#122 - Methodologies for Analysis (with Christopher Crowley)

[00:00:00]

[00:00:12] **G Mark Hardy:** Hello, and welcome to another episode of CISO Tradecraft®, the podcast that provides you with the information, knowledge, and wisdom to be a more effective cybersecurity leader. My name is G Mark Hardy. I'm your host for today, and we're going to be talking about methodologies for detecting breaches and on our show today.

I've got a good friend of mine, a man I deeply, respect Chris Crowley. Chris, welcome to the show.

[00:00:33] **Christopher Crowley:** thank you very much. Really appreciate you having me

[00:00:36] **G Mark Hardy:** Well, Chris, for folks who have not had a chance to run into you, tell us a little bit about your background. What are you doing now and what did you do, and things like that.

[00:00:44] **Christopher Crowley:** Yeah. So right now I'm doing consulting. I do a solo consulting gig through my company Montance®LLC. I teach a class through that. That class is on cybersecurity operations. So I think that's for me the expression of all of the effort that I've put [00:01:00] in through my career to try to solve the problem of misalignment of cybersecurity operations with an organization teach that.

I teach for SANS Institute. I do a lot of research. I tend to do talks at a lot of different cyber events. In the past I've worked for government agencies, number department of Defense in the US. I've worked in private organizations. I have scale basically across. Probably every market sector I have worked in an organization either as an employee, a dedicated contractor, or a contractor as a consultant.

So, my background really started in computers when I was a kid. I grew up near Boston, Massachusetts. And my first job where I was like actually getting a paycheck, I mean, I did some stuff where I was doing mail merges from like my mom's friend to, for, for marketing when I was a, a kid on my Apple, two Gs.

But my first [00:02:00] job, I actually started when I was 15 years old and I did Unix. Specifically Altrix Systems administration and VMS Systems Administration primarily was there to run tape backup at the end of the day after the engineers had worked on stuff. So this was a bunch of geeks who had left BBN Bolt Beranek and Newman.

Maybe it's a company some people have heard of because they felt like it was getting too corporate in the eighties, . So

[00:02:27] **G Mark Hardy:** well,

bbn, they were the

[00:02:28] **Christopher Crowley:** the background of where I came from.

[00:02:30] **G Mark Hardy:** they turned off the internet back in 1988 when the Morris Room came out. Cause I remember they had like the gigantic off switch that later governments that are totalitarian wanted one of those

[00:02:41] **Christopher Crowley:** yeah, exactly. It, it, it's really interesting, like BBN, I'm sorry, I don't want to talk about this too much, but like, some of the stuff that I talked to about Talked with those guys about is like they would have stuff where they were doing network monitoring. And this is again, in the eighties, this is exactly what we're talking about.

And they provided the internet to everybody in that area. [00:03:00] And so they would have, these were the systems administrators that I was talking to that were. Helped to start that and so they , they were talking about stuff where it was like, yeah, when there was a network outage, they would figure out if it was a problem or not by basically calling the Fax Line that they knew was the fax line at that customer.

And then verifying that the Fax actually responded or not. And if the Fax didn't respond in a timely fashion, then they would they would basically say, oh, there's probably a power outage in that location, . So it's just like these very clever, intuitive things as a 15 year old kid dealing with a bunch of computer professionals at that time.

So that's where I came from. Right. So,

[00:03:39] **G Mark Hardy:** Pretty neat. So we got an early start and like a lot of us are probably in our teen years working on things far outside of the scope of what you'd expect. But back then it worked because there nobody cared.

[00:03:52] **Christopher Crowley:** Yeah, exactly.

[00:03:53] **G Mark Hardy:** you haven't been able to grow a mustache yet. It's

[00:03:55] **Christopher Crowley:** and they were given and they were giving kids like me root on all the systems, which maybe [00:04:00] in retrospect they cared about, but they didn't really understand , just how, how curious a 15 year old on a, on a wide open network would be. So lots of NNTP and and FTP time for me and lots of access to data.

So it's probably why I'm as weird as I am today.

[00:04:17] **G Mark Hardy:** Yeah. So no, no. So, so that set the hook. And so let's fast forward a little bit and now we get to the point where we're in the corporate world. We're, we're supposed to hold down a job and look responsible and, and work a nine to five paycheck and commute and get a house with a picket fence and everything else like that.

But that wasn't necessarily going to be your future, was it?

[00:04:38] **Christopher Crowley:** No, and I mean, I did that for a very long time and then ultimately decided that I wanted to be more self-directed. And the segue to that was really when I left New Orleans in 2005 because of Katrina I had already been starting to move into that arena of SANS Institute and teaching for them.

Based on my first experience with at attending a night session. in New Orleans, probably in [00:05:00] 2002. I'm from a sans event rolling through town. I was like, oh, this is really useful information. And the first class that I took technically was Hal's Unix class where I basically had this had this set of books.

And then I took intrusion detection because when I was working at Tulane University, We were having problems in the early 2000 era. That was when a lot of things started coming to the forefront of cybersecurity. So I established a anti-spam program , because we needed it , right?

That, that kind of thing. Where, where basically we were able to stand up defenses that didn't exist before. So an interesting era of cybersecurity.

[00:05:39] **G Mark Hardy:** Yeah, I remember doing stuff back in the day. I had a client, oh, what were they? They. , we were using Novell NetWare. And so of course you come in that way. It was all dial up practically, and you end up in the F drive. And I always tell people the F Drive and Novell network, for those who weren't there was like that train station in the Matrix where [00:06:00] you just, you can't get out of it.

You go around and you stay there. And at the time the question was, well, how do we make sure that. Somebody dialing into our system as up to date antivirus and it's like, well, okay, let's solve that problem because no one had solved it before. And I end up writing a whole bunch of DOS batch files that would basically kick off and it would query it.

And if the date would start and off it went and great. And now we come back and you look 15 years later and there's entire industries around being able to go ahead and do stuff like that. But like you, you just see a problem and you fix it. And then in a lot of our cases, we just moved on to the next problem, or, popped in our bicycle and rode home. And then somebody else.

[00:06:36] **Christopher Crowley:** Or fix it well enough. Fix it well enough, and that's it.

[00:06:40] **G Mark Hardy:** Yeah, and it's, it's good enough. I, one thing, trading stories back and forth, so, at a different customer they were doing CC mail. I dunno if you remember

that

[00:06:48] **Christopher Crowley:** I do,

[00:06:48] **G Mark Hardy:** the day. And what CC mail does that, they're mail for, they called them routers. They weren't really routers network sense, but I was working for consulting firm and they did a, a firm fixed price bid.

For one of [00:07:00] the statewide utilities in, in Maryland. And so what happened was is that we were going to run it on OS two. Because that was going to be the really first multitasking operating system. And they said, Hey, we've got all the stuff that'll work there. And we figured, well this is great. We can just have seven of these things running and that will take place and we'll have all these different sessions and it'll work.

Well, turnout was vapor wears with the, we're supposed to run, but we're stuck on fixed price. And now I gotta roll out 180 of these things for the price of eight or seven. And we figured it out and but one of the problems was, that you had to leave this thing setting out there at one of these stations all around there, these utility companies.

And it's an unattended terminal with admin access that doesn't sound like a recipe for long-term success. And so I said, well, okay, well what would I do if I had to solve this because I got to have a lot of money extra. So I said, let's set up a little suma force. I had one machine that was running. Put it in their data center.

And what would happen was you had a shared drive and if you needed something done, like to do [00:08:00] a file backup or a consolidation, which all those required admin privilege, you just set a little flag in there, which is a text file with, what am I, I'm station number 47 and I need this, and so, Once every minute this thing would scan through.

Are there any files in there? Ooh, there's one. Go do this. That the other thing, delete the file, and off you go. And so what happened then is that the worst you could do is if you got a hold of one of these machines out there is you could trigger a privileged action that was supposed to happen anyway, like backup, but you couldn't have arbitrary command control.

Okay? No big deal. Fast forward three years. , I'm on my own now. I get my contract with the same thing and I'm migrating their EDI system from Windows into Linux and. So when I get into the data center, it's like, wait a minute. I remember I walked over and looked at the screen. It was my program. It was still running.

Three years later, it had not stopped because I found little TSR for those who don't remember, Terminate and Stay Resident code. That basically [00:09:00] said if you saw C colon, back slash greater than on the screen. It meant that somehow it popped out of the batch loop, reboot the machine, and that's what kept it back and going again.

And it's just little innovations like that, that you look at 'em from almost 20 years back and you go, how do we pull this stuff off? Because today, Let's face it, we become tool masters. If you're going to do threat hunting or something like that, or detect a breach, you're not sitting there like it was in the early days.

Mono a mono banging away on the keyboard, seeing who could get whomever.

[00:09:31] **Christopher Crowley:** Well, it's an interesting construct. The other way to think about that. Computer systems are specifically designed to be flexible information processing, and they are not designed to be dedicated task processing. So when you build something with a little bit of resilience in it, the, the computer system, the information system is specifically designed.

To be able to absorb new and adaptive constructs that you feed to it, which is exactly what [00:10:00] you're talking about. I've got this general purpose thing, and if I make it just a little bit more flexible, then it, it likes that and it, and it goes down that direction, which is actually one of the reasons why attackers are so difficult to remove from our systems because everything in the system is designed to be flexible and adaptive.

And if they're more flexible and they're more adaptive, then they will beat us. on our own systems because we cannot keep up with them.

[00:10:26] **G Mark Hardy:** and that's a very good point because, What happens also in our mindsets is that like any other trade, if you go back and you look at the traditional trades, I want to go ahead and go to trade school. I'm become a plumber, an electrician, a carpenter, or something like that. I can learn a basic set of skills, but what's going to happen over time is I'm going to perfect those skills.

I'm get better and better at that. I can become ultimately a master at what I'm doing. and, and so there's, a well defined career path. But what we're happening is as we're watching this whole thing blossom and emerge and change around us, we don't know what's going to be there five or 10 [00:11:00] years from now.

If somebody told you two years ago that you're going to be doing work fighting a chat G P T, automated attack tool, that's science fiction. And yet here we are playing around with this stuff.

[00:11:12] **Christopher Crowley:** It's really interesting. I had a conversation decades ago with a microscopist because when I went to undergraduate he was he was, Retired and he was working in a laboratory that I was basically a maintainer of microscopy lab. And I remember having this conversation with him one time and I wanted to talk with him about some problems.

And he said, well, what do you want to talk about? And I basically said, well, like interpersonal stuff and politics. And he said, Hmm. I'll come and talk to you, but we're not going to talk about politics and we're, and we're not going to

talk about interpersonal relations. I'll teach you methodology because by teaching you methodology that will actually serve you far greater than sort of this momentary concern that you have.

So it's a, it's a lesson that I learned a long time ago that was really nice in terms of what to [00:12:00] focus on and the important aspects of things.

[00:12:02] **G Mark Hardy:** Well, since you mentioned methodology, and coincidentally, it doesn't happen to be the title of the episode. We might want to to get to that after this long preamble. Hopefully everybody's still with us, but this is always fascinating. That's why I love having friends on the show because you, you learn about some amazing backgrounds and hopefully other people can say, well, that's pretty cool and I'd love to keep listening to this guy.

So when we talk about methodology, let's kind of start well with the basics. Like, what do we mean by methodology?

[00:12:27] **Christopher Crowley:** Mm-hmm. . Yeah. So methodology generally is, is an approach to problem solving. You have, you want to fix something, you want to address an issue. It is what rigor will we apply to this so that we can be directed in our effort?

[00:12:43] **G Mark Hardy:** Okay. And so what we have then is more of a strategy for problem solving as compared to, well, I guess even randomly trying things, pushing buttons like Homer Simpson, eeny meeny miny moe. Technically is going to be a methodology. It's not [00:13:00] a recommended

[00:13:00] **Christopher Crowley:** push buttons and I just push buttons until it's until it's actually working is a methodology. It's not our preferred methodology, , but, but it is a methodology which some people can employ or it's the only method that some people can employ is a way to say that. So we like, we like other methods.

And it's something that I actually want to say. First of all, the reason why we would choose to have a methodology, which may be a foregone conclusion to a lot of people, but if we articulated it helps us to be able to decide what the right one is, is the reason why we have a methodology is to make sure that we have an overlooked stuff and to make sure that a number of different people involved in a team can work together in an efficient way to come up with stable, reproducible quality in the results that they're that they're tasked with in the face of novel circumstances. Exactly what you're saying. We don't exactly, know this and so we're [00:14:00] able to be adaptive.

And apply our methodology and circumstances that we maybe we hadn't thought about before, which is why we have the method to make sure that we address all of the concerns and have a, a stable point of reference for, for things in the future.

[00:14:16] **G Mark Hardy:** We'll talk a little bit about some ideas and methodologies, because of course, as I say, kind of. , the base case, which is well just push buttons till something works. And we say, and not preferred, but it, it does qualify as a methodology because you can define it and you could follow the process.

And again, a methodology is not guaranteed to give you a, an answer that you like. , right? You, you may come up with the saying that, Hey, my hypothesis was wrong. And I, there's no way I can get there. So, so let's keep with this definition a little bit about what do you mean by a hypothesis?

[00:14:50] **Christopher Crowley:** Okay, so, so hypothesis is a, is a, a, a structuring of a, of a statement around a circumstance. [00:15:00] And usually the hypothesis is an expression of your thought of what might be a plausible explanation based on some, some data that you have. So let me, let me jump into a couple of things. So first of all, there are different techniques that we leverage in order to articulate hypotheses.

So scientific method is one that every, everybody goes back to all the time. And a lot of times people talk about this in the context of cybersecurity. I don't like the discussion of scientific method in the context of cybersecurity, because in practice we don't actually follow the scientific method, which would be an articulation of a hypothesis, and then a construction of an experiment intended to disprove that hypothesis.

And then if we disprove that hypothesis, then we throw it away and we seek another hypothesis. And so it's this sort of iterative approach. A stable hypothesis, which eventually [00:16:00] becomes Ethereum because we haven't been able to disprove it. And, and in practice we don't actually work like that in cybersecurity because we don't have the time and the resources to be able to come up with these sort of more, more grand notions of, of theories.

We usually have limited time. We have to come up with. Best explanation of available data as opposed to a sort of more theoretical, expansive expression of a construct that allows us to work in the future. That's not what we need in cyber. What we usually need in cyber is this is something that is intentional right now that is good enough for the purpose. And, and this is largely in the

context of investigation. Now if we're talking about detection engineering, say, then maybe, maybe we could do a different approach where it's more of an engineered approach and we might leverage some sort of a, a hypothesis construct to be able to say, here's the thing that [00:17:00] we're going to say we can do.

And now lets try to break it as much as possible until we can say, look, we've tried everything and we can't break it. Okay. That's the one that's closer to scientific method than what we do in investigation and sort of more immediate responsive action. So let me just continue to say one of the thing I mentioned a methodology.

I don't. Talking about scientific method, I do like analysis of competing hypotheses from Richards J. Heuer, Jr. And the Central Intelligence Agency book. Psychology of Intelligence Analysis. He talks about this analysis of competing hypotheses, and he kind of presents this methodology, which I'm going to give you a very simplified picture of that.

He says, you have some situation that you need to deal with. You don't have full data, you don't have clear parameters of accuracy. You need to come up with the best explanation. , what do we do? We look at the data that we have, we brainstorm some, some potential explanations. We come up with an expansive list [00:18:00] of potential explanations.

Based on that, we go find the rest of the data that might be available to us, and then after we sort of reconcile the data that might be available to us. In cyber, we talk about this as visibility, right? We seek greater visibility. Then we have these hypotheses and we kind of consolidate and work on it.

And then this is something that. Almost never seen people in cybersecurity. Do we have rigor in our assessment of the hypotheses in light of the data in a quantifiable way. And, and basically this is pretty simple. At at, at its simplest, make a spreadsheet. Here are the columns of my hypotheses, here's the data that I have.

I'm going to score it based on each data element for the specific hypothesis. And then I've articulated, and in Heuer's term, externalized. The construct of my investigation and my analysis, which allows me to do sensitivity assessment [00:19:00] of what points of data do I really depend on, and it also allows me to compare between different analysts what different people think, and I then have a way to express either. My analysts agree or my analysts disagree, and if I have an individual assessments, that's an individual assessment that's performed

between a number of different people, and I can do that comparison. If they're all sensitive to the same data elements and they all come up with the same conclusions, based on the same dataset and the same hypotheses, well then I have very high confidence in the expression of that without having to go and spend a tremendous amount of time, I can do this relatively quickly. So it's a faster method, which introduces its own challenges, but absolutely provides value in the sort of situations that we tend to encounter in cybersecurity.

[00:19:58] **G Mark Hardy:** So when you talk about that and [00:20:00] building that spreadsheet and saying, here's what my requirements are, here's the data, let's fit it. I think some folks might be thinking, okay, that how, sounds a little bit like the MITRE ATT&CK® Framework, where they got tactics and techniques, and that might also be like a NIST 800-171 or an ISO 27,001 and I'm just going down the list. Going down the list, but that's not really what we're talking

[00:20:19] **Christopher Crowley:** about

That's not what we're talking about. So you said two things. You said, here are my requirements and let's fit it. And it's actually far more open-ended than that. I will say that MITRE ATT&CK® Framework is also just a two-dimensional array, right? I mean, so it's like we use two-dimensional arrays a lot for our decision making, and it's a perfectly valid approach.

It's a simple approach. We could add additional dimensions if we wanted to, but for the purpose of the sort of timeliness that we usually need in an investigation and a response and a decision that we would make, two dimensions is going to be good enough. But let me just sort of characterize this. More intentionally for the [00:21:00] terms that you use to make sure people understand it.

Number one is we brainstorm hypotheses, and the hypotheses would be columns. In the spreadsheet we identify available data. And the data is the rows in the spreadsheet, and then the analyst's task is to assess based on some available quantity per row, a weighted ranked expression of which hypothesis is most supported based on that specific data element that there.

So this is the structure that he establishes. We're not fitting it to something that's an expectation. We're articulating our thinking on what we know based on the data that we have and the hypotheses that we have agreed are within scope. For this particular exercise of thought,

[00:21:54] **G Mark Hardy:** That's where going to be. The big difference between doing this compliance type of approach that I was talking [00:22:00] about is that you have a set goal.

[00:22:01] **Christopher Crowley:** nope. No goal.

[00:22:03] **G Mark Hardy:** to these type of methodologies where you say, I've got a situation and I think this might be the answer to this or this, but you're going to find out and you're going to be data driven to determine what is really happening here.

[00:22:14] **Christopher Crowley:** and flexible. And so the hypothesis, if you read Heuer's explanation of it, the hypothesis generation stage is intended to be incredibly flexible. And he talks about the things that analysts tend to do as soon as they have a certain thought, that's the thought that they're stuck on, and they keep looking for data that.

Matches the hypothesis that they've already arrived at as a foregone conclusion is the best explanation, and this is one of the reasons why he says you need to use this methodology, is if I say, okay, brainstorm all your explanations, and the analyst comes up with two and it's a. A with a slight variation, and I go back and look at the analyst work, well then I know that they've encountered this foregone conclusion problem.

But if I go and look at the analyst work and they've come up with 20 different [00:23:00] hypotheses, and then they work through the notion of consolidating this and finding other data elements based on the hypotheses that were added at the brainstorming phase, they've done a better job. And I have more confidence in the work that they did in terms of that analysis.

So you want to, you want to get get into like a specific example Cause we're, we're talking about this and I think that this is a, a good one to help to illustrate it because I can talk about it theoretically, but we can, we can do a very specific one.

[00:23:29] **G Mark Hardy:** All right. So, yeah, one last thought on that one. I was on a call last night and someone was mentioning that he was working at a high level organization in in the government. And he said when they came up with a idea or whatever and when everybody thought it was a great idea, it was almost universally wrong.

[00:23:45] **Christopher Crowley:** Yes,

[00:23:46] **G Mark Hardy:** was a combination of group think, a lack of testing the hypothesis, a lack of rigor. And, and so he had learned as a senior executive is that when you bring something in and everybody says, yes, you need a disruptor, you need to find some way [00:24:00] to say, okay, push the reset button and then re, control alt, delete your group, and said no.

And so if we look at the history books, my understanding was that was what Bobby Kennedy's role was back in the Cuban Missile Crisis.

[00:24:12] **Christopher Crowley:** There's a team, there's a team role name for that, and I think it's the 10th Man or something like that, where it's like, there's it, it there is a role that the only thing that you're allowed to do is disagree with everything.

[00:24:23] **G Mark Hardy:** Right. And I, I mean some people I know then that's their just mo all day long We have other names

[00:24:31] **Christopher Crowley:** Sure. Exactly. Don't always invite them to the party, but Yeah, exactly. Yeah. From a, from a, the, the notion of being stuck in the oh, well that must be it because those smart people are, are addressing it then Yeah. Having somebody that is the disruptor is in is quite valuable, but just. To that end as well, that disruptor would say, okay, no, that's not the hypothesis.

No, that's not the hypothesis. We don't want that. We want let's, let's get all of the things in the disruptor's role in the hypothesis [00:25:00] brainstorming would be, here's another one. You haven't thought about aliens from Mars. You haven't thought about zombie apocalypse, or whatever it is. Just these sort of like real outside expressions to help.

Break people out of the, the, the constraint of non-creative thinking.

[00:25:18] **G Mark Hardy:** And that's really important as a leader. I think I mentioned it one or two shows ago about, I put an officer on my, team who always saw things differently than I did. And I told her, I said, you're going to have the toughest job here because you're going to be out voted nine out of 10 times.

But what out of 10 times you're going to save the ship. You're going to come up with something that everybody else did not think of. And because we have a respectful exchange of ideas and you bring this thing out and everybody's going

to go. Whoa. Wow. But as they say, when you're nine out of 10 times, you're getting voted off the island.

It's, it's tough emotionally. And so for those disruptors or whatever title we give them, you have to also make sure they've got the resilience to [00:26:00] be able to interact in a way where you know that your one lost record is always going to be very terrible. But when you win, it's a decisive win. I mean, you, you strike out nine out of 10 times, but then you hit a grand slam.

[00:26:11] **Christopher Crowley:** And

[00:26:12] **G Mark Hardy:** nine more times in a grand slam. That's

[00:26:14] **Christopher Crowley:** Yeah. And in the expression of analysis though, we don't, so to me, in my thinking, I don't want to talk about win or lose because in these sorts of, Questions we will never know. We will never know, right or wrong. It's not a did the, did this happen or did that happen? It's a, there's something that happened and we're trying to explain it and we will never know authoritatively if we are right or wrong.

That's just the, that's just the circumstance that we find ourselves in and that's why Heuer says we need to have this structure. Okay. And so again, this is one there. Let's one, one other thing I do want to mention it's not one that I use a lot, but it's one that I've been looking at more [00:27:00] is decision theory.

And so it is a methodology of being able to actually identify optimal decisions. And I still don't think that that's right for cybersecurity operations, but it's just one other one. If someone's hearing this and like, okay, well let me go get into these different methodologies.

Comparing and contrasting decision theory from analysis and competing with hypotheses from scientific method is absolutely a valuable thing to do. So you can pick the right aspect of what you want your staff to do. And you mentioned Mitre ATT&CK® Framework. Also really good for things, investigation, response detection, engineering defensive gap analysis, all sorts of things you can do with that too.

But I feel like we're more in the kind, like the hunting investigation kind of mindset when we use ACH and competing hypotheses.

[00:27:45] **G Mark Hardy:** So let's, let's take a threat hunting type of an example. Let's say there's been a potential for a breach and you're trying to

figure out we're concerned that maybe somebody has exfilled data from our organization. Maybe we use that as our. Are null hypothesis. [00:28:00] And though from that we'll say, okay one of them could be.

They're using DNS and so we'll use our favorite Port 53. So let's, let's start that. For example, let's say that was on the table and you say, well, we're, they're getting stuff out, but we're not sure how they're getting it out, but we're pretty sure that there's something going on. So what would be a, an methodology approach that we could use being a little bit more specific, looking at DNS so that we can talk about it from a threat hunting.

[00:28:26] **Christopher Crowley:** Sure, sure. So, so part of, we're, we're sort of combining the structures of threat hunting as a, as an act, and then the explanation of, hey, there's a problem. Right. We, we, we have some clue that, that an attacker has taken things out. We need to investigate that. And one of the potential explanations is that they're using DNS command and control to be able to do that.

And so now we're [00:29:00] looking for data that would assist us in investigating that. So, so now we need visibility aspects of where does that data live? And immediately you mentioned UDP 53. So we could look at network traffic on UDP 53 and TCP 53 because zone transfers and it doesn't have to be zone transfer can still run over TCP 53.

We would look

[00:29:22] **G Mark Hardy:** it was over a thousand bites or something like that. It

[00:29:23] **Christopher Crowley:** Yeah, exactly. We would look for data. We would look for data on the DNS servers. We should also look for data on the host themselves. There might be sys logging I'm sorry, windows event logging or all of these different aspects. We could look at cache data on the host.

So now we have all these different potential data sources. By the way there are two other there are multiple other DNS protocols, but there are two other ones that we should be concerned about with respect to exfiltration of data. That's DNS over HTTPS and DNS over TLS. So DNS over TLS is going to be a Port [00:30:00] 853 that we need to concern ourself with DNS over HTTPS is actually over port 443 or any port that you want it to be, and it looks like a fully normal HTTPS connection.

So all of these are places that now we need to consider if there's data moving through that. Then as we go and gain visibility into the data that's available to us, then we would look for attributes within that data for elements that help us to make a decision. This is benign and I can remove it from my view and investigation.

This is suspicious and I should continue to characterize it. Or this is malicious and, and now I have a problem. And so the expression there in our threat hunting is, Interesting for most analysts because again, they're, they're tuned to, we are not, not just analysts, but human beings are tuned [00:31:00] to. Soon as you see something that kind of matches a pattern, you're going to click into that pattern and make sure that you articulate very quickly.

I'm right, , right? So, so the, let's

[00:31:12] **G Mark Hardy:** Confirmation bias, I think

[00:31:13] **Christopher Crowley:** Yeah, yeah, yeah. Let's, let's talk about some of the elements that we typically use in our DNS investigation in order to kind of assess if this is malicious or suspicious. So one thing that we would do with this is say, okay, let's look at history and look at all the things that we've looked at previously. Say that that's automatically benign. Is, is that a hundred percent accurate? No. Is it useful? Absolutely. So I could just basically say the domains that I have have seen. In my environment, 14 days in previous, 28 days in previous, 180 days in previous. Let me just exclude those. Now, you would [00:32:00] need to have either a table that lets you do that quickly or you would need to build that.

Table for exclusion. A actually, quick aside, I wrote a script that basically did this a decade ago to deal with DNS as a basically a vector of, of indication that there's a problem. And one thing that I did. Was, I just said, let me record everything every single day in terms of what we're what we're seeing.

And then on any day where I have something new, a domain, which I have never talked to before, let me look at it. And in a really big network that gets noisy, but there's still value to it because you could do that. Plus additional characteristics. So what would those additional characteristics be? Very high volume of communication, long host names weird response records.

Lots of invalid for a period of time, followed by lots of valid for a period of time. And maybe even sort of a switching back and forth in terms [00:33:00] of responses. In DNS, we would call it an NX domain, basically failed to resolve

and then all of a sudden the responses successfully resolved and then maybe flapping back and forth between them is a weird behavior, right?

Re-registration of domains to new authoritative name servers. Happens, but generally doesn't happen a lot in terms of domains that are in place. But this might be an indication of some sort of a domain takeover, maybe even for a for a short period of time. So just some thoughts about what an analyst would then go use as that preliminary differentiator now in the methodology in the methodical approach. That, that Heuer describes, all of that would be included. All of those data elements, anything that I'm thinking about, let's pull all of those into my, into my data attributes and still hypotheses. Thus far, were kind of [00:34:00] DNS. DNS Command and control DNS for exfill. And we could actually, if we wanted to stay within the DNS range, we could say DNS over HTTPS, DNS over TLS, DNS command and control using a number of different tools.

And we could actually just have a, a very detailed and granular set of hypotheses within the specific hypothesis of DNS in order to help to weight. What, what we're doing there, so part of this investigation is let's go find more data.

Yep.

[00:34:31] **G Mark Hardy:** Right. And so what we have then is all these different hypotheses would be, if you will, our columns there in the, in the spreadsheet where we're saying, Okay. This has got a lot of weird stuff. This is a very long query. This has non alpha numeric characters in the DNS query. When you're constrained to that a lot of fails and then it starts to succeed.

No, no. And, and then we just fit the data to the hypothesis rather than fit the hypothesis to the data or, which is what are we

[00:34:57] **Christopher Crowley:** Well, well, let me, let me say, instead of [00:35:00] fitting,

uh,

[00:35:00] **G Mark Hardy:** know, as soon as they said that word, I said, okay, I want, this is going to

[00:35:04] **Christopher Crowley:** yeah, yeah. So, so let me, let me say, okay, let's, let's presume we did our brainstorming. We've got a whole bunch of

hypotheses we've worked through this data collection for a while now the data elements, all sorts of them, all sorts of data elements are in rows.

Now, what I want an analyst to do is to consider each row, each data element. In turn, as independently as that human's brain can isolate that data element. . I want them to take a number of points and let's say that we have. Five hypotheses. So I'm going to say per row. Each analyst is allowed to assign a maximum of 20 points per row, and I want the analyst to grade the hypothesis on a weighted [00:36:00] basis for each data row.

And then each analyst should take each data row in turn and say, does this hypothesis explain this at the absolute best to the exclusion of all the other hypotheses. Okay? That hypothesis gets 20 points in that row, and the other hypotheses get zero points. Now, if hypothesis A is not feasible based on that data element.

But hypothesis B is kind of feasible and it's about as hypo as equivalent as hypothesis C, but hypothesis D is really good and hypothesis F it is really not feasible. Then my points in that row would be B would get five. C would get five and because I think D is really the best explanation, it would get 10.

So now it's 0 5, 5 10. And if I do this in turn and I'm basically just like ranking it based on my thought of that data element [00:37:00] explaining that particular hypothesis. As I do that, in turn at the bottom, my summary, my, summation function is going to say Chris thinks. B is the, is the most weighted explanation.

But he also thinks that C and D are, are probable explanations in A and F. Not really.

[00:37:22] **G Mark Hardy:** Now when you mentioned that you put a zero in there because you say it, it can't happen to me that. It adds up, but it might get overweight by other things that could happen. So in doing that approach, would I ever get something where I would put a big, like a negative infinity in there and just simply say, you know what, that disqualifies this because it's utterly impossible for that hypothesis to have happened with this fact set.

[00:37:44] **Christopher Crowley:** That A, actually, according to Heuer's method, and I didn't go into all the details, you would've excluded that hypothesis at the consolidation phase before you were even doing this grading. But sure, maybe, maybe at some point you're like, no, actually this. Irrevocably refutes this. But that's not [00:38:00] the that's not the expression typically.

It's non-negative numbers. Right. , it's like there's the constraint in, in the way that I'm describing it, but, but certainly there are exclusions that happen in our, in our generation of, of our hypotheses to where we, he calls it refining the matrix. You remove stuff, you knock it out, and maybe you have to go back and adjust the matrix after you're doing the analysis and you, and you knock it out later.

Yeah.

[00:38:26] **G Mark Hardy:** Got it. So essentially what we're talking about then is having a methodology. We have some competing hypotheses.

[00:38:32] **Christopher Crowley:** Mm-hmm.

[00:38:32] **G Mark Hardy:** take and collect the data. We then use the data and we score, if you will, the hypotheses upon, does this hypothesis account for this? Does it account for it extraordinarily well? Does it, it could happen or no, that's not, and and although I mentioned the zero or negative Infinity Zero may be more powerful because you might say, no, this hypothesis or this data doesn't account for this hypothesis, [00:39:00] but it doesn't necessarily totally disqualify

[00:39:03] **Christopher Crowley:** Yes. Yes, exactly. Exactly. And then, and then think about reporting this. Think about reporting this where you can actually now articulate your confidence in a specific explanation. And also express your confidence in other explanations, which you have considered. And, and the reporting becomes a lot. More robust in, in, in my opinion.

And, and I think for people who would be hearing that, you would say, we considered this, this, and this, and this. And our analysts assess their confidence. And you could also, if you wanted to get fancy with the math, you could also show the distribution from a specific explanation between multiple different analysts who have assessed the same hypotheses in light of the same data.

And we actually you remember this advertisement four out of five dentists agree, which is exactly what you were talking about earlier. Nine out of 10 of my, my officers agree, [00:40:00] but think about that. If rather than it's okay, well you got bounced out because you, you didn't.

Now if every. Had to articulate their work in expressing the, the why. The why, the why I think this, and this is exactly what Heuer says we need this for, is it because it, it externalizes that because some analysts are just going to yell

louder than others and some analysts are going to keep talking until nobody else is willing to argue against that analyst because they just don't take a breath.

[00:40:34] **G Mark Hardy:** And, and you've gotta also think of the group dynamics because if you've got somebody who's new and they come up and said, Hey, I've got. Yeah. And one of your old curmudgeons said, well, that's the dumbest thing I ever heard of you, Moran, or

[00:40:46] **Christopher Crowley:** Or whatever.

Yeah, yeah,

[00:40:47] **G Mark Hardy:** he's never going to speak up

[00:40:48] **Christopher Crowley:** crushes people's drive to, to express, especially people. Who either don't have confidence or just want to be analytical and they don't want to [00:41:00] fight the verbal argument. This actually helps to equalize that.

[00:41:05] **G Mark Hardy:** And the other thing that occurs to me is we talked this through and we're getting pretty close to the end of the show here, and, and it feels like we just go on forever, but we, we

[00:41:11] **Christopher Crowley:** Sorry, I could go on forever.

[00:41:13] **G Mark Hardy:** In fact we're going to bring it back on the show. There's no, there's no question about it. But

[00:41:16] **Christopher Crowley:** do more of a strategic view in the in the

[00:41:18] **G Mark Hardy:** that I do to this, because we have multiple analysts and we're looking at the same data and we're scoring different hypotheses.

But over time what I would do is a SOC leader or as a CISO is I was going to say, you know what? This person's got a better track record than that person, than this one. It doesn't necessarily mean I'm going to use that for hiring or firing decisions, but I might put a little bit more weight

[00:41:39] **Christopher Crowley:** You certainly could do that.

[00:41:40] **G Mark Hardy:** right more often.

[00:41:42] **Christopher Crowley:** You could do, you could do individual waiting. One other thing that you, you haven't mentioned, but I absolutely think that you would do in a SOC, is you would actually automate and orchestrate a number of these tasks already. Because if we keep having to go get that data, if we keep building hypotheses through brainstorming, Actually we can, [00:42:00] we can expedite a lot of this stuff, which is the context that I came up with.

Use of analysis of competing hypotheses originally was actually in an orchestration or automation context in order to help us to move faster and eliminate certain things that are, are repeated, but still maintain good analysis, still maintain the human being, and I can actually express where I want the human being in the process of doing the ACH and where I want the system automating the task.

[00:42:29] **G Mark Hardy:** It almost sounds like we could then go ahead and if we're looking at competing SOAR tools, that we could apply our methodology to say, Hmm, which one of us give us the injects when we want the injects, as compared to it just automatically goes past that. I had that conversation yesterday with a MSSP that they had an automate solution that they had, and I said, I don't get a chance to trigger my automated playbook until after your human analyst has made a decision to say, yeah, this looks genuine.

If I'm in a situation where I think I'm under attack, I'll take some false [00:43:00] positives and I'll put 'em on my table. So I'll look at 'em at the same time you look at 'em at your table. Cause you might have five other clients who are lighting up the switchboard. And so I need a bypass up and over that last step.

And they said, I've been asking for that for over a year. And they, I said, yeah, we're built, we just built that. So sometimes if you're a squeaky wheel, you can get vendors to, to adjust to what you're doing. Any, any last thoughts before you wrap up on this show

[00:43:20] **Christopher Crowley:** Just really appreciate the folks considering what I've said and thinking about how to fold that into their environment. I think it's an important aspect of people doing good cyber work.

[00:43:30] **G Mark Hardy:** Well, that's awesome. So, Chris Crowley, thank you very much for being part of the show. We'll put links to how people define you on Montance® LLC, montance.com into our show notes, as well as some links to some of these methodologies. For those of us who are just listening to

us on the podcast, hey, tune into YouTube. We're starting doing that at the beginning of this year, and we're going to make other short videos available soon, so subscribe over there as well so that we can go ahead and help people get the word out. But this is CISO Tradecraft. Until next time, stay safe out there.