[THEME MUSIC]

SPEAKER 1: Welcome to season five of the Prodcast, Google's podcast about site reliability, engineering, and production software. This season, we are continuing our theme of friends and trends. It's all about what's coming up in the SRE space, from New technology to modernizing processes. And of course, the most important part is the friends we made along the way. So happy listening and may all your incidents be novel.

SPEAKER 2: You missed a page from Telebot.

STEVE MCGHEE: Welcome back, everyone, to the Prodcast. This is Google's podcast and I guess video thing, videocast, about SRE and production engineering. And we're live in New York, even. Amazing.

FLORIAN RATHGEBER: For the first time, we're recording in person.

STEVE MCGHEE: That's right. It's pretty cool. I'm Steve. I've been here a bunch. Florian you know as well. Florian's been here a bunch. We have a new friend. Who are you, new friend?

CARLA GEISSER: My name is Carla Geisser. I worked at Google as a site reliability engineer from 2004 until 2015-ish, mostly on storage systems. There's a quote from me in the SRE book. And now I run a company that does what we are calling crisis engineering. The company is called Layer Aleph.

STEVE MCGHEE: Cool. What is a crisis and how do you engineer it? What's going on with that?

CARLA GEISSER: So we made up this term because we needed a thing to call what we had done for our careers, including at Google, which was enter a situation full of ambiguity, full of circumstances that were surprising to people, and perhaps information that was noisy and use a combination of management type skills and software engineering and systems engineering skills to find a way out of the maze. And we needed a word for that, because that took me 30 seconds to say. And so we described it as crisis engineering.

FLORIAN RATHGEBER: That's interesting, because what you just said to me somehow also seems to apply to a regular, quote unquote, incident. So what makes a crisis then different from an incident as an SRE might know it?

CARLA GEISSER: The place that I draw the boundary is that an incident is a thing where there is probably some set of instructions you can pull off the shelf that is maybe a similar thing that has happened before, or even a full on playbook, because it happens often enough. And you can turn to page 37 in the manual and say, now I'm going to follow these steps, obviously with modifications, because it's not going to be quite the same.
But there is a playbook that exists typically for a thing that is an incident but not yet a crisis.

FLORIAN RATHGEBER: Cool. So an incident is something you either have a playbook or you could reasonably have a playbook. And for a crisis, you couldn't.

CARLA GEISSER: Yes. And we have a taxonomy for the elements that we think are necessary to be in a situation that is a crisis. There are five of them. The first one is fundamental surprise, which is, this is a thing that is a novel experience for everyone, and it has to be truly novel. So the place that this term kind of gets used most is actually in science fiction writing, because it's a great setup for a science fiction novel. Your cat wakes you up one morning, that part's normal, and tells you that the Earth is about to be invaded and invites you onto their spaceship to evacuate.

STEVE MCGHEE: That's novel.

CARLA GEISSER: That is fundamental surprise.

STEVE MCGHEE: Got it.

CARLA GEISSER: So, for example, during the Amazon outage, several people's smart

mattresses failed. That's novel.

STEVE MCGHEE: That's novel.

CARLA GEISSER: That is a new thing that you did not expect the computers way out in the cloud to have control over.

FLORIAN RATHGEBER: You could argue whether they should.

CARLA GEISSER: Well, yes.

FLORIAN RATHGEBER: Topic for a separate episode.

CARLA GEISSER: So fundamental surprises, one. Broken critical functions is another one. That one's the one we typically have in any crisis and in most incidents. Something that is core to your business is not happening right now. That one's easy. The other one is high visibility. So people are watching the outcome of this thing. And that can be either an internal thing or an external thing, particularly inside a large organization.

The thing you might start to experience is that every meeting becomes about this incident, even if it originally was not meant to be about this incident. If the news media is involved, obviously that is definite high visibility. So that is another criteria. A rigid deadline, like a real, actual deadline. Holiday shopping season, streaming for the Super Bowl, launching a satellite. Those are real deadlines. Internal tech company deadlines don't count.

STEVE MCGHEE: They don't count. OK, good to know.

CARLA GEISSER: Yeah, those are fake and made up, and they move all the time, as we know. So rigid timeline. And then the final one is hard to talk about. It's what we call perception breakdown or a failure of sense making. This is when the people are unable to produce a useful picture of what is going on, because either the information they're getting is late, or it's noisy, or it's overwhelming, or different parts of the organization are seeing different things happening. The simplest version is it works on my machine. That's the beginning of perception breakdown. And then we go from there.

FLORIAN RATHGEBER: It feels a bit like headless chicken syndrome.

CARLA GEISSER: Yes.

STEVE MCGHEE: Yeah, and it feels fundamental surprise. And the last one seemed quite related as well. In an organization who maybe thinks they understand production but actually doesn't. Is that a common thing that you've seen?

CARLA GEISSER: Yes. In particular in-- we increasingly like to say that the core of modern civilization now has computers in it, whether we like it or not, but a lot of organizations have not yet admitted that to themselves. And so you get a lot of organizations, especially in government, where they're processing health care benefits or unemployment claims, and they haven't yet gotten to the place where they understand that the computer is actually controlling all the decisions. They still think there is a policy working group that can control what is happening. And it can, but first, you have to tell the computer what the policy working group decided.

FLORIAN RATHGEBER: Is it a matter of them not admitting to themselves, or could they actually be genuinely unaware?

CARLA GEISSER: It's often a little of both.

STEVE MCGHEE: I know that we've seen this-- so we work in more of the cloud space, less of the government space, of course, but a thing that we see a lot is the people in charge are the people who used to be doing the job. And so often, they'll assume that the system is clearly just the same as it used to be, only bigger or something, whereas fundamentally it's far more complex or something fundamentally is different about it. And so they'll attack the problem as if it's the old one. But in reality, it's something totally different. Is that a similar kind of vibe?

CARLA GEISSER: Yes, and one of the kinds of perception breakdown you see is information that is out of date or stale, and that is a version of that where the people who think they have

their hands on the controls are holding the controller from five iterations ago. And so it's not actually connected to anything that drives anymore.

STEVE MCGHEE: Yeah. The example that we use a lot is if you have a modern distributed system, but you're thinking about sort of a static stack of things, your responses to failure are very different. And also your demands of your expectations of this thing can't be expected to fail versus a distributed system, which is like, it's made to fail. It's OK to fail. That's the whole point of it. You're just looking through the wrong lens entirely. So that seems really common.

CARLA GEISSER: Yeah, a version of that that we see a lot in our work with government is what volume of data is considered too hard for a computer to deal with? And people often have a extremely low estimate of that, because they're used to what will fit on a spreadsheet on their laptop, when in fact, reprocessing 10 million unemployment records might actually be very possible for the mainframe. You can just do that, but it's not inside people's belief of a way you could deal with the system.

FLORIAN RATHGEBER: Back in the olden days, when the term big data was sort of becoming popular, I seem to remember there was this story of the three V's, wasn't it? Volume, velocity, voracity, something like that. Does that ring a bell?

CARLA GEISSER: It sounds like a real thing. I'll believe you.

STEVE MCGHEE: So let's back up a little bit. So this all came out of experiences that you all have had. So this is Layer Aleph. Did I say that right? And that came out of your work with what? Take it from there. Where did this come from?

CARLA GEISSER: So it came from our work initially on the healthcare.gov crisis, which at this point probably nobody remembers.

STEVE MCGHEE: I remember.

CARLA GEISSER: You remember it. When the federal government launched a website to allow people to sign up for health care, and it did not work. And then Mikey built a team to help resolve that, which was a combination of managing contractors and then also debugging this hilariously complex distributed system.

STEVE MCGHEE: And there are videos of talks that came later which describe that whole thing. So we can link to those so people can review that, because it's part of the record. It's a good part of the story for sure.

FLORIAN RATHGEBER: Yeah, and even as a non-US citizen, I have memories of that. Or I have a-- yeah, I have heard of this.

STEVE MCGHEE: The best part of that, I recall, is the monitoring system was CNN. I think that kind of tells the story of the first day. I think that was--

FLORIAN RATHGEBER: Enough said.

STEVE MCGHEE: Enough said. Yeah. So then what? Clearly time has passed since then.

CARLA GEISSER: Yes, time has passed. After that, we did a series of similar projects inside the federal government and then also with various state governments. And somewhere along the way, we formed this company, which I think is about seven years old now. And we now sell this sort of technical crisis management incident response thing to anyone who wants to use it, which typically is often state governments.

It's sometimes private companies who are encountering what we would call a planned crisis. The version of this that we did that I had a lot of fun on was a company, a very small startup that was about to be acquired, and they needed their systems to not fall over during the finalization of that acquisition contract. But they knew that there would be a big PR around it, and they knew that--

STEVE MCGHEE: The date. That's the number three, I think.

CARLA GEISSER: Exactly. There was a date and there was a clear consequence.

FLORIAN RATHGEBER: Heightened scrutiny.

CARLA GEISSER: Heightened scrutiny.

FLORIAN RATHGEBER: Attention.

STEVE MCGHEE: Visibility.

FLORIAN RATHGEBER: High visibility.

CARLA GEISSER: Yes, definitely high visibility. And because they had been operating on a very small budget for a very long time, the set of people who actually knew how the systems worked was very small. And so there wasn't really anyone who knew whether it could scale, how it could scale, whether it would survive this upcoming increase in attention. And so we worked on that with them.

STEVE MCGHEE: Maybe this is a silly question, but are the things that you end up fixing always technical or never technical? I know the three of the four of you in the company, I've met at least three of you, and you're all former SREs. I don't know if the fourth one maybe also is too. But is it an SRE thing by nature, or is this just a coincidence that you're here?

CARLA GEISSER: I think the part of it that is SRE by nature is the ability to look at a very big system and start to connect the dots with no information. And the system includes the people often. So a lot of the work we do is--

FLORIAN RATHGEBER: Does it ever not?

CARLA GEISSER: It almost always includes--

STEVE MCGHEE: That was a leading question. Of course it includes--

CARLA GEISSER: It always includes the people. So a lot of the work we end up doing is technical, in that we are helping with visibility of the technical systems that exist. So things like building database queries that tell you actually how fast is data flowing through the system, and where are particular kinds of records getting stuck in this pipeline, which is often something-- there's a breakdown of perception.

Different people in an organization will look at a complicated process, like applying for unemployment benefits, and they'll say, well, clearly it's the people over there who are blocking all the applications. And you can just look in the database and see where stuff is getting stuck, but nobody has.

STEVE MCGHEE: So having the technical capabilities helps you uncover some of these problems, which may not even be technical to begin with, at least. And then maybe it's a coupling of other things down there as well.

CARLA GEISSER: So yeah, a lot of the technical work we do is about exposing the full system to the people who can then make better decisions, I would say.

STEVE MCGHEE: Right. It's all just software, as we like to say.

CARLA GEISSER: Yes, it's all just software.

STEVE MCGHEE: That's all it is. Come on. It's not that hard.

FLORIAN RATHGEBER: And presumably, this may be the first time that these people see that system end to end through this lens.

CARLA GEISSER: Almost always. That is actually the value we are providing as part of the thing we call crisis engineering is building them a very clear map of what is going on right now from the beginning to the end of whatever their process is.

STEVE MCGHEE: I like to say that when I left Google, I left Google and came back at a point, and that was when I left Google was-- well, I worked basically from university to Google to somewhere else and then back again. And during that somewhere else is where I learned the phrase stay in your lane. I had never heard that before. And my response was, please do not tell an SRE to stay in their lane. This doesn't work in our brains. This sounds kind of similar to that kind of concept.

FLORIAN RATHGEBER: SREs are looking left and right.

STEVE MCGHEE: Yeah, exactly. Yeah. Instead, yeah, we're actually encouraged to look in other lanes.

CARLA GEISSER: Yes. And I think that is a unique skill, especially in larger companies and larger organizations. And so that is a huge thing that we bring into the problem space. Plus the technical ability to look at a computer system and either say this is good or this is a problem.

STEVE MCGHEE: Gotcha.

CARLA GEISSER: Which a lot of the decision makers won't have that instantaneous reflex to be like, oh, this thing is actually fine, even though everyone's whining. They're whining for good reasons, but it's perfectly good computers in there. Versus this thing is a liability, and we need to fix it in order to take the next step.

STEVE MCGHEE: Gotcha. Are there cases along the way that you can talk about more, like examples of things where you saw this isn't just a one off? This is a pattern that keeps happening.

CARLA GEISSER: I would say the pattern in government, and to some degree in large corporations, is everyone really, really wants to blame the oldest computer system they have for all of their problems, and they are frequently in the middle of some decade-long transition away from that thing, which is often a mainframe. And our attitude is the mainframe is actually fine. And what instead they have built is a stack of complexity where about every five years, they take whatever is the hot new technology, build a few new things in that, and then get distracted and never complete that migration. So now they have basically one of each kind of technology that has ever been popular over the last 20 years.

STEVE MCGHEE: Wow.

CARLA GEISSER: And that happens all the time.

FLORIAN RATHGEBER: Is it also that they keep building? They don't want to change that old system, and they keep building stuff around it to deal with the quirks of the old system? Because I had that exact experience at a previous job.

CARLA GEISSER: It's a little of-- usually the place they have started is they have a grand modernization scheme where they are definitely going to replace the old thing. And so then they start with maybe a little piece, if you're lucky, or sometimes they start with the grand modernisation plan and get about 60% of the way through. And then you just have a hodgepodge of stuff that's glued onto the side. And you still have the original system, and now you have this other stuff. And we do it again and again and again.

And so drilling through that complexity, I think, is also a skill that tends to coincide with people who describe themselves as SREs. We don't care what kind of computer is in there, as long as we can follow a chain of events. Anyone can run tcpdump. Anyone can look at a log file. It doesn't matter who created the log file. It's all just computers.

FLORIAN RATHGEBER: As long as we can reason about the system and we can understand it, then we don't care what the system is.

CARLA GEISSER: Yes.

STEVE MCGHEE: So the type of people who watch or listen or whatever to this thing, this podcast video thingy that we do, tend to be SREs who don't work at Google or SRE-adjacent folk. So what should they do with these ideas? Is it once they go to work for some large federal government that needs trouble, or is it more applicable than that?

CARLA GEISSER: I think the most useful part of these ideas for anybody who's just coming to work and doing their job is figuring out if a crisis is actually happening. Because if a crisis is happening, you can do interesting and novel things, because leadership is going to be willing to do-- they're going to be willing to break rules. They're going to be willing to move faster. They're

going to be willing to approve exceptions if you're in a real crisis using those criteria I described earlier. If you're not in a real crisis.

FLORIAN RATHGEBER: Yes, that's the one.

CARLA GEISSER: You just have to do your job like a normal person. And you can be mad at everything that's going on that is definitely going to be a crisis someday. But nothing special is going to happen until you hit that tipping point. And so I think that is the both useful and also pragmatic and depressing advice we give to a lot of people where they can see the sky about to fall and they're like, if we just get everybody to care about it, we can prevent the sky from falling. And you can't.

STEVE MCGHEE: So this is the don't-waste-your-time advice.

CARLA GEISSER: This is the don't-waste-your-time advice. You can absolutely take notes and start to learn how the system works and build up your allies who will help you when the crisis actually happens. But until you hit three of those five criteria, nobody's behavior is going to change.

STEVE MCGHEE: Got it.

FLORIAN RATHGEBER: Right. So you mentioned hitting that tipping point. Who actually decides when it starts being a crisis? And who declares? Is there someone that declares the crisis?

STEVE MCGHEE: Is there a crisis commander?

FLORIAN RATHGEBER: Yeah, exactly. Or is it just everyone instantly knows it's a crisis? How do you decide?

CARLA GEISSER: I mean, the way we decide is, because we are external consultants, is somebody has decided to pay the four of us to come help them. And that is enough of a signal.

FLORIAN RATHGEBER: Straightforward criteria.

CARLA GEISSER: Well, it's interesting, because we often have to talk to clients about this, about their willingness to do that and what it will entail once we start working with them. So that's our signal. Within an organization, there has to be someone who is, frankly, scared enough to be willing to break rules. And so you can nudge them using the information you have. But ultimately, someone who has the levers on power needs to be willing to do something new.

FLORIAN RATHGEBER: So once you are engaged, by definition, it's already in a crisis state?

CARLA GEISSER: Or close enough that people are willing to do something new and interesting.

STEVE MCGHEE: I presume that you learned these rules by when it didn't work. You got engaged and you're like, it's not moving, or like we keep getting frozen out or whatever.

CARLA GEISSER: Exactly. Yeah, we kind of back computed the crisis criteria from the experience of places where someone said, this is really important, we definitely need to do something. And we were like, OK, we will definitely help you do something. And then nothing changed.

STEVE MCGHEE: So this is a good way to determine if it's worth continuing or not kind of thing. But we're not trying to encourage people for, oh, if you have two out of three, just cause the third or the fourth one to necessarily.

CARLA GEISSER: I personally do not like that approach. I consider it to be kind of a dark art.

STEVE MCGHEE: Yeah, OK, fair.

CARLA GEISSER: I suspect you all and many people listening to this have worked for leadership who is willing to create a crisis for their own ends. A version that I have seen actually be successful is we're moving everything to the cloud. We have ended our contract with our data center provider two years from now.

STEVE MCGHEE: You just made a hard deadline.

CARLA GEISSER: Congratulations. You have a hard deadline.

STEVE MCGHEE: We made one. Great.

CARLA GEISSER: And sure, you could extend that contract, but it's going to cost four times as much. I don't know if that works, but it's definitely a thing that people try to do to get their migrations to go faster.

STEVE MCGHEE: Motivation.

CARLA GEISSER: Right.

STEVE MCGHEE: OK, well, this is really good. I think this also calls in the phrase never waste a good crisis. Similar kind of thing.

CARLA GEISSER: Yes.

STEVE MCGHEE: I hope you put that on all your literature as you go forward. I'm sure I'm not the first person to suggest that.

CARLA GEISSER: No. I mean, a upside down version of the way we think about this is, great, you shouldn't waste a good crisis. So in that moment, what exactly are you doing? And that's the rest of the story, which we are not talking about right in this moment. But maybe Mikey can fill you in on part two.

STEVE MCGHEE: It sounds pretty good.

FLORIAN RATHGEBER: It's also a bit like when you transition from somebody should do something, we must do something right now.

CARLA GEISSER: Yes.

FLORIAN RATHGEBER: Kind of moment.

CARLA GEISSER: And usually the we must do something right now, because of that taxonomy I described, is everybody gets that feeling at roughly the same time, because it becomes impossible to ignore.

FLORIAN RATHGEBER: So I found the example interesting that you mentioned with we have this migration to the cloud and then we cancel the contract with our DC provider. Are you saying that this is basically, at this point, it becomes a manufactured crisis?

CARLA GEISSER: I mean, it is a management technique that senior leaders definitely try to get the migration to move faster, and it can become a manufactured crisis. I think the risk there is that everybody saw you do it. And so it might work once and then it burns any trust with the people in leadership after that.

STEVE MCGHEE: But people will do that for other reasons as well, which spins off this thing which we now have a word for, which it is a crisis due to the fact that it was already visible. Now there's this deadline, blah, blah, blah. That totally makes sense. I think we're almost at time. Thank you very much, Carla, for coming all the way over here from part of New York to another part of New York. It was fun. We had trouble finding this studio in a maze full of elevators. So if you're ever in New York, there are so many elevators in this town. I cannot believe it. It's pretty amazing. Is there any extra things you want to say to the world, the audience, about what to focus on or how to follow you on Strava or anything weird like that?

CARLA GEISSER: I guess just one final thought for this crisis thing is, there's a bunch of crises that are going to happen that you can predict, and you should just assume that you will have a brief suspension of the rules of engagement during that moment. So organizations that do the holiday shopping season, which they might be starting to prepare for now.

FLORIAN RATHGEBER: They better do.

CARLA GEISSER: They better do it right now. That is all of the circumstances of crisis. And you know they're going to happen.

STEVE MCGHEE: Every year.

FLORIAN RATHGEBER: Yes, exactly.

CARLA GEISSER: And frequently, your leadership will write you a ticket in advance that says,

you are allowed to enforce and break the following rules in order to hit this special time of the year. So I think that is a useful thing to rely on. When we say crisis, people get kind of spun up that it's definitely surprising or it's definitely big and newsworthy, and sometimes it's just an alignment of circumstances that you know is going to happen. But you should use that alignment of circumstances.

FLORIAN RATHGEBER: Does it then actually fit the novelty criteria?

CARLA GEISSER: Doesn't have to. You only need most of them, not all of them.

FLORIAN RATHGEBER: Yeah, that makes sense.

CARLA GEISSER: It's definitely not novel, but it's high visibility, definitely deadline, and probably going to cause a breakdown in understanding of how your system works. Pretty likely. And risk of broken core functions also always there. So you got most of them.

STEVE MCGHEE: I have a random question, and then we can finally close. I already said goodbye. Do you think there are well-functioning companies that just don't have crises on the regular? Or is this a pervasive thing? Does everyone come across this? They just already have this capability in-house potentially without calling it what they're calling it.

FLORIAN RATHGEBER: So I guess could there be a company that cannot hit enough of the criteria? Is that what you're saying?

STEVE MCGHEE: Or they just regularly handle it.

CARLA GEISSER: I think my take, and you'll have to ask Mikey, is that this is actually part of the cycle of change inside of an organization. You have to go through these moments periodically for anything substantive to change.

STEVE MCGHEE: So if you're not having crises, maybe you're not something.

CARLA GEISSER: Yes. I think, basically, the big stuff requires this set of circumstances and a set of brief suspension of normal behavior in order to then get to a new place.

STEVE MCGHEE: Gotcha.

CARLA GEISSER: Yes.

STEVE MCGHEE: All right. Thank you very much. Yeah. The end.

FLORIAN RATHGEBER: Thanks, everyone.

CARLA GEISSER: Nice talking to you.

STEVE MCGHEE: You, too.

SPEAKER 1: You've been listening to the Prodcast, Google's podcast on site reliability engineering. Visit us on the web at sre.google, where you can find books, papers, workshops, videos, and more about SRE. This season is brought to you by our hosts, Jordan Greenberg, Steve McGhee, Florian Rathgeber, and Matt Siegler, with contributions from many SREs behind the scenes. The Prodcast is produced by Paul Guglielmino and Salim Virji. The Prodcast theme is "Telebot" by Javi Beltran and Jordan Greenberg.

SPEAKER 2: You missed a page from Telebot.