2025-01-17--t11-02-05pm--guest437619--bill-enhanced-v2-90p

[00:00:00] **G Mark Hardy:** Hey, today I'm going to give you the dirt on a really interesting threat model called includes no dirt. And I'm rather surprised because it's probably the best model I've ever seen in terms of fitting in terms that spell something that really means something. And I've got the originator area of it here, right here on the show.

So stay tuned for CISO Tradecraft. You're going to love this one.

[00:00:30] **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, and I'm going to be your host for today. And I've got Bill Dougherty today on the call. Thank you very much, sir, for joining us.

[00:00:46] **Bill Dougherty:** Hey, G Mark. It's a pleasure to join you here today.

[00:00:49] **G Mark Hardy:** And I love the background. We were talking before the show, he's out in Napa Valley and I'm not, but anyway, that looks pretty cool. But before we get rolling, a quick reminder, because coming up shortly is [00:01:00] CruiseCon. It's the flagship of cybersecurity from the 8th to the 13th of February coming out of Port Canaveral, Florida. And Admiral Mike Rogers will be there. I'm going to be there. A bunch of senior CISOs are going to be there. What a great opportunity for networking and getting a chance to get away from the cold and the snow if you're from up north.

So go to cruisecon. com and if you decide you want to come join us. Save yourself some money. Use the code CISOTRADECRAFT10 and you'll be able to get a discount.

Here's an awesome tool that will help you leverage AI to detect security vulnerabilities in your software. ZeroPath is a SaaS tool that can help you secure JavaScript, Python, Go, Java, C#, and PHP, and their wall of fame lists the open source vulnerabilities that they found already, and you can put the same world class tool to work for you.

Schedule a personalized demo today at zeropath.com.

Okay, back to IncludesNoDirt. But first of all, let's [00:02:00] talk a little bit about yourself. Bill, who are you and how did you end up doing this?

[00:02:05] **Bill Dougherty:** Well, my name is Bill Dougherty. I'm the CISO for a digital healthcare company called Omada Health based in San Francisco. Been here about eight and a half years. And Omada, we specialize in, cardiometabolic diseases. So type two diabetes, hypertension, things like that. We also have a physical therapy program, all of it delivered online through a smartphone app with cellular connected devices like glucometers and scales and blood pressure cuffs.

pretty cool company. prior to here, I was the CTO and CISO for a data center company for five years. I did a stint in site operations at stubhub. com, part of eBay, and then bounced around the Northern California, tech scene for more years than I'd care to admit,

[00:02:50] **G Mark Hardy:** it sounds like you've had some great opportunities and things like that from a career perspective. And I was just talking to somebody a little bit earlier today on LinkedIn who is interviewing for a VP [00:03:00] of security job. And he said, Hey, go back and listen to, I think it was episode 106 of CISO Tradecraft, which I recorded a couple of years ago.

About how to get your first CISO job, and I'm thinking, there's some good stuff in there, and sometimes you forget about things that you've written, and then you come back a couple years later and you go, whoa, not too bad. Other times you look at stuff and you go, I'm so sorry, my name's Annette. But you wrote this a little while ago, didn't you?

This includes no dirt. Can you tell me a little bit about the background here?

[00:03:27] **Bill Dougherty:** yeah, we, I think we published this in 2019, you know, time's been compressed the last few years, the genesis of it, Patrick Curry, who's the, our VP of compliance and I, we, we're doing our risk assessment at the end of 2018. And in healthcare, you're required to do an annual risk assessment. And actually nobody tells you what that means.

Nobody tells you how they just say, you know, there's a checkbox with your auditor that says, did you do an annual risk assessment? And at the end of 2018, we decided that we [00:04:00] wanted to start doing threat modeling. I've done a little bit of threat modeling in other places. I thought it was a good idea.

but we couldn't find a model that we liked. So, and there's some really good models out there, like a, STRIDE, which has been around forever. I think it came out of Microsoft. guy named Adam Shostack wrote the definitive book on it, and it's amazing, and if you want to get into threat modeling, buy Adam Shostack's book before you read my white paper.

It's much better.

[00:04:28] **G Mark Hardy:** And he's a great guy. And by the way, for people who have not heard of STRIDE, shame on you, but, if you remember your trivia, spoofing, tampering, repudiation, information, disclosure, denial of service and elevation of privilege.

[00:04:42] **Bill Dougherty:** That is correct.

[00:04:44] **G Mark Hardy:** computer security threats, basically technical type threats that you could go ahead and have a problem of, and that is one of the domains of information or things we've got to worry about as a CISO.

[00:04:56] **Bill Dougherty:** Right. So, STRIDE was really developed to, to [00:05:00] focus on application development,

and it really focused on security issues. In healthcare, I also have to worry about compliance issues and I have to worry about privacy issues. And sometimes, the security issue runs smack dab in conflict with a privacy or compliance issue.

Let me give you an example. we tend to, in the security world, want to authenticate and authorize everybody and we want non repudiation. We want to know who did what when. There are systems you sometimes have to put in place that require anonymous reporting. And so, knowing who did what when, Actually runs in conflict, and so when you're trying to do a risk assessment or a threat model, you have to know what the goals of the system are, and then you have to figure out, does the system matches goals?

So we sat down and said, How do we combine, security threat modeling using strata or base with some of the goals for compliance? [00:06:00] Some of the goals for privacy looked around. Found a privacy model called Linddun, which is also really very good for privacy things. And,

[00:06:10] **G Mark Hardy:** Trivia Time stands for. Yeah, linking, identifying, non repudiation, detecting, data disclosure, unawareness, and non compliance.

[00:06:21] **Bill Dougherty:** you, you've got that memorized far better than I do. I used to be able to rattle that stuff off. so we, we, we said, okay, well, let's see if we can combine these. then we, we took it a couple of steps further, and really had two big ideas.

first big idea was that it had to be broader than just application development.

So, so we started with, we want to model a system, and a system can be as broad or as narrow as you want to define it. So a system could be a new application, system could be a network. for us, we do, we're kind of a SaaS [00:07:00] first SaaS only type company. So a system is the vendor we're bringing on board for a new SaaS product.

Or we, we've used this for some M&A activities. And so a system could be the thing we're thinking about buying. So it can be as broad or as narrow, but we

[00:07:20] **G Mark Hardy:** pretty flexible then,

[00:07:21] **Bill Dougherty:** And, but we want this very repeatable process. The second was it had, if we're going to go very broad, as our definition of a system, and we want to model lots of systems, it has to be fast.

And so we sat down and said, well, how do we make this fast? And one of the ways we, we made it fast is we looked at the various goals of, of a system that you'd be looking at in a threat model and said, well, if we have controls in place that meet these and the new system adopts those controls, I don't need to spend a lot of time, modeling it again.

So as an example, we have a fairly robust authentication [00:08:00] mechanism. with multi factor, we adhere to all the rules of NIST 800 63B. it's very robust. If we're bringing out a new SaaS vendor, and it will integrate to that existing authentication mechanism for single sign on with SAML, then I can just check that box and move on with my life.

I don't need to go any deeper. In terms of a modeling standpoint. So, and what that meant was that we could put together checklists and say, tell me what the new thing is. And then for authentication, does it adhere to our existing standards? Yes or no. If yes, move on. If no, tell me more. And, and that meant that we could model a lot of things very, very quickly.

[00:08:44] **G Mark Hardy:** So it's almost like an object oriented modeling approach where you've gone ahead and said, Hey, we've already defined the

system. We know it's in there. We're satisfied that it meets our requirements. So if I drop this thing into a larger structure, I already know what that's doing. I don't [00:09:00] have to go ahead and do it a second, a third or a fourth time.

[00:09:02] **Bill Dougherty:** Yeah, absolutely. So the, the name of the models includes no dirt, which is, I take no credit from the name, Patrick, my coauthor, he took all of the principles that we got from STRIDE and from Linddun and other places and put them in an acronym generator and spit out includes no dirt ended up being kind of a cool name for us because.

If you think about it, we don't want any dirt in our system, dirt being a metaphor for, you know, risk or cruft or whatever. So ideally you'd want your system to be dirt free. so we, we call it the shorthand is the no dirt model.

[00:09:37] **G Mark Hardy:** Yeah, and there's 14 words in there. I'm not going to try to remember those because I don't. Although I did go through the list and I figured normally I find when somebody has a model and it fits to a word or pronunciation, I go yeah, somebody had to bend something out of shape a little bit to make it sound right.

And there's actually, it doesn't look all that bad, so well done. But if there's an acronym model that works, I need to go figure out what that is at the end of the [00:10:00] show. And maybe you'll find something that stands for CISO Tradecraft. Like,

[00:10:04] **Bill Dougherty:** we'll work on that for you.

[00:10:05] **G Mark Hardy:** all right, he sounds good. We got a deal. So how does this no dirt model that you have balanced both your technical and non technical requirements?

Because it sounds like this thing is capable of looking at more than just IT systems from a technical perspective.

[00:10:22] **Bill Dougherty:** to be clear, it is mostly focused on technical, not non technical. It's just technical across multiple domains. So technical across security and privacy. And we're, we're not really using it to model non technical systems, although presumably you could, but it, how it makes that balance is just including all the main, the, the main principles that we're trying to achieve and figuring out for a specific system, the person who's modeling needs to know what system does.

And they need to know, does this [00:11:00] apply or not? So, and the starting point for each of those, principles is essentially a series of yes or no questions. Like, do, does the system need to authenticate people yes or no? Pretty simple question. If yes, does the system use our existing authentication mechanism?

Yes or no? Again, and you can see how you could get through that really, really quickly on stuff that is just a repeat. the benefit of that is you, you're not reinventing the wheel every time. And for things that are kind of important and complex, you can get down to the meat of the thing very, very quickly.

[00:11:41] **G Mark Hardy:** And, that's very helpful. I have sitting on my desk, running things, a, and it's not really a threat model, but it is in a way from an insurance renewal application with all the questions of during the past three months, 12 months of this happened, has this been any changes? How many of this, how many of that?

And then of course, first thing you hope is I hope they [00:12:00] didn't change it too much from last year, because you can copy it over, but here, what we're looking at. It's to say that is looking at technical items and systems could be technical and non technical systems, but those areas and again, looking at the terms that you have in there, the includes no dirt, it is a technical thing, but is it really designed for someone who needs to be a network architect or an engineer to be able to use this or can a non technical person make good sense of this process?

[00:12:26] **Bill Dougherty:** certainly one of the goals was that a non technical person could read the output of it and understand it. Okay, that isn't to say that I would expect a non technical person to be able to do all the analysis.

[00:12:41] **G Mark Hardy:** yeah,

[00:12:43] **Bill Dougherty:** you have to be fairly, versed in the system you're modeling in order to model it, no matter what that system is.

[00:12:50] **G Mark Hardy:** that would make sense.

[00:12:51] **Bill Dougherty:** the output should be human readable. For the most part, I have my security engineers doing the [00:13:00] threat modeling, but they can go model an application that my engineers are doing without having to be experts in writing the code, or they can go model a network without having to be network engineers.

What they need to know is what are the important goals and the principles, and then ask questions like, how are you addressing this?

[00:13:20] **G Mark Hardy:** Okay. Got it. And so what I find also interesting about your model is that you've created, on the cover of your paper, you got a little Venn diagram looking at security, privacy and compliance. And most of us tend to focus on security models and things such as that. And we did good at that. And then, and we cover that because that's really what we focus on.

And privacy, of course, is a big deal. Sometimes we say, there is an entire privacy department or the lawyers take care of that. And we have a tangential. Of course. Both you and I know security and privacy are not the same thing, but then of course, add in there the compliance. And so does that get driven by different frameworks for the compliance requirement?

And as such, do you have [00:14:00] questionnaires for each one or does someone say, Hey, I've got to deal with PCI DSS, so I've got to go to the drawing board and come up with the right questions or is there a repository of things to help out with the major?

[00:14:11] **Bill Dougherty:** I would say that the most accurate way to think about this is I'm ultimately modeling against my own control framework and my control framework is guiding guided by what I have to comply with. So I'm a healthcare company. I have to comply with HIPAA. I also have to comply with my customer contracts and a variety of other obligations.

we are HITRUST certified, so we have to comply with HITRUST. so, and every company has something to have to comply with, right? So it's not that there is a separate questionnaire, it's that, for the most part, the things I have to comply with are baked into these principles that we are dealing with.

So, good example, HIPAA security rule requires, protected health information to be encrypted at rest and in [00:15:00] transit.

[00:15:01] **G Mark Hardy:** Okay.

[00:15:02] **Bill Dougherty:** so within, right, right, but, so within the system, it's going to ask, Is the data being encrypted at rest and in transit,

sometimes that answer is yes, and sometimes that answer is no. Earlier on, it's what kind of data are we modeling? Because not all of the systems in my company contain health information. I have employee systems, I have email, I

have all kinds of other systems as well. So, tell me what your system does, tell me what kind of data that's in there. And, but then, we're always looking at things like baseline controls.

We want all of our data encrypted at rest and in transit. So if, if you tell me, yes, again, I likely get to move on. If you tell me no, well, now I need to go dig deeper.

depending on how built into the model was a scoring mechanism. So if you, if encryption is important to me and you say it's [00:16:00] encrypted, then I don't add any points to the risk.

If you tell me it's not, I do add points and that. Increases the score and the higher score is worse from a risk perspective.

[00:16:11] **G Mark Hardy:** It was like golf. Okay.

[00:16:13] **Bill Dougherty:** So when one of my people gives me a completed threat model, the very first thing I do is I look at the score. And if it scores super low, it's like, okay, this probably isn't something that needs a whole lot of my time.

As the CISO, if it's scored super high, I probably want to pay more attention to it. So you can imagine that if you have a repeatable process with a reliable scoring mechanism, you can use then allow all of the leaders, the manager and the VP and the CISO to triage things.

[00:16:44] **G Mark Hardy:** And that sounds like a real time saver and a real way to go ahead and focus effort. And it's interesting when you think about this, is this in IT tool? Is it a IT security tool? Is it a GRC tool, or is the answer yes.

[00:16:59] **Bill Dougherty:** Actually [00:17:00] it's yes. In fact, if you read in the paper, you'll see that we instrumented the, the questionnaire in the GRC tool or that the scores are always captured.

[00:17:11] **G Mark Hardy:** Got it.

[00:17:12] **Bill Dougherty:** which means like for us, we've got hundreds of vendors and every single one of those vendors has a risk score based on a threat model, based on what kind of data we're putting in there and what kind of contract we have and where the data is and what the thing is supposed to do.

And so I can tell you my risky vendors from my non risky vendors for every single vendor that we share data with.

[00:17:31] **G Mark Hardy:** And that was what I was going to get to next was in terms of who's the subject of this is that you're saying, I'm first I'm thinking, okay, I look at my own environment because typically you look at your own risk and things like that. But what I'm hearing from you is that this is a great way to go ahead and potentially do risk modeling of third parties,

[00:17:48] **Bill Dougherty:** it is for us, and again, I, I said at the outset that we are a, we're like a SaaS first, SaaS only type company being a 14 year old company. So with the [00:18:00] exception of the systems we write for our own product. All of our corporate systems are SaaS. So, if I'm not good at third party risk management, then I'm not good at IT or security at all, because there are lots of companies that have some portion of my data.

[00:18:17] **G Mark Hardy:** Touche. So that makes good sense. And that's not always true for every organization. It's nice to be able to have that standardized approach because I'm on the other end of it, where I get all these questionnaires from either third parties or insurance companies or whatever. It's can't you guys just ask me the same question every time?

And then I'll go ahead and say, here, copy, paste, and off we go. so

[00:18:38] **Bill Dougherty:** we have to send out questionnaires to all of our vendors. We also get questionnaires from all of our customers. So, security questionnaires are the bane of my existence. I hate them. I could spend the next hour ranting on a soapbox about how stupid and pointless security questionnaires are.

[00:18:55] **G Mark Hardy:** too, bad you can't say to whatever customer it said, sent you [00:19:00] something, said, wait a minute. We got something even better than what you asked. It's our model. And here's our self scoring, by the way, or even if you did it externally, would this be done in this case, like self scoring, would you actually have your third party auditor come and said, we're going to use your own checklist for you.

So you can then go ahead and present that to all of your customers and get out of the word business of filling out these stupid forms.

[00:19:20] **Bill Dougherty:** that, that's what a HITRUST are supposed to be for. we, we go through all of the, the pain and expense of, a third party risk assessment and then we don't rely on it.

[00:19:34] **G Mark Hardy:** I work with organizations that are a little bit too small to make a cost justification for a SOC 2.

[00:19:39] **Bill Dougherty:** also, built into the, the white paper and includes no dirt is a stripped down security questionnaire that we also use. I think we've added some questions since we originally published this. The basic philosophy was don't ask a question unless if you didn't like the answer, you would refuse to do business with the vendor other than some [00:20:00] very basic things like tell me what your system does, what type of data you have and give me a system, you know, a data diagram.

But otherwise, if I if I ask you 10 questions about your password policy. And I didn't like any of your answers, and I'm still willing to do business with you. Then why did I ask you about

[00:20:17] **G Mark Hardy:** Might have asked both questions in the first place. Yeah, I'm looking at some of these samples that you have in the appendices and things such as that. And it's, you've got some scoring in there. If it's, if you say yes, get this, if no, then go on to that or required, not required, et cetera. So it really is more, almost a methodology in a way as well as a scoring system.

You're not just left to guess, like someone says, Hey, are NIST SP 800-171 Okay, that's an awful lot of controls to go through, but maybe some of those things aren't going to be critical to your relationship that you have with your SaaS provider. Just don't even bother asking them.

[00:20:52] **Bill Dougherty:** right. How does, how does NIST 800 171 apply to this particular system?

[00:20:58] **G Mark Hardy:** Yeah. And if it does [00:21:00] now, here's the interesting Don Rumsfeld question. What about the unknown unknowns, the questions you don't ask, and therefore you don't know the answers to?

[00:21:10] **Bill Dougherty:** there will always be unknown unknowns. The, the best way to uncover them, I think, is to spend more money, more time and effort on what the system is and what it's supposed to do. And let's climb on things that you just know it complies. So we spent a lot more time focusing on tell me exactly the data that's going in and exactly the data that's coming out and where

is it going to be stored and, and what is unique about this system, not and less time on password policies or encryption or other things because we can just check those boxes.

So I can evaluate somebody without sending them a 2000 questions SIG questionnaire. Of which nobody reads the answers anyway. You just feed it into your GRC or your [00:22:00] GRC kicks out a score and then you do something and more time on, show me a data flow diagram. And we, because we use this so often in our third party risk assessment and we do it during the procurement process, before we buy the thing, what happens is one of the results of this process is a list of things that we either want our business partner to implement.

Or we want our legal team to negotiate in the terms and conditions.

so as an example, we have a policy as well as some obligations to keep our data within the United States. So a pretty common question is where, where are your data centers located? And will you, will you attest to storing, processing and accessing only from the U.

S.? And sometimes the answer is yes, sometimes the answer is no, sometimes the answer is well, we've got 10 data centers around the world and, and so we'll go back to legal and go, okay, legal, [00:23:00] we want a data locality term in the contract, and then we'll go fight with the vendor if we have to on that, but it kind of exposes that up front before we ever signed the contract, it's a lot easier to deal with that up front.

[00:23:14] **G Mark Hardy:** I like that. No.

[00:23:15] **Bill Dougherty:** same thing, like I, I had a SaaS request come to me yesterday and I asked, well, how do they do authentication? And the authentication was OAuth. I said, well, do they have a SAML option, do they? And they said, yes, but it's more money. It's a wonderful. single sign on tax and said, go get me SAML and tell them we're not paying extra for it.

And let's make that part of the negotiation. You want a deal? Okay. We, yes, I understand that's your higher tier. I don't care. I, I want SAML so that I can automate the provisioning and deprovisioning through my authentication mechanism. And so by having the right set of questions up front, regardless of who's on my team is evaluating the vendor, they're following a standard process.

We know what that process is. And we can [00:24:00] surface these things quickly,

[00:24:01] **G Mark Hardy:** And it seems also good from a negotiation perspective, as you know the stuff that you have to have and know your throwaways,

[00:24:07] **Bill Dougherty:** right?

[00:24:08] **G Mark Hardy:** and that's always important to go ahead and get those, figured out in advance. Now, if I'm a healthcare organization like you are, and you're subject to HIPAA or HITRUST, then what you say is, if I go through, this includes no dirt model, does that mean that I'm now Know that the third party is HIPAA compliant or is that an entirely separate exercise?

Are these orthogonal to each other? what's the value of doing all these things?

[00:24:36] **Bill Dougherty:** So the, the answer is, I'd say it's orthogonal, because the first, the starting question is, does that third party access health information?

And if it is, then we have to pay more attention to certain rules and just the fact that they would have PHI in it, raises the risk score for [00:25:00] us.

[00:25:00] **G Mark Hardy:** Okay.

[00:25:01] **Bill Dougherty:** If the, if the first question is, are you going to have health information answers?

Yes. That's an immediately signal to my legal team. Oh, we're going to need a business associates agreement. Okay. we're going to be less willing to budge on certain things like data locality. I can budge on data locality for data. I don't care about, I can't for health information. I can, I can budge on password rules for systems I don't care about.

I can't, if it has health information. So it. It gives us, we're asking a set of questions based on kind of our overall policies that is telling us which systems are higher or lower risk. And that then guides what risks we're willing to accept and what risks we're not willing to accept. So again, yes, I got, it was funny.

I got two SaaS requests yesterday, both of which came to me and said they needed, they had OAuth. One was like for five users in a training system that I [00:26:00] really don't care about, it has none of my data. And I said, fine, I'll ask, fine, I don't care. The other one's going to have a lot of information. And then, okay, great.

This one I care about being able, asking the right questions up front, let's you triage it and go, this is worth my time. This is not worth my time.

[00:26:15] **G Mark Hardy:** So we've got these existing models that are out there like the STRIDE and the Linddun and things such as that, that will help us with security modeling and privacy modeling. but you've did this, you said the first draft was over five years old. And are you still in version 1. 0 is because most standards never stay 1.0 except maybe something like a Bitcoin. We still keep using that.

[00:26:39] **Bill Dougherty:** I don't know that we've really versioned it. Like when we first did this, the reason we published it is because we were young in the industry. And, and at the time, like. Digital healthcare was still kind of a nascent thing and no one was telling us how to do it, which means no one was telling anybody else in our field, how to do it and healthcare is the most breached industry there is.[00:27:00]

And we kind of felt an obligation to put something out, to help all of our peers that were at digital healthcare startups, or getting lots of money trying to figure out how to do this stuff. we have tweaked the model internally. We haven't republished it. I don't know if we will, cause there's a certain amount of effort they have to go through just to get.

You know, legal approval to, to, put something out with your logo on it. but we've, we haven't changed the principles. We have modified some of the questions, some of the scoring we've changed a little bit, the depth will go into certain areas. we've gone from what started off as basically like a questionnaire form that you would fill out to more like a Google docs.

So it's a little more robust. So we've, we've changed it certainly in the five or six years we've been doing it. I'm working right now to look at new emerging [00:28:00] risks relative to AI. And what I'm trying to figure out is, does it actually fit in the principles of includes no dirt, or are there new principles? At the moment I think they fit, however, there are new threats and there are new attack vectors, but the principles remain the same. And so, of the stuff related to AI is so unique compared to what we've been doing for the past 30 years. That I'm considering revving this document to actually include a full set section in there on how to think about AI.

[00:28:38] **G Mark Hardy:** so we'll put a, at the end, it'll be in includes no dirt. Hey, it'll be your Canadian model.

[00:28:43] **Bill Dougherty:** It'll become Canadian.

[00:28:47] **G Mark Hardy:** But, yeah, because that was one of the things I was thinking about with the AI is that things are changing very quickly. And as a result, what we think the threats are of six months ago may not be the ones we have today might be different six months from now. So [00:29:00] it requires some fluidity in terms of our ability to manage and measure that.

But the idea of doing threat modeling. If organizations are not doing threat modeling, it seems that with AI, you really just, if there's anything to get you off of top dead center and get you moving toward it, it might be that. Ah.

[00:29:19] **Bill Dougherty:** yes, because it is requiring, I would say, some new ways of thinking. And I'll give you an example. I'm willing to bet that if I were to ask any of your security practitioners who are watching this podcast, do you do software development on production data? Every single one of them would say, Oh, God, no.

Like, we never allowed devs to write code against production data. Like, we, we have fake data, we have a way to sanitize it, we have some process in place so that they can go play in a dev sandbox, do the job they need to do. They'll move it to staging, they'll move it to production, [00:30:00] but you don't, like, basic, you know, security 101.

You don't write code against production. Okay, what is model training? Is that software development or is that production? You can't train models effectively unless you're using real data, okay? So now we have to figure out What does that look like in our world is because it's really not, if I'm doing pure model training or, or, or if I'm doing prompt engineering, these are in one way more akin to software development, but I, and I certainly haven't pushed them out to production yet. But they're not software development the way we've done it traditionally, there's something new, there's something different as it needs to have all the security controls of production, but it also needs to have the [00:31:00] fluidity and off lineness of development.

[00:31:05] **G Mark Hardy:** That's a very interesting insight and profound in a way, because as you say, that makes perfect sense, is that as you build out your AI models, you have to train on something and that training technically sounds like dev to me. But if you train on useless information, you'll never get anything useful on the other end.

So you don't really know if it's going to work.

[00:31:24] **Bill Dougherty:** So, so now let's take this a step further, because I'm in health care, health care, we have a principle called minimum necessary. Which means that when you're dealing with protected health information, you use the minimum amount of data and you expose it to the minimum number of people to accomplish the, the purpose. And the purpose has to be an allowable purpose, such as providing care or, or doing billing or claims or, or running your healthcare operation. Okay. And certainly in healthcare, there are totally [00:32:00] legitimate purposes to use health information to train them all. The question is, when you're training a model, what is the minimum necessary amount of data?

[00:32:11] **G Mark Hardy:** And using training models, you want as much data as you can get in there.

[00:32:14] **Bill Dougherty:** You do and you don't. Because there's also a cost for every additional data element. And, it depends on what the purpose of the model is.

like I, I, we started this talking about systems and what is the purpose of the system, when you start thinking about training a model, what is the purpose of the model? So, it, depending on the model I'm training, I might need identifiers like a name, or I might not.

if you think about in my world where we have lots of different types of data, we have biometric data coming from devices and we have interactions with the care team, and we have lots of data. If I'm doing something that really only involves biometrics, then that's the only [00:33:00] data I need. I don't necessarily need the other stuff. So, we're trying to adapt how we do things, and based on how Applying a set of 25 year old rules under HIPAA and high tech to this new thing. Now, how do I adjust my threat model and make this analysis repeatable? Because we're not going to stop. We're just going to keep doing it.

[00:33:26] **G Mark Hardy:** And so for. The environment you describe, HIPAA, HITRUST, things like that, I get that. What if I'm not in healthcare at all? Do I just say, Hey, this is way too specific for healthcare, or does this have general applicability?

[00:33:42] **Bill Dougherty:** as someone who came out of e commerce and has, worked at a lot of different companies, I think it's got wide applicability because the same basic principles apply across the board. if you accept credit card payments. You're supposed to under PCI protect cardholder information.[00:34:00]

You shouldn't train your model on cardholder information unless you need cardholder information for your model. It's the same basic concept, and, and you certainly have always had this concept of masking. card numbers, if you're in financial services, you've always had this concept of masking the social security number.

You don't, you give that number to just anybody. So the same basic principles apply, and we also have privacy rules that apply to everybody that's not in healthcare. So like CCPA has a carve out if you're in healthcare, but if you're not in healthcare, CCPA applies. And so you still have a compliance obligation.

You still have a set of rules you have to adhere to. The point for us was to put out a baseline model and then make it as easy as possible for people to cut and paste and then modify it to their own purpose. So for, to me, Includes no dirt is an, is a choose your own adventure kind of thing. Like I don't expect that we got it right for everybody.

We [00:35:00] wanted everyone thinking in a certain way and then go make it yours and instrument it for whatever works inside your organization, because we didn't get it right for everybody. We barely got it right for ourselves.

[00:35:12] **G Mark Hardy:** the thing is, it's a great progress in something. So I think for the, sake of at least getting into the recording, let me go ahead and get the include no dirt terms, even though we're not going to dive in each one of them with the time remaining, but identifiability, which is really an anonymity, it's a privacy issue, right?

the non repudiation, creating, if you will, the avoidability of the plausible deniability. I didn't do that. yes, you did. You've got non repudiation there. That's also a privacy issue. A clinical error, which is the only one that I saw that really labeled you as healthcare, that's your aha, you flipped over your card.

The correct application of clinical standards, that's a compliance and one could argue that depending on what industry you're on, it could be the banking standards or it could be some other manufacturing standards, et cetera. But it didn't make the acronym spell all [00:36:00] so I'll defend you on that one, linkability, versus unlinkability, what are we referring to here?

The ability to link certain data together to come up with a conclusion, like who's the record really is this or,

[00:36:13] **Bill Dougherty:** Exactly. It's the ability to link disparate data sets together and it's directly related to, the identifiability or non identifiability of PHI. Correct.

[00:36:24] **G Mark Hardy:** it. Okay. Unlicensed activity. Looking at on a compliance perspective, there's a proper credentials, a licensure, a denial of service. Absolutely, an available issue of the CIA. Now we're into security, privilege elevation, spoofing. These are good security stuff. Non compliant to policy or obligations.

We're back to compliance here. And of course, non compliance is a big deal. With regard to assessing and really, if you're looking at a model that says, Hey, I've got a HIPAA requirement, I could collapse a lot of that into there to say, Am I compliant? [00:37:00] the next one I thought was rather interesting overuse, and that tends to back to what you said about minimum necessary. That you want to make sure from a compliance perspective that you're not putting too much in there because now you're disclosing information perhaps on a patient that should not be in that record. Data error is an integrity issue from the CIA. Information disclosure, C for confidentiality. repudiation, you had non repudiation before with a concern about plausible deniability.

And now you got it. Repudiation, with the non repudiation, the first was privacy, the second was security. Was that just to fit the model to get the last letter in or?

[00:37:43] **Bill Dougherty:** No, it really,

[00:37:45] **G Mark Hardy:** things,

[00:37:45] **Bill Dougherty:** it really goes back to the, to the tension between privacy and security. And when do you want, there are times when you want plausible deniability and there's times when you want non repudiation, they're different goals.

[00:37:59] **G Mark Hardy:** Exactly. [00:38:00] And the last one, the T for the no dirt, tampering. Again, integrity, being able to protect from that. So data error, meaning that information was changed, probably inadvertently or whatever, in the tampering. Somebody actually went in there and tried to modify stuff. And these are pretty good ways of looking, very good ways really, of comprehensively looking at threats and risks that might be out there.

And from a threat perspective, you could then go ahead and say, Hey, how do we model this? How do we protect ourselves against this particular threat? If we're worried about spoofing, for example, then we say, Hey, my goal is authentication. And then I could dive into the model, I could read a little bit more.

And in your paper, which I'll include the link. In the show notes, but it's basically http://www.includesnodirt.com/nodirt.Pdf 37 page pdf and scrolling as we talk. So no, I haven't memorized all these things, but it does nice examples at the end. how you would fill these things out, the point scoring and things.

And as you indicated, entities could [00:39:00] modify these to better meet your model score. And I see you end up with a low, medium, or high. It's a little bit like going to a Zagat restaurant. Oh, this is a 22. It must be pretty good. Or it's a 15. you get food poisoning or however they do that.

[00:39:13] **Bill Dougherty:** So the other thing I'd point out because there was a lot that didn't fit into the, the includes no dirt risks. There's also in the paper, a miscellaneous category, which is basically our generic catchall for everything else that depending on the thing you are modeling, you might put in there. So physical risks, environmental risks and criminal risks and disaster risks and regulatory, there's a lot of stuff in there, competitive risks.

It's really, what is it that you want to pay attention to?

so like disaster risks. If I was, if I was standing up a new factory, I want to know all the environmental risks of where it's going to go. Is it an earthquake zone, a flood zone? Is it subject [00:40:00] to hurricanes and tornadoes? But and we would account for that if that was the thing we were modeling.

But if the thing we're modeling is A new SaaS product where we may or may not be going in that deep.

[00:40:13] **G Mark Hardy:** Got it. Understand. Hey, is there any thoughts that you can think of as we wrap up here in the show that you'd like to leave our watchers or listeners with or recommendations for how they could go ahead and better incorporate threat modeling into their business practice?

[00:40:27] **Bill Dougherty:** Yeah. A couple of things. we've made this paper free download. There's no copyright on it. There's no license on it. You're free to use it. We want you to use it. We love hearing from people who have been using

it. I heard from a buddy of mine who took this and turn it into an LLM. So he's got a chat bot that asks

[00:40:44] **G Mark Hardy:** Oh, wow.

[00:40:47] **Bill Dougherty:** one of the things that was an, an initial stumbling block, and it was such an important thing for us that we. Spent a lot of time on it. And it's right at the beginning of the paper on page four is we created a taxonomy and [00:41:00] because what Patrick and I realized was that the compliance people and the security people were using the same words to mean different things. And so we published a taxonomy. It's ours. I'm not saying it's right for everybody, but you should have a taxonomy in your organization so that. Everybody who's using this has an understanding of this, what you mean, what is a risk, and what is a threat. like, if I ask some people what a threat is, they're going to identify a system, or they're going to identify a vulnerability.

We define a threat as an actor, someone who acts. it could be a system, it could be a piece of code, or it could be a person, but you can think of like insider risk, insider or threats. What they do is not a threat, it's an attack vector. And so, we found it very, very important to document a taxonomy and make sure everybody knew exactly what we're saying [00:42:00] because a threat is not the same thing as a vulnerability.

And a vulnerability is not the same thing as a risk to us. But oftentimes we use these terms interchangeable when we're, when we're kind of speaking off the cuff. but if you're going to get through like a repeatable scalable modeling process, you better use the same term every time.

[00:42:20] **G Mark Hardy:** Got it. And I've got them right here. And I, agree with you. We had this on the pre show. We're talking about a threat and I agree it's an actor or principle. It's something, it's a source of where that danger is coming from rather than the attack vector, which is the method. And then, of course, a vulnerability is something that exposes your asset to that particular threat via that attack vector.

And now that nomenclature kind of loops in, he says, I get it. And by agreeing on that common language, that common taxonomy, if you will, it's a little bit easier, a lot easier, in fact, sometimes to communicate, even with executives and managers who seem to think things are different, but if you can lock them down to this, it makes good sense.[00:43:00]

this has been a really valuable episode. I have learned a lot. As I say, I was able to print this thing out because I had to read it on the airplane flying back and forth. And I think that's great work. So if you do update this, that would be wonderful. If not, I think what's out there right now, the current version is certainly good.

So for everybody who wants to take a look at it, again, the link is in the show notes. I already mentioned where to find it on the website. last thoughts or should we wrap up?

[00:43:28] **Bill Dougherty:** really just pleasure being here and I appreciate you, you know, even after this is five or six years out that you found it valuable and you're highlighting it again because I still think there's a lot of people out there that could benefit from this. and I'll say what I started at the beginning.

If you really want to go deep diving into threat modeling. I think it's called Practical Threat Modeling by Adam Shostack. It's a great book. the, the good ideas in this paper, for me, I stole from him.

[00:43:58] **G Mark Hardy:** we'll give Adam a shout out [00:44:00] there. I have gentlemen I've known for a number of years, and I think he's a great security practitioner and a great guy from the time I've spent with him. so thank you very much. So we've been talking about includes no dirt, a practical approach to threat modeling for digital healthcare and beyond with Bill Dougherty so thank you so much for being part of our show for our listeners.

If you like CISO Tradecraft, we're a lot more than podcasts. If you're not following us already, we'll do If you're listening to us on your favorite podcast channel, give us. Five stars if you can, because that helps boost our ratings and then other people will find us if you're on YouTube, appreciate it.

But also, we have a Substack newsletter and we have other ways we put out information on LinkedIn. We'll go ahead and have a steady stream of high signal, low noise, useful information to help you along with your career. So we hope you've enjoyed this episode very much. If so, drop us a note and let us know, and let Bill know, perhaps, that, hey, we love this, we'd like to see more, and who knows, maybe you'll get encouraged to write the next version of the model.

So until next time, this is your host, G. Mark Hardy. Thank you for listening to CISO Tradecraft, and stay safe out there.