Predicting Quality Assurance with Software Metrics and Security Methods

featuring Dr. Carol Woody as Interviewed by Will Hayes

Will Hayes: 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.

My name is Will Hayes. I am a principal engineer in the SEI's Software Solutions Division. Today, I am pleased to introduce Dr. Carol Woody, technical manager of CERT Cybersecurity Engineering Group.

Carol, thanks for joining us today. If we could start with a little bit about your background and the things you have accomplished in your career, and what you are bringing to us here at the SEI.

Carol Woody: Well, it has been kind of an interesting trip. I was actually trained as a systems software engineer. When the SEI hired me in—and it has been 15 years, hard to believe—I was brought in to work with the security people and the risk management groups to begin to figure out how do we start to tackle the real challenges of software assurance, essentially building the software so that it has operational security.

Will: There is a particular project we want to talk about today and it involves using metrics in the context of looking at assurance. Can you tell us the background on that work?

Carol: We have been developing methods and practices for key areas. I am going to make sure I do not leave any out: security requirements, software risk management, and supply chain risk management are three of the key areas in addition to training how to do these.

Measuring the effectiveness of their use is our real next challenge. We have lots of qualitative measures but really establishing a quantitative mechanism to begin to really seriously manage this area has been almost the Holy Grail that we have been searching for. But we think we are beginning to make headway in that area.

Measuring the effectiveness of their use is our real next challenge. We have lots of qualitative measures but really establishing a quantitative mechanism to begin to really seriously manage this area has been almost the Holy Grail that we have been searching for. But we think we are beginning to make headway in that area.

Will: There some compelling numbers, I think, that we are seeing in some of the publications in terms of high-level languages that are common today. You have got some statistics that help us gauge the magnitude of the problem. Can you talk a little bit about that?

Carol: Through our research we were able to identify a connection between security vulnerabilities and quality defects. There has been a tremendous body of research—primarily I think Capers Jones is one that has a lot of the metrics on that—in terms of tracking defect densities across various types of software products.

What we have determined through our research is that 1 to 5 percent of the defects are actually vulnerabilities or should be considered vulnerabilities. Now there is a caveat on that; we have very limited data. Very few people are willing to share data about security. What we have been able to identify is primarily linked to our operating systems. There has been some deep analysis in terms of literally looking at each defect and determining what category they are and that is where the 1 to 5 percent numbers come in.

Will: This is an empirical benchmark that has been derived from data and systems that staff at the SEI and CERT have looked at.

Carol: We are leveraging a lot that has been done out in the academic environment. Then we have looked at it and really linked it to several of the quality projects that our group monitors. I believe there are about 100 projects that they have detailed data on. This is primarily through the TSP [Team Software Process] process.

Will: The Team Software Process is a very data-rich process that is been deployed out and we have got a pretty good…. So you are really able to leverage that bit of research in field practice and bring it into a different realm for us then.

Carol: Well, we are and we are not. It was quite interesting because we were able to identify five projects out of these 100 that had excellent security and safety results. We are assuming safety is closely tied, it seems to be closely related to security. Well, if quality were the total answer, all 100 of them would have excellent results.

So we did a deep dive on those five and really identified the linkage of critical security analyses and practices throughout their lifecycles so that they were capturing the vulnerabilities early as well as the defects and then addressing them primarily long before the system was fielded so that they would have very good operational results.

We have a high correlation at least in that small data set from the use of very good methods and practices, that then support improving the addressing of vulnerabilities throughout the lifecycle.

Will: There is a term that we are using a lot now a days, assurance [For a definition of software assurance, please see CNSS Instruction No. 4009; DoDi 5200.44 p.12] , and it seems to relate strongly to this thinking about things up front and understanding the qualities as we are building them. Can you elaborate on assurance as a topic a little bit for us?

Carol: Well, what we are really looking for is assuring that the software functions as intended and only as intended. So that relates to your requirements and how you define what you want it to do, but it also relates to the components that you are buying and assembling to make sure that they function the way you expect them and do not do other things that are unexpected, which may allow an attacker to cause instability or to actually breech certain parts of your system.

Will: Given where we are in the technological evolution relating to software-dependent systems, the possibility that everything would be greenfield and built from scratch is really diminishing.

Carol: I think it is almost zero now.

Will: Almost zero. Many people feel that these are really assembly, not in the assembly language context, but assembly projects where we are taking known components with assurance cases tied to them.

Carol: *Integration* is the key word, and it is integration within the context of where you are fielding the system that determines what it can do and the question is what should it do. So you have got things in the configuration and implementation area that you know you can do from an operational security perspective, but a lot of the challenges we are dealing with in the security world are vulnerabilities that are designed in and then others that are coded in that carry on into this environment. *How do you recognize those and figure out how to either mitigate them or put in place ways that you can identify and recover from them quickly*.

Will: What is next in this line of research? Where are you heading?

Carol: That is an interesting challenge. Obviously, five data points we do not consider to be a strong enough case to really ensure that we know what we are doing. We are in the midst of working with several high-maturity organizations. These groups actually already collect a lot of metrics, so we have got some sort of starting baseline as to where they are in terms of their lifecycle.

We are looking at integrating.We are training them on our policies and our practices and actually integrating it into their policies and procedures. So we are looking at establishing a repeatable mechanism across their projects. Then we will be starting to look at their metrics to see is this consistent because then that will start to strengthen our message.

Essentially we have got this core idea that shows a lot of promise that we are now driving out through some key contact organizations that are interested in working with us. Hopefully we will have a follow-on story to strengthen our message soon.

Will: Like so many efforts at the SEI, there is a handshake and a marriage between research and practice. It is very interesting to see how practice drives the inspirations and research can enable some things in practice that are really ground breakers.

That is great that you are able to work with so many other technologies from the SEI and have a real nice sense of synergy, a word we might not choose to use in a lot of forms. There really is a nice confluence of concepts here, and it is feeding empirical data to your effort. That is great.

Carol: Well, there is that. But, in reality, software assurance is a collaboration because it is the combination of the engineering, the development, the acquisition, and the implementation that all determine what you really can do with the systems and software.

Will: That sounds like a nice vision for the work to push people forward and to have them think about and the role they are playing.

Carol: Yes. Security has traditionally been kind of parallel and thought about separately, but if we are going to really rely on these symptoms, and we expect them to have software assurance, it has got to be fully integrated.

Will: You are tapping into the old adage that some folks under pressure might say, *We never have time to do it correctly, but we always find time to do it over*. Instead of having it parallel, if we integrate it we are really making better use of our time.

Carol, thank you very much for joining us today. Carol has published a blog post on this subject. You can find that on our website at insights.sei.cmu.edu. You can just search for *Carol Woody*, and you will find that blog post as well as many other publications on this area.

As always, a transcript of this podcast will be available on the SEI website at sei.cmu.edu/podcasts and on Carnegie Mellon University's iTunes U site. If you have any questions or wish to follow up, please do not hesitate to contact us at info@sei.cmu.edu. Thank you.