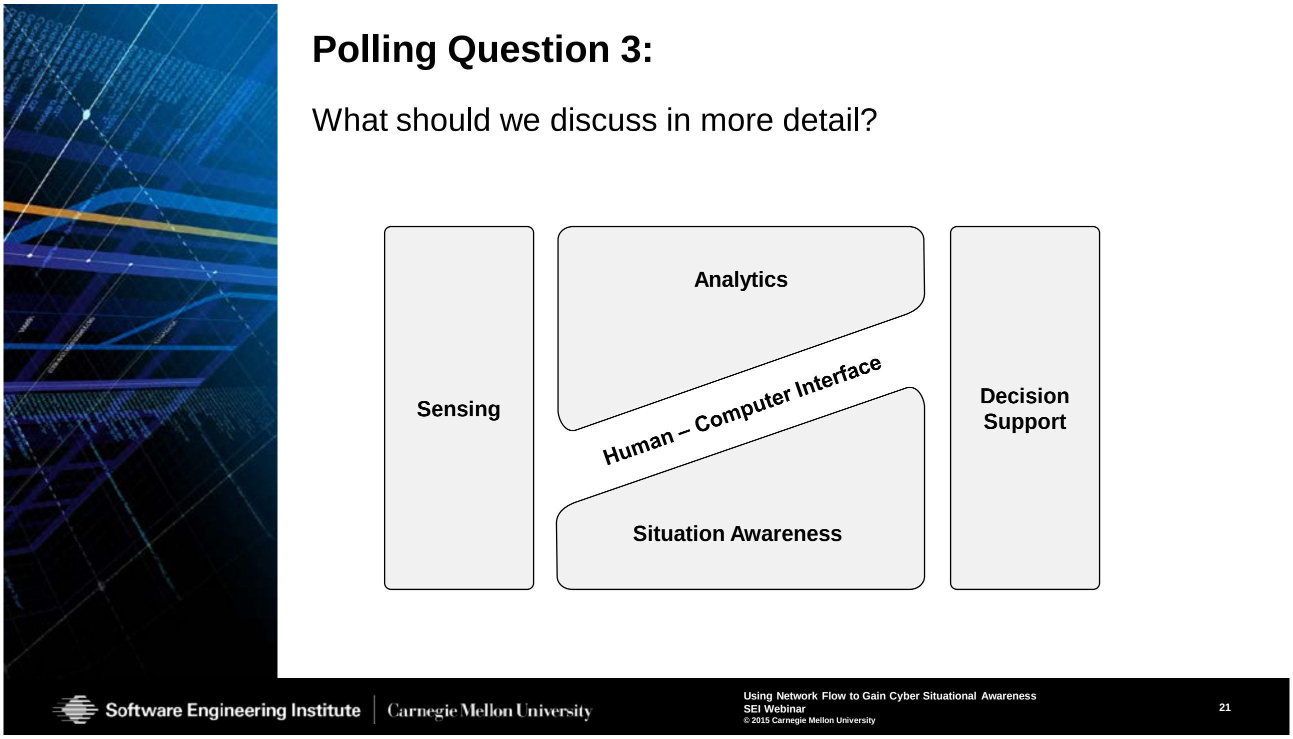
22

**022 The analytic side of it where you can put analytics into your Logstash, into your Elasticsearch instance, and it will work on the data for you. Logstash provides you a way to get data into that environment. And then Kibana presents you that human-computer interface. One of the things we didn't talk about is how you actually tune that interface to support situational awareness. And there's feedback there. Drilldown becomes critical in maintaining the environment. All those are provided by that instance. It becomes a really, really nice tool.

There are other commercial products out there. I'm not terribly familiar with a lot of them. I know most of the folks that we deal with have had a lot of success using ArcSight to, again, bring that knowledge from the system back to the user and present the user with what they want and also support that drilldown.

I think probably in a final answer to that question, one of the most critical things that, whenever you look at a tool, the thing that you must have is that drilldown ability. I remember working with an analyst one time, and we were doing an analysis of a SEIM tool. And he said to me, "I hate environments that aren't clicky." And it took me a while to process that. But then I realized what he was talking about is there's-- you need to have that ability to drilldown. Even though the system may present you with very rich situational awareness information, you need to drilldown from time to time into the more detailed, more granular information just to gain confidence in what the system is presenting you and understand better what's going on in the environment.

Polling Question 3:



Polling Question 3:

What should we discuss in more detail?

Sensing

Analytics

Human – Computer Interface

Situation Awareness

Decision Support

Software Engineering Institute | Carnegie Mellon University

Using Network Flow to Gain Cyber Situational Awareness
SEI Webinar
© 2015 Carnegie Mellon University

21

**021 Shane McGraw: Okay. Question from Ron, while we're on tools, asking, "Are there tools that examine flow records in real time or just retrospectively?"

Sid Faber: So, that's a great question. Are there tools that examine flow records in real time? I know for our part within the SiLK toolset and the network situational awareness tool suite that CERT provides open source, we have a tool that's called Pipeline. I mentioned it briefly beforehand. But Pipeline will take you flows as they come in live off the wire and do some very basic analytics off of them and then present those in a way that the user can consume them. So, that streaming analytics fits.

There are other tools in that space. You'll have to forgive me. I don't recall them off the top of my head because we tend to use our tools in most of our implementations. But I know this is a rich area. And again, for that live thing, that category that you're looking for is that streaming analytics. And also recognize in general you're talking about data plane analytics, which is very different, again, from control plane analytics and also usually different from inventory as well.

Shane McGraw: Okay, another one from Laurie asking, "Where does sensing end and analytics begin? What is the dividing line?"

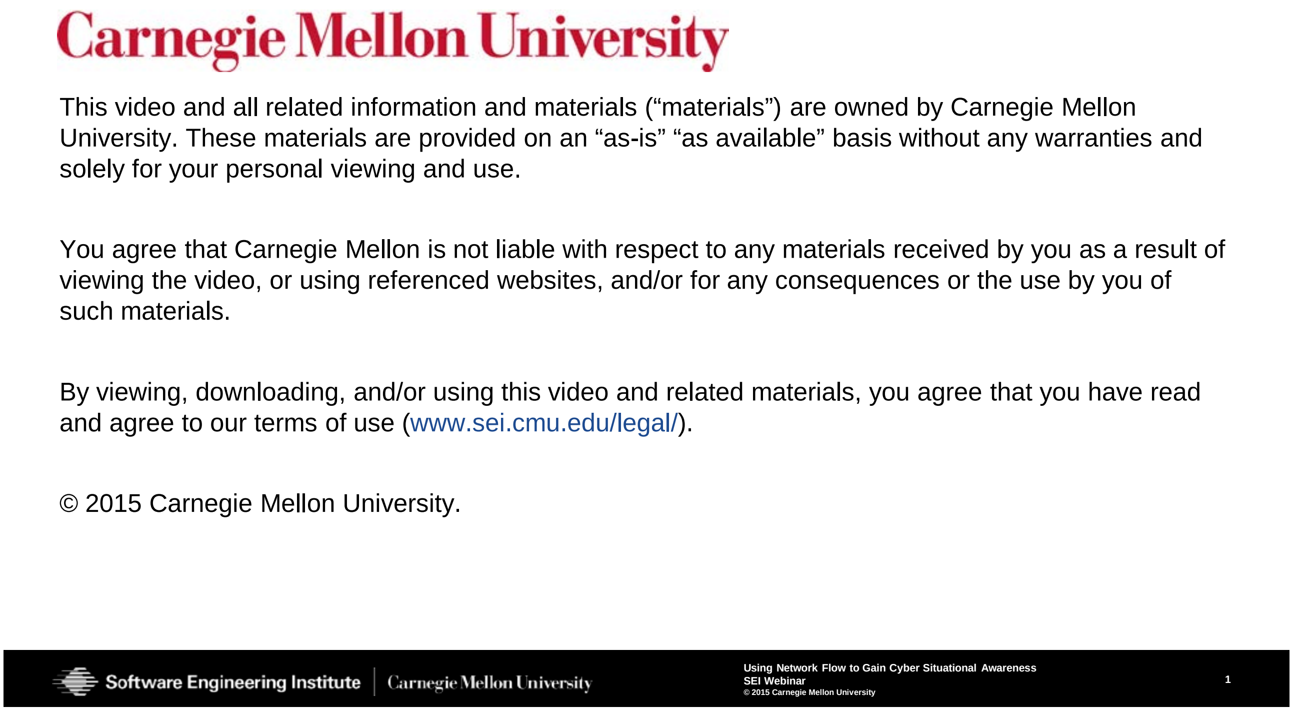
Sid Faber: So, that's a great question. Where does sensing end and analytics begin? The reason why it's such a great question is because we look at where we're heading. In today's environment, we're getting more and more connectivity, more and more networking. The idea of sensing every packet that flows on the wire is-- will not happen in the near future even if it can happen today. So, that means that I need to opportunistically put my sensors where I can. And I need to push some of this out to my sensors. So, what we want to see in the near future is actually start pushing those analytics out to those sensors.

As far as terminology goes, I would like to say that sensing creates measurements. Measurements are combined with analytics to create metrics. So, I would say sensing ends when I have measurements. And once I have those measurements, I combine those with analytics to create metrics. But again, if you're talking about the actual physical system, what is a sensor? More and more, we need to look at pushing some of those very basic analytics all the way out to the edge simply because we need to. We can no longer afford to bring everything back to a central location.

That begs a longer discussion about how analytics have to be structured in order to support that. there's a lot of really neat science in there. And I'd love to talk more about that if the opportunity permits.

Shane McGraw: Sid, excellent presentation. Folks, we're about 2:29. So, we're going to wrap up our questions there and just remind everybody, upon exiting today's survey that you please fill out the surveys as your feedback is always greatly appreciated. And lastly, just a reminder, our next webinar will be next Tuesday, December 1st where our topic will be structuring the chief information security officer organization by Nader Mehravari and Julia Allen. And everyone will get an invite to that. Thanks and have a great day.

Carnegie Mellon University



Carnegie Mellon University

This video and all related information and materials ("materials") are owned by Carnegie Mellon University. These materials are provided on an "as-is" "as available" basis without any warranties and solely for your personal viewing and use.

You agree that Carnegie Mellon is not liable with respect to any materials received by you as a result of viewing the video, or using referenced websites, and/or for any consequences or the use by you of such materials.

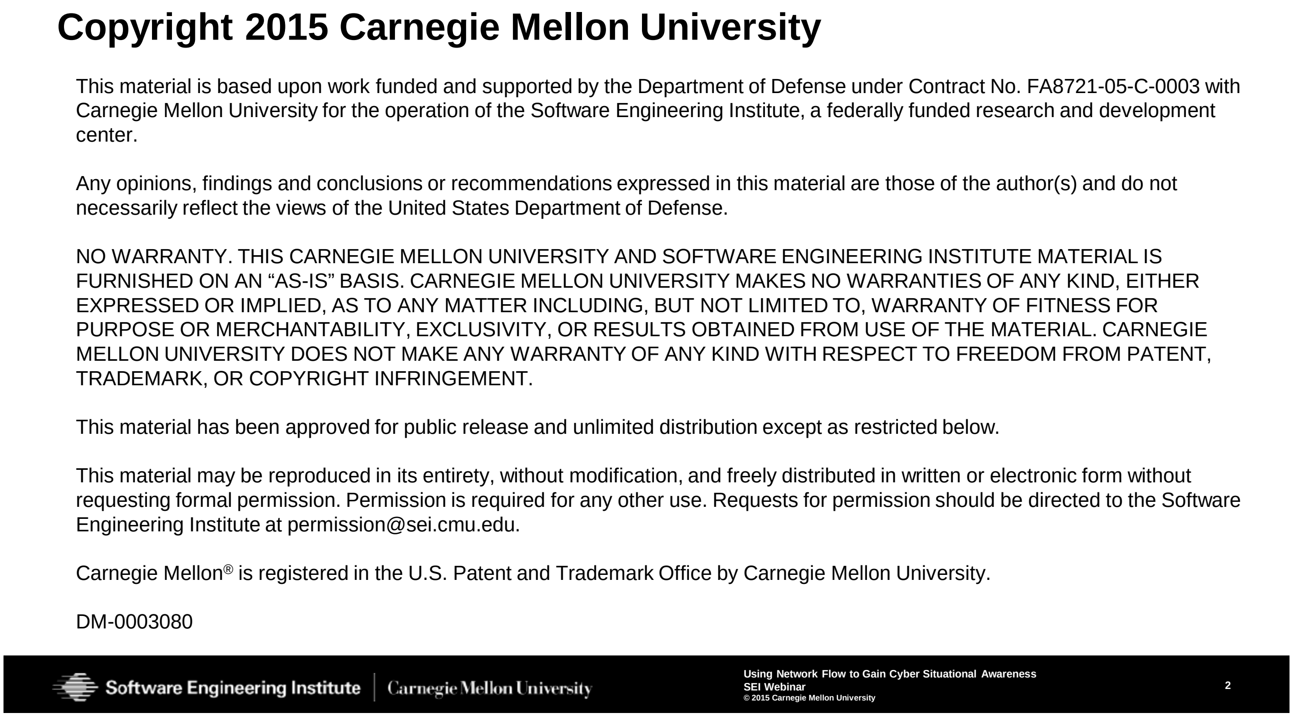
By viewing, downloading, and/or using this video and related materials, you agree that you have read and agree to our terms of use (www.sei.cmu.edu/legal/).

Software Engineering Institute | Carnegie Mellon University

Using Network Flow to Gain Cyber Situational Awareness
SEI Webinar
© 2015 Carnegie Mellon University

1

Copyright 2016 Carnegie Mellon University



Copyright 2015 Carnegie Mellon University

This material is based upon work funded and supported by the Department of Defense under Contract No. FA8721-05-C-0003 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center.

Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the United States Department of Defense.

NO WARRANTY. THIS CARNEGIE MELLON UNIVERSITY AND SOFTWARE ENGINEERING INSTITUTE MATERIAL IS FURNISHED ON AN "AS-IS" BASIS. CARNEGIE MELLON UNIVERSITY MAKES NO WARRANTIES OF ANY KIND, EITHER EXPRESSED OR IMPLIED, AS TO ANY MATTER INCLUDING, BUT NOT LIMITED TO, WARRANTY OF FITNESS FOR PURPOSE OR MERCHANTABILITY, EXCLUSIVITY, OR RESULTS OBTAINED FROM USE OF THE MATERIAL. CARNEGIE MELLON UNIVERSITY DOES NOT MAKE ANY WARRANTY OF ANY KIND WITH RESPECT TO FREEDOM FROM PATENT, TRADEMARK, OR COPYRIGHT INFRINGEMENT.

This material has been approved for public release and unlimited distribution except as restricted below.

This material may be reproduced in its entirety, without modification, and freely distributed in written or electronic form without requesting formal permission. Permission is required for any other use. Requests for permission should be directed to the Software Engineering Institute at permission@sei.cmu.edu.

Carnegie Mellon® is registered in the U.S. Patent and Trademark Office by Carnegie Mellon University.

DM-0003080

Software Engineering Institute | Carnegie Mellon University

Using Network Flow to Gain Cyber Situational Awareness
SEI Webinar
© 2015 Carnegie Mellon University

2