

SEI Podcasts

Conversations in Software Engineering

User-Centric Metrics for Agile

featuring Will Hayes and Pat Place as Interviewed by Suzanne Miller

Welcome to the SEI Podcast Series, a production of the Carnegie Mellon University Software Engineering Institute. The SEI is a federally funded research and development center sponsored by the U.S. Department of Defense. A transcript of today's podcast is posted on the SEI website at sei.cmu.edu/podcasts.

Suzanne Miller: Welcome to the SEI Podcast Series. My name is [Suzanne Miller](#), and I am a principal researcher in the SEI Software Solutions Division. Today I am joined by my friends and colleagues, [Will Hayes](#) and [Pat Place](#), who work with me in the Continuous Deployment of Capability group in the SEI Software Solutions Division (SSD). We are here to talk about their recent efforts to reframe measurement in terms of [user stories](#) to improve their relevance and utility.

Welcome, Will and Pat.

Will Hayes: Thank you.

Pat Place: Hello.

Suzanne: I know that both of you have done other podcasts with us, but as you know, we like to give our audience a little bit of background on our speakers. I would like to start with each of you telling the audience that isn't familiar with you already, a little bit about yourself and what brought you to the SEI and the work that you do here. Why don't we start with Pat on that?

Pat Place: Gosh, so what brought me to the SEI was, actually, interest in what was called at the time, formal methods, and that was almost 35 years ago. I have been here a long time and have not stayed doing that all that time. But at the time, we were interested in things like formalism and how you could use that to describe software, and I had appropriate skills. I think what keeps me at the SEI is the constant ability to change, to change directions, to look at new things. I won't give the sordid history of all the things I have done. But for the last five or six years, I have been really involved with the use of [Agile](#) thinking in government programs, and that is what keeps me excited.

Suzanne: OK, and Will, give us a little bit of yourself and why you are here.

Will Hayes: Like Pat, I have been here a little over three decades. My initial assignment in the SEI was with a team called the Empirical Methods project. I was studying to be a psychometrician, if you will, at the beginning of my career. That is the science of assigning numbers and symbols to traits that are pretty elusive and difficult to measure, kind of appropriate for the topic we are pursuing today. I had a long experience working with many different organizations around the world in the CMMI Program. Now, I am a principal engineer, and I lead the Agile Transformation team with my colleagues here. The goal of our team is really to transform what Agile means for the Department of Defense, more so than developing Agile for its own sake. We are interested in exploiting it for a very challenging set of situations where it is applied today.

Suzanne: Yes, and everybody be nice to Will because he is my boss.

Pat: Mine too.

Suzanne: Before we get into user stories per se, let us first discuss what is the current approach to capturing measures in the development and acquisition of software-reliant systems, so we can have a contrast when we talk about user stories. Will, do you want to start with that?

Will: Sure. Often the practice of measurement is an outwardly focused activity; that is, I collect data for someone else's consumption. Often the boundaries that we consider come from a pick list. So there is a graph that has been used, or an index that is popular, or some base measure over which there is a fervent debate happening somewhere in the corner of the Internet about the veracity of those base measures. But often it is a set of things that are chosen among, as opposed to a construct that is designed to

meet a specific need. There is a lot of great work under the rubric of [Goal Question Metric](#), [which] comes from [Vic Basili at the University of Maryland](#), really nice conversations that have been structured around *If I am to collect some data, I should understand the purpose to which that data is being put*. And that purpose should be in pursuit of goals of the team, the division, or the enterprise. Often the veracity of measures is judged by the logical trace. It is sometimes an academic exercise rather than a notion of understanding the utility of the results that we are measuring.

Suzanne: Pat, did you want to add any comments to that?

Pat: Yes, just a little. I think a lot of the programs I have seen were collecting a whole bunch of metrics because those are the ones we have always collected. We just keep collecting them, not with necessarily any sense of purpose. The other thing I have seen at times is where we collect metrics because we are trying to catch the contractor out. We are trying to somehow say, *Hey, you are cheating us in some way*, which really bugs me from an agilistic thinking sense, where I want to be collaborative, not combative. So yes, obviously we want to make sure that the government is getting value for money, but something that is very accusatory may not be of any help.

Will: Yes. Just to build on that, there is almost a forensic accounting mindset sometimes, when we talk about what measures should be gathered; as if somebody on the periphery could reconstruct all the complex things that happen in a large program and make some proof of the perspective they walked into the conversation with, and that is troubling.

Suzanne: Well so, you are trying to do something different. You are trying to help us with understanding the purpose of measures as we use them and collect them. You outline in your recent [blog post](#) one of the things that we brought in with Agile software development, which we have extended beyond software development as well, is this concept of user stories; which, if you think about a requirement being a *what*, then a user story is not only defining the *what*, but it also defines the *who* needs this and more importantly, *why* they want that function. So it starts to get at that goal question metric kind of paradigm, inherently, in the user story. But you have actually taken it beyond requirements for products into using user stories as a way of describing measurement. Let's talk about what do user stories allow you to measure that typical measurement requirements would not give you? We'll start with Pat this time.

Pat: I don't know that we are measuring anything different by using this

format because I am still looking at a measure, but we are bringing in the *Who wants that measure?* and *Why they want that measure?* I think this is psychologically very important. There are a lot of times people feel that they are judged by the measurements they are given. Well, yes, they are. But that they might be judged unfairly, or the data they provide is going to be used in a way that is unexpected. If I say I want a particular measure, and *I am going to do this with it, and this is why I need that measure*, it reduces some of that fear that you might otherwise see other than just saying, *Here is the measure*. I cannot say strongly enough, it gives the *why* of creating the measure. which in one of our discussions with one of our customers they said, *Yes, we forget that all the time. We forget why we are doing this stuff, and this just brings it front and center*.

Suzanne: Will, did you want to add anything to that?

Will: Yes, when I first learned the notion of what a user story is, I was taught that it is the promise of a future conversation. That it forms a basis for speaking in greater detail with more nuance of what is the business problem being solved by the implementation we are going to choose, knowing ultimately that there is more than one way to approach any problem. So the notion that we might choose something different than we have done before because we have a clearer understanding of the need to be met, that really appealed to me. That moves us away from the pick-list mentality that is often associated with measurement programs.

Suzanne: OK, and I am aligned with both of you on the importance of that *why*. Will, I'm glad you brought up the idea that sometimes having the why actually changes the what and can make us think differently about, *Well, what is it? Is this the best way for us to get the why?* is our initial thinking about how to get to that, but it is not always the best thinking once we start having that conversation. So I appreciate that. When you are thinking about this, can one of you give me an example of identifying a user story in a software system that would translate into a measurement? What would a user story for measurement look like? Will, do you want to start us off?

Will: Sure. As a program manager, I need to understand the time it takes to run through the regression test suite so that I may budget for the time horizons required to get to a final build and a delivery. There may be lots of things that contribute to budgeting. There may be lots of things that contribute to timelines. There may be lots of things that contribute to the duration of the test suite running and the number of test cases, the kind of technology available, the diversity of the product that is being analyzed

through that lens. All of those things matter. There may be multiple metrics that bear on answering the need of that program manager. A fairly broad conversation gets opened up by just one statement.

Suzanne: I can see where that is a very different way of talking about this instead of just, *Show me your schedule graph that gives me the time*. I don't know that you are using that for budgeting, but, when you tell me that I have this concern about budgeting, then we can open up that conversation. That is really what we are talking about. Pat, did you want to add anything or did you want to give a different example?

Pat: Yes, I was thanking you for asking me second. It gave me a chance to think. Yes, I was thinking something like, as a quality assurance engineer, I need to know the percentage of coverage for my test cases, so that I can be assured that my test suite is in fact covering all of my code. [That is] kind of a very low-level metric there, but one that is crucially important to somebody who is a quality assurance engineer. Maybe I am predicting your next questions, but this brings in the *who*, which I think is the other side of it: Not everybody needs every metric, but identifying who needs each metric is crucial.

Suzanne: Have you run into cases when you have a broader audience, would you suggest naming all the roles, as program managers and cost account managers, *We need blah, blah, blah*, or do you tend to make the user stories singular in terms of the role?

Pat: Go ahead, Will.

Will: The connection to who is going to consume the information is essential. I think one of the sources of challenge that we experience today is the lack of perceived ownership, that this is a graph we send to somebody out there for purposes that we may not fully appreciate. I think it is really important to be clear about who owns it and, therefore, the kind of responsibility they take on. It is one of the things that I think Pat emphasized well in the blog post: to talk about the level of trust and, kind of, erasing the fear factor when it is clear what consumption is going to come from, or how the data we are producing is to be consumed and what perspective is taken, and so if that changes, there is a conversation we can have about it.

Pat: If I gave my answer to that same question, it would be a little different. That is because, what I usually find is different roles have a different *why* for maybe the same data, so that a PM's [program manager] interest in the data

is often different than the QA [quality assurance] engineer's interest in the data, even though it is the same data.

Suzanne: That is very interesting, because that allows you to understand better... I have a concept that I have used for decades called powerful measures. Powerful measures are ones that are essentially base measures that can be used in different ways for different purposes. It can reduce the amount of data collection. That is why I call them powerful is because I can get multiple uses out of something. When I recombine with other things I get information that is useful to different roles, different stakeholders. I take your point that the *why* is really the thing that distinguishes what is the power that these measures have when we are talking about what the QA engineer needs versus the PM versus the cost account manager versus the team leads. I appreciate that viewpoint. This is a really pretty simple concept, which is nice. We don't always have simple concepts to offer to the world. Have you tested this in any organizational settings that you work in? What kinds of feedback have you gotten when you have used this in customer settings?

Pat: The answer is very much so. In fact, what I quoted earlier of somebody saying, *We all often forget the why*, came precisely out of one of these engagements with a customer organization. They were really very thrilled with this notion because it also allowed them to curate the measures they are currently getting. I will tell you, in one particular organization the contractor is providing a large number of measures. We went through as many people as we could find in the program office, including one of the people who was supposed to be receiving a whole bunch of these measures. And his response was, *Yes, I don't actually look at that data. It is not useful to me*. In this particular instance, it looks like we are asking the contractor to provide data that nobody on the program office side cares about. Well, there is a cost to doing that. By highlighting the fact that nobody is looking at it, perhaps we can say to the contractor, *You don't need to provide us that anymore*.

Will: The really clever way that Pat managed one of those conversations was to take that dashboard or that graphic that was being used and imposing a user story on it. In using the name of that role that was receiving it, and the conjecture at that point about how it would be used, there was almost a suggestion to that person of *What you might do with this* that they recognized *That isn't what I do*. So the conversation was a very different one than, *Can you interpret this graph, or Is there some adverse or positive inference that could be drawn from it someday if it showed some pattern?* That kind of arbitrariness

wasn't really part of the conversation. It was much more a purpose-driven conversation, which really was effective.

Suzanne: I can see where the people who put measurement requirements on a contract are often not the people who are going to end up using those measures for management and understanding the effectiveness of what is going on. If I were to ask people to write their requirements when they are putting them on an SOW [statement of work] in this way, that would probably provide a much richer conversation even at the beginning. You know, so that you could say, *No, we are not going to do that. So, we don't need this measure. We need these, and we don't have anybody.* In the *who* of all these user stories, the PCAM [physical capital asset management], the cost account manager is missing. So what does a cost account manager need? I am envisioning a lot of different ways that these conversations can help at different points in an acquisition. Now you are making me think of how I might use some of this with one of my customers. In the course of doing this, we always like to give people cautionary notes as well. Are there drawbacks to doing this that you have encountered, and how did you overcome them?

Will: In one of my recent presentations, I used the analogy of, *If all you have is a hammer, everything looks like a nail.* In people's eagerness to apply what they see as a fairly logical approach, we find there are times when we kind of need to rein them back in and say, *Well, yes, there is an elaborate story there, but recognize that that space that you are talking about is well addressed already.* So you really want to think about the economic consequence to adding yet another measure. It is true that anything that you care about may deserve more than one lens on it, but just because you can think of a lens doesn't mean it gives you the visibility you need. Due diligence has some balance points, I think.

Suzanne: Sure. Anything from you on that, Pat? What have you seen?

Pat: I haven't seen huge drawbacks yet. I think we are still early days in making this a widely used approach. I think the management of the metrics, how you decide to cycle new ones in and old ones out, is something that we have yet to fully address. I think keeping the *why* with the *what* absolutely points you to...Just as you said earlier, the data we are getting isn't helping us make the decision we need to make, so let's change the data we are getting. But, I haven't seen any sort of huge difficulties. I suppose the one that surprised me was people said, *Oh, it is user stories, therefore we can put it on the backlog and it is done.* It is because something on the backlog, when it comes off the backlog for the product, it is done. Whereas these are not

done. These are just ways of expressing something that is an ongoing measure. But people kind of had this interpretation, because we used the phrase user story, that it is a one-and-done. The only one-and-doneness of it is putting in place the measure. That is a story that belongs on the backlog, but then the actual data, it has to be provided on a continual basis.

Suzanne: And used on a continual basis.

Pat: Yes.

Will: If I could amplify what Pat just said, if you think about the progression of a product from its initial concept development to the time when you might retire that system, the needs you have for insight about progress evolve. If you are forced to choose just one set of perspectives and live with their imperfection throughout the timeline that you are responsible for, you are depriving yourself of the insight you need. This ability to change your mind or to revise the perspective is really important. Just want to amplify.

Suzanne: No, that is a very fair point. I run into that where it is like, *Well, we are not doing these activities anymore, we are more focused on testing activities right now and integration activities. Yes, we are still in doing incremental development, but our focus is different.* So we need to make sure that we amend what we are measuring to really align with that.

You talked early in the podcast about GQM, Goal/Question/Metric, and now we are talking about user stories. There are certainly some similarities between the GQM approach and the user story approach. But what are some similarities and differences in your minds between those two, so that people that might be using GQM already might be able to make a decision as to whether or not they want to include user stories in their practice? You want to start, Will?

Will: Sure. If you think about GQM, it is eminently logical, right? We have goals for our enterprise, for the work we are doing. Those goals beget questions, and those questions can be answered by data. That cascading logic, like so many waterfalls, makes a lot of sense. One of the patterns that we see arising from that compelling logic, though, is that the veracity of the graph I am drawing with Excel, in terms of serving the role it is supposed to serve, is judged by the logical linkage to the question, the logical link to the goal, as opposed to the informative value for the decision maker and the caliber of decision they need to make. So the level of precision I need, the level of relevance to where I am in the enterprise and what outcome I am

trying to assure, that is a much more nuanced, a much more user-driven kind of perspective. We find that the promise of a future conversation that is represented by user stories allows us to get to that nuance serving the user-centric purpose and avoid some of the traps of, *Well, this is logical. It maps to the goals and the questions. You can't tell me I'm wrong. We are going to proceed with this.* It is really meeting a need more than answering a particular question.

Suzanne: Pat, you had a very nice way of talking about this. What is your way of talking about the differences?

Pat: Yes. I mean, from my perspective, GQM is absolutely wonderful. I like it because it is focused on goal. I think metrics as user stories are similarly focused on goals. Each *why* is a goal. But for me, the biggest difference is that the GQM is very top down if you will. We start with the organizational business goals, and all of our metrics have to somehow be coerced into that. I see this approach as being very egalitarian, that everyone can contribute a why. People down at the bottom of an organization have goals. They have things they have to do just as much as people at the top. They don't feel that they have to coerce their specific goals into one of the large-scale organizational goals. I like that. Probably someone in the metrics program will want to try and map individual goals that are user stories up to the large-scale ones. They might even find from the egalitarian bottom-up approach that the organization is missing some top-level goals. I am pretty sure that we'll have covered all of the top-level goals going down.

Suzanne: Yes, true. When you think about this simple concept, we like to talk to people about, *How do you transition this?* And so, this is actually one of the things that I think many of our audience members would go, *Oh, I know measurement, I know user stories, I am in. I will read [the blog post](#), and off we go.* But what are some other resources that we may offer to help them transition into thinking this way and to be more mindful of measurement and Agile settings in general? Will, you want to take that one first?

Will: Yes. So, we have the real pleasure of working with a community of practice that we have sustained for more than a decade now, an Agile Collaboration Group that meets monthly in a Zoom meeting. We have like-minded people from a great diversity of organization types that share their experience, firsthand experience. Sometimes we have people doing a trial run of a conference presentation they are planning to give—a really nice insider group that can talk about such things. [We would certainly welcome more membership in that.](#) There is also a community under practical

software and systems measurement that is been doing work that touches on these areas for a very long time. They have published a [continuous iterative development measurement framework](#) some time back that provides some nice examples and a list of things to talk through. They are also presently focused on a digital engineering measurement framework, which is probably something that a large percentage of our audience for a podcast like this would find interesting. That is a topic where SEI is doing some research presently as well. *How do we apply model-based systems engineering in the settings where we are working?* We are going to have a lot to say about that.

Suzanne: That takes me to my next question, which is, what is next? Give us a little bit more about that. You just teased it, Will. So just give us a little bit more about where you think model-based engineering fits with Agile and Agile measurement. Pat, I will also ask you to say what is next for you, because not all of our interests are exactly the same.

Will: So, what does an iterative and incremental implementation of [MBSE \[Model Based System Engineering\]](#) look like? Many people, when they understand the power of the digital models that they can create to represent the capability they are looking to deliver to the user, they have the whole model in mind. Often the practice of MBSE and modeling falls into the big modeling upfront trap, as somebody said in a recent internet posting. We trap ourselves into *nothing is visible until everything is done* in the model. What does the first increment of a digital model for a large-scale software system look like? It is kind of this similar question that we got asked a decade and a half ago: *What does one week worth of an aircraft carrier look like? How can we apply Agile to such a massive problem?* Very interesting conversation. Look for more of that soon.

Suzanne: Pat, what are you thinking about? What are the things that we are going to bring you back to talk about?

Pat: Well, I mean, for a while now, I have been looking at how you do earned value in an Agile setting. I think there is a lot to be mined in that area, and there is a lot to be said. Then, I have got a sort of a little idea going on about, *Are we asking the right questions when we talk about MVPs [minimum viable products]?* But, it is still a bit half-baked.

Suzanne: Well, we'll give you some time.

Pat: Yes.

Suzanne: It doesn't have to be this week.

Pat: Yes, it won't be this week.

Suzanne: I want to thank both of you for having this conversation with me today. It is always fun to talk about these things. I hope that our audience has also enjoyed this as much as I have. We will be including, as people know, the links and the transcripts to any of the resources that we have mentioned during the podcast, including the [blog post](#) that relates to this topic as well, for those that want to read the blog post. Finally, I want to make a reminder to our audience that our podcasts are available lots of places, SoundCloud, Apple, and my favorite, of course, the SEI's YouTube channel. I want to thank you all for joining us today. I want to thank Will and Pat again for having this great conversation.

Thanks for joining us. This episode is available where you download podcasts, including [SoundCloud](#), [Stitcher](#), [TuneIn Radio](#), [Google Podcasts](#), and [Apple Podcasts](#). It is also available on the SEI website at sei.cmu.edu/podcasts and the [SEI's YouTube channel](#). This copyrighted work is made available through the Software Engineering Institute, a federally funded research and development center sponsored by the U.S. Department of Defense. For more information about the SEI and this work, please visit www.sei.cmu.edu. As always, if you have any questions, please do not hesitate to email us at info@sei.cmu.edu. Thank you.