1
00:00:02,000 --> 00:00:12,000
. Good morning, everyone. Welcome to the open security controls assessment language. The 10th workshop in the series

2
00:00:12,000 --> 00:00:22,000
. I'm your host today. We're going to also serve as the all call strategic director. This workshop will be recorded

3
00:00:21,000 --> 00:00:29,000
, so we're would greatly appreciate if you mute your microphones in the process

4
00:00:29,000 --> 00:00:39,000
. Our guests today are coming from the Google Cloud CISO, and we have Vikram Khare

5
00:00:39,000 --> 00:00:49,000
, the director of Continuous Assurance Engineering and Valentin Mihai, Technical Lead. And they're going to share with us the OSCAL adoption

6
00:00:49,000 --> 00:00:59,000
for continuous assurance and beyond. So with that, help me welcome Vikram and Valentine

7
00:00:58,000 --> 00:01:06,000
The microphone is yours?
Let me take over real quickly.

8
00:01:06,000 --> 00:01:16,000
All right, everyone. The topic we're going to cover today is also adoption for continued

9
00:01:16,000 --> 00:01:19,000
assurance and beyond

10
00:01:20,000 --> 00:01:29,000
agenda. This will be about how we're moving towards continuous assurance. Some of the adoption challenges, specifically around data alignment

11
00:01:30,000 --> 00:01:39,000
, a process and systems overview and a little bit more on what some of our future plans are, what we're thinking along the lines of we could aspire to

12
00:01:39,000 --> 00:01:49,000
do if all of this works out. But in terms of introductions, as Michael mentioned, I'm Vikram Khare. I work for Google Cloud and the CISO

13
00:01:49,000 --> 00:01:59,000
office, and I'm the director of Continuous Assurance Engineering. And you'll also be hearing from Val Mihai, who's our technical lead for continuous assurance

14
00:01:59,000 --> 00:02:08,000
. So why? Why do we care about OSCAL? Really, the case for OSCAL is for us, it's not that we

15
00:02:08,000 --> 00:02:18,000
want to go up and adopt an emerging standard, it's that we really want to enable continuous assurance. And for us, ASCO is a very robust way

16
00:02:17,000 --> 00:02:28,000
of exchanging risk and controls data with within different systems themselves between potentially our customers and even regulators.

17
00:02:28,000 --> 00:02:38,000
and partners. We find it's a very comprehensive taxonomy for GRC systems and also continuous controls monitoring platforms. And we're also

18
00:02:38,000 --> 00:02:48,000
interested in the tooling ecosystem that's being built with OSCAL. We think that this could really help drive interoperability and standardization. So the fact that we've seen vulnerability

19
00:02:48,000 --> 00:02:58,000
management systems and also GRC systems coming out in support of OSCAL is a very positive sign for us. And most importantly,

20

00:02:58,000 --> 00:03:08,000
that this is a standardized way of conducting security and compliance audits. You know, if you operate a regulated environment, you'll be dealing

21
00:03:08,000 --> 00:03:18,000
with third party auditors, regulatory agencies and potentially even your customers and shared audits and anything that can be done

22
00:03:17,000 --> 00:03:28,000
to standardize how you present that data to those auditors, to the people who are requesting the audits can allow us to then begin automation

23
00:03:28,000 --> 00:03:38,000
. And I think if anyone has ever worked on an engineering project, you know that you have to have a good set of requirements and a good idea of what can be automated before you can

24
00:03:38,000 --> 00:03:48,000
really do anything like that. And I guess it's also important for us to talk about what we mean by continuous assurance to us. I'll kind of share with you like

25
00:03:48,000 --> 00:03:57,000
what our opinion is. Continuous assurance is kind of an umbrella term for a lot of industry buzzwords. If you hear about things like compliance, privacy, security by

26
00:03:57,000 --> 00:04:07,000
design, what those things are doing is they're helping you build continuous assurance. If you hear about terms like policy as code again,

27
00:04:07,000 --> 00:04:17,000
a well constructed policy is aligned to a set of controls that are being routinely tested in the organization, and ideally those controls are

28
00:04:17,000 --> 00:04:21,000
running in an automated manner

29
00:04:21,000 --> 00:04:31,000
. But a more tangible idea of what all these different buzzwords mean for us is that for every control, you have a set of assurance activities

30
00:04:31,000 --> 00:04:40,000
that are fully automated and that you also have real time monitoring of the controls based on objectively defined metrics. And I think the last thing we want to talk about

31
00:04:41,000 --> 00:04:50,000
is is that we want to create compliance failures like system downtime. A lot of times what happens is that in regulated environments, there's a level of subjectivity that

32
00:04:50,000 --> 00:05:00,000
you meet a requirement, do not meet a requirement, a failure and a compliance obligation. While important, it may not be treated with the same urgency as like a security

33
00:05:00,000 --> 00:05:10,000
breach or a server going down. And we want to kind of shift things and the mindset such that any time we do have a compliance violation, it's treated with the same

34
00:05:10,000 --> 00:05:15,000
sense of urgency that, like a server outage, would be

35
00:05:15,000 --> 00:05:25,000
. So one of the challenges we've run into with the adoption of Arsenal and moving towards continuous assurance. I think some of the minor challenges we've run into

36
00:05:25,000 --> 00:05:35,000
is the timeline. There's obviously public sector requirements to support OSCAL final assessments, and so that requires some engineering work that

37
00:05:34,000 --> 00:05:44,000
we're doing. It also requires us to rethink how we are cataloging controls. So the onboarding of the controls has to be rethought of

38
00:05:44,000 --> 00:05:54,000
. We also need to think through how we do data change management on our controls. The other minor challenges are that obviously our scale is an emerging standard, so

39
00:05:54,000 --> 00:06:04,000
it's still in the process of being developed and finalized. And also, we have to standardize a lot of our taxonomy and rethink some of the data models that we have internally

40
00:06:04,000 --> 00:06:14,000
. What are the major challenges we run into just because of the size and scale that we operate and there are third party software challenges? So while even though there is

41
00:06:13,000 --> 00:06:23,000
kind of an ecosystem of tools being developed around also, it's not necessarily the case that we would be adopting them. Organization of the data is

42
00:06:24,000 --> 00:06:33,000
very challenging for us and we'll get into what we're doing there. And lastly, any sort of technical that needs to be done away with. So if you are doing

43
00:06:33,000 --> 00:06:43,000
continuous monitoring as soon as you're finding gaps, you have to go deal with them. You know, typically a control may be tested on a bi

44
00:06:43,000 --> 00:06:53,000
annual basis or an annual basis. If you move towards continuous monitoring of those controls, you're really increasing. The frequency to be ideally would be real time

45
00:06:53,000 --> 00:07:03,000
. But typically, most people in the industry would say that a control is continuously monitored when it's tested on at least a monthly frequency. And once you do that, you'll find the gaps more quickly and you'll have

46

00:07:03,000 --> 00:07:13,000
to get them remediated more quickly. So how are we moving towards continuous assurance? Really, the first phase of all of this is we've just

47
00:07:13,000 --> 00:07:22,000
we've completed like a proof of concept. You know, we are basically using templates and scripts to see if we can generate large-scale file

48
00:07:22,000 --> 00:07:32,000
assessments. We're moving into what we call the all scale builder phase of an MVP, where it'll be more UI driven and web driven

49
00:07:31,000 --> 00:07:41,000
. We're also going through a very extensive data pre population exercise and our GRC systems, you know, making sure that the asset inventory, the vulnerability

50
00:07:42,000 --> 00:07:52,000
management information is all centralized correctly. And we're also thinking through usability enhancements. And then finally, it'll be a

51
00:07:52,000 --> 00:08:01,000
base for will be maturing the whole process really. At Phase three, what we're doing is a lot of the data collection will be done in a somewhat manual manner. Some of it automated, some

52
00:08:01,000 --> 00:08:11,000
of it manual. And really, with phase four, we want to have like a two way integration system with our GRC system, where most of the data collection that needs to be done

53
00:08:11,000 --> 00:08:17,000
is fully automated, including how we externalize the data to auditors

54
00:08:17,000 --> 00:08:27,000
. So in terms of aligning the data, we have to really think through the entire data lifecycle here. And that begins with like how we capture our requirements later on

55
00:08:26,000 --> 00:08:36,000

. In the slides, you'll see Val POC through a piece on the overall end to end process about how we get like internal compliance requirements and external

56
00:08:36,000 --> 00:08:46,000
compliance requirements. The actual compliance requirements are all around regulatory decomposition, and with Austell, you do have a way of generating machine readable

57
00:08:46,000 --> 00:08:56,000
formats for different regulations, not just fed ramp. And so for us, we have to start thinking about how we're managing our security and compliance requirements. Really

58
00:08:56,000 --> 00:09:06,000
at that intake level and that begins with the regulatory decomposition. We need it to come up with a more granular structure for defining controls. We need a way to like draw

59
00:09:06,000 --> 00:09:16,000
like a straight line from what are the regulatory requirements we have to how where the controls are actually being implemented and then into specific metrics for measuring

60
00:09:16,000 --> 00:09:26,000
the control performance. And you'll see like how we think through control metrics and the aggregation process we have to go through. Another big part of

61
00:09:26,000 --> 00:09:35,000
this is really refining the asset, the asset model. With the adoption of scale, we really have to get much more granular and detail on what we're doing

62
00:09:35,000 --> 00:09:45,000
. And so, you know, establishing an asset taxonomy that takes into account tooling and automation and everything else has become very important for us.

63
00:09:45,000 --> 00:09:55,000
And then finally, the last pieces that we're aggregating data from disparate sources, you know, I think you've heard me talk about how

this requires more granularity. So I'll say it again,

64
00:09:55,000 --> 00:10:04,000
The hospital definition of a control requires a higher level of granularity. So internally, we've rewritten our data models and we are also looking at like

65
00:10:04,000 --> 00:10:14,000
populating some of the data that we're keeping an external resources. We consolidate it into the GRC system. Probably one of the most important lessons learned. We

66
00:10:13,000 --> 00:10:23,000
have around ourselves is that when you think about it, it's adoption. You really have to think about one getting it adopted into your GRC system first

67
00:10:23,000 --> 00:10:32,000
. So with that said, how does this actually look when we get a specific control like, for example, this is

68
00:10:32,000 --> 00:10:42,000
a control that simply states that a cloud service provider has an eddy program, an anti-virus program. We'll look at that and we may have like supporting controls

69
00:10:43,000 --> 00:10:52,000
that we have defined in our GRC system that we have these scans running daily and prod that all these findings are resolved in 24 hours

70
00:10:52,000 --> 00:11:02,000
and that all all incoming data into a data center is scanned on an annual basis. Now this is a hypothetical. This isn't how we actually run it, but let's say

71
00:11:02,000 --> 00:11:11,000
that that breaks down into like two different centers within the US, two different data center. We would then look at like the control implementations as you kind of

72

00:11:11,000 --> 00:11:21,000
like, shift to the right on the slide, you'll see the boxes in the blue with the control implementation. And this will have details like in North America, the data center, CSP

73
00:11:21,000 --> 00:11:31,000
is running scans using certain software on all production systems. It'll talk about where the AV scan results are aggregating

74
00:11:31,000 --> 00:11:41,000
data. And if you go down, you'll see like there's another set of blue boxes there around cloud FCA. And basically, what we're doing there is we could potentially be running a different

75
00:11:41,000 --> 00:11:51,000
eighties, right? And so what we have here is kind of like a breakdown of like what is the actual requirement that we have to meet and how we have like aligned it? According to OSCAL,

76
00:11:51,000 --> 00:12:01,000
internally. And the value of all this is that one of the values of all this is that this level of granularity that you see here

77
00:12:01,000 --> 00:12:11,000
, we basically have like different with this kind of granularity, we can actually do the automation we'll actually know and have it clearly documented that there are different antivirus

78
00:12:11,000 --> 00:12:21,000
systems and different data centers that can help us make sure that we're doing the automation correctly and it can excuse me for the wrong slide. It can also help us make sure that

79
00:12:21,000 --> 00:12:31,000
we're looking at the continuous controls monitoring correctly. So you know what you have here in terms of how we can measure a continuous assurance

80
00:12:31,000 --> 00:12:41,000
is on the left. That initial requirement. The CSP has an aid program. It's going to be an aggregate measurement. You know, it'll be either red

81
00:12:41,000 --> 00:12:50,000
yellow or green red, meaning that the control is failing. Yellow means that it's in a warning state and green me that it's worse now. What is that determination of red

82
00:12:51,000 --> 00:13:00,000
, yellow, green? It's an accurate of what you see on the right, which are the different ECM metrics. This could be the percentage of scans that are being completed

83
00:13:00,000 --> 00:13:10,000
, the time it takes to complete all the vulnerability remediation and whether all incoming data is scanned. We can assign weights to them and we

84
00:13:10,000 --> 00:13:20,000
can figure out just exactly what state the control is functioning again. And with that, I'm going to hand this over to balance the clock about our

85
00:13:19,000 --> 00:13:24,000
high level process and just take you to the rest of the presentation

86
00:13:24,000 --> 00:13:34,000
. Thanks for the call. So when we when we look at our process, got it broken down into, let's say, three

87
00:13:34,000 --> 00:13:44,000
three key top three key areas, right? We have the intake process. We have a cataloguing and onboarding process as well as a continuation

88
00:13:44,000 --> 00:13:54,000
continuous assurance phase and fundamentally during the intake process. The focus is on consuming various data sources, such as external compliance reviews,

89
00:13:54,000 --> 00:14:04,000
internal compliance reviews, our risk assessment methodology, external regulations, best practices, external standards,

90
00:14:04,000 --> 00:14:14,000
etc., as well as our own control, maturity and control lifecycle process. And all of those things contribute to something called a control

91
00:14:14,000 --> 00:14:24,000
graph. And the idea meaning is that Bill's requirements fundamentally don't align 100 percent with the control catalog in the current capabilities

92
00:14:25,000 --> 00:14:34,000
that we're measuring and we're assessing. So as part of that controls gap definition, we then pivot into the creation of a control so we go through the whole process of the control

93
00:14:34,000 --> 00:14:44,000
definition. And as Vikram had highlighted on the previous slides, there's a we have internal best practices in terms of the level of granularity and how those controls need to be defined

94
00:14:44,000 --> 00:14:54,000
, right? So there's this concept that the control needs to be granular enough that the control objective, drawing it all the way through to

95
00:14:54,000 --> 00:15:04,000
the actual metrics during the assessments and the poems and the remediation that it's all directly linear, you can see the the connection between

96
00:15:04,000 --> 00:15:14,000
them. And a lot of this is going to be fundamentally driven by the scope, right? So essentially, the scope is going to define the control implementations

97
00:15:14,000 --> 00:15:24,000
that's going to impact the control definitions and having that over to the control creation modification process, right? Once once the controls

98
00:15:23,000 --> 00:15:30,000
have been identified, we look at a couple of different options. One,

is this a control that has

99
00:15:30,000 --> 00:15:40,000
, let's say, the right tooling that is currently already in place, in which case we can go directly to the control being onboarded and we can start to observe it and measure it in other scenarios. We have these two diverging

100
00:15:41,000 --> 00:15:50,000
paths, meaning that either we don't have, for example, definitions for those measurements or there is an additional requirement because it's a manual process

101
00:15:50,000 --> 00:15:59,000
that assurance automation be inserted, in other words, to make the process observable in order to be able to generate metrics to drive that automate

102
00:15:59,000 --> 00:16:09,000
. But once we've gone through this whole phase of integrating, identifying the gaps, doing the controlled definition, going through the creation

103
00:16:09,000 --> 00:16:19,000
process, the metrics onboarding and getting to the observability, we now have a control that's ready to be consumed and that consumption actually approaches

104
00:16:19,000 --> 00:16:28,000
into two different areas. One, we can start to externalize and create our audit artifacts, for example, the scale package, because as part of that control, onboarding

105
00:16:28,000 --> 00:16:38,000
, we've had to set the control objectives. We've got to identify the systems both. We've had to understand the implementation, which is essentially the fundamental building blocks that are required by all skill

106
00:16:38,000 --> 00:16:48,000
. And then all of that that data and metadata, as Vikram was highlighting earlier, enables us to be able to turn on continuous

monitoring because we have source systems, we have targets

107
00:16:48,000 --> 00:16:57,000
, systems we have in scope assets. We have all of these elements that now feed into it. And now we can simply apply thresholds for things like alerting and notifications.

108
00:16:58,000 --> 00:17:07,000
identify responsible parties and really moving towards what we call control, reliability, engineering right

109
00:17:07,000 --> 00:17:17,000
? Anybody's familiar with Ussery same type of concept where we look at control failures, we can start to look at them the same way that we would with any other sort of failure or outage or event

110
00:17:16,000 --> 00:17:22,000
that happens within our services. We look at the at slightly different

111
00:17:21,000 --> 00:17:31,000
. So in order for us to build the backend infrastructure

112
00:17:32,000 --> 00:17:39,000
, we really focused on categorizing the systems and understanding how these different elements need to come together.

113
00:17:39,000 --> 00:17:49,000
When we look at it, we have really five five buckets for core that are, I would say, almost unique to the core process. So obviously, we have a whole lot of inputs

114
00:17:49,000 --> 00:17:59,000
, right? So various control signals, policies, risk data, third party services, customer commitments, contracts, external regulators, industry, best practices,

115
00:17:59,000 --> 00:18:09,000
our own internal practices, best practices and policies. All of those things can be considered inputs. All of that stuff gets aggregated in our systems of record

116
00:18:09,000 --> 00:18:19,000
. Our Systems of Record Act as the primary aggregation points in order for us to be able to establish the right relationship model between all of these various inputs that

117
00:18:19,000 --> 00:18:29,000
can have different types of information and to be able to correlated and align it to that scale model. So some of these could be controlled limitations and impose

118
00:18:29,000 --> 00:18:38,000
some of these could be requirements that we would then map the controls. Some of that could be related to components and systems. Other things could actually be the metrics themselves. So what we do

119
00:18:38,000 --> 00:18:48,000
is we gather all of this information and these and these systems of record, which are fundamentally the let's say about the heart and soul of the way that we perform

120
00:18:48,000 --> 00:18:57,000
the camera. And from that piece, the data can either be consumed in data analytics or directly piped out the customer. When we talk about the analytics side of it,

121
00:18:57,000 --> 00:19:07,000
we're really focused on things like the risk assessments to be able to do our own risk assessments and risk quantification to do any sort of upstream alerting

122
00:19:07,000 --> 00:19:17,000
and monitoring control to threat mapping and modeling with all of these things are part of that data analytics ecosystem. So basically, how are we joining this data? Are we contextualizing

123
00:19:17,000 --> 00:19:27,000
it, et cetera? And other data directly from the systems of record or from these analytics are then used for end user outputs? Are those entities your outputs can be things like

124
00:19:27,000 --> 00:19:36,000

generating an SSP, going a dashboard and audit report, sharing assessment plans, as well as end user outputs being for our own internal

125
00:19:36,000 --> 00:19:46,000
consumption. So for our own risk quantification, compliance assessments, those types of things, and obviously because all of this data has a significant

126
00:19:46,000 --> 00:19:56,000
sort of compliance and legal angle to it, we do have this concept of an evidence repository and this is partially the sort of data

127
00:19:56,000 --> 00:20:06,000
retention requirements, but it's also to help us manage to the right works. The idea being that we want to get to a place where anything that's shared externally

128
00:20:06,000 --> 00:20:16,000
can be traced back to the bits and bytes internally that helped us make those assessments and those assertions right. So this is sort of the vision of the ecosystem that that we're in the process

129
00:20:16,000 --> 00:20:25,000
of building and developing and putting together. The support can be given to the next logical

130
00:20:26,000 --> 00:20:35,000
. So the areas where we're seeing a lot of opportunity as we've been working through this and as we've been developing this ecosystem of tools right

131
00:20:36,000 --> 00:20:45,000
, a couple of areas come up. So one, obviously the immediate one would be the audit 12 reduction. So obviously said we'd we photograph

132
00:20:44,000 --> 00:20:54,000
is very interested in how we're actively exploring it, but it would be great. Want to see this adoption go beyond the US public sector?

133
00:20:55,000 --> 00:21:05,000

right? And also see the expansion of scale to include CCF right and really getting into the metrics side of it to be able to

134
00:21:05,000 --> 00:21:14,000
not only share sort of the assessment plans and all that and all about aspects of it, but see if there's a way for us to shorten the actual audit and evidence gathered together and say that

135
00:21:14,000 --> 00:21:24,000
right? And obviously, you know, the more regulatory bodies that adopt these metrics in this whole process, the more we could scale, the more can be refined, the more

136
00:21:23,000 --> 00:21:26,000
can be perfected right

137
00:21:26,000 --> 00:21:36,000
? Additionally, you know, having more participation and collaboration from regular regulators and standards bodies to help us establish a robust data sharing architecture. So

138
00:21:36,000 --> 00:21:46,000
we have ideas, we have mechanisms on how we could do things. But you know, they're not industry standards. And to understand how others

139
00:21:46,000 --> 00:21:56,000
are intending to consume the data and helping them, helping us align with what they are intending to do, it would be something that would be of great interest to us

140
00:21:55,000 --> 00:22:05,000
. The other aspect of it is this idea of the analysis world reduction. So as I mentioned during the intake piece, as we decompose regulations, as we decompose requirements

141
00:22:05,000 --> 00:22:11,000
, we end up building a pretty robust catalog of controls right

142
00:22:11,000 --> 00:22:21,000
? Currently, the process is relatively manual or historically it's been relatively flat. What we're looking to do is to find mechanisms

to introduce AI and other technologies

143
00:22:21,000 --> 00:22:31,000
to accelerate that process, to basically feed it a PDF capital controls catalog and help spit out and give us some sort of rapid assessment as to how we align the specific

144
00:22:31,000 --> 00:22:41,000
frameworks and regulations, right? That's that's one way to kind of reduce the total reduction in that. And the other way is also to see if there's opportunities for us to facilitate

145
00:22:41,000 --> 00:22:51,000
it by providing tools and technologies to regulators and to others that would allow them to convert the existing regulations into machine readable

146
00:22:52,000 --> 00:23:01,000
formats. Or we provide the technologies that do that to have them review and sign off, right? So this can help drive some of that standardization and catalogs like we're seeing

147
00:23:01,000 --> 00:23:11,000
that today with that ramp, but it's not something that we're seeing universally right. That's, you know, that's that's a big area for us where we see a lot of opportunity to drive

148
00:23:11,000 --> 00:23:19,000
standardization and minimize confusion associated with the potential interpretation of a regulation

149
00:23:19,000 --> 00:23:29,000
. The other thing is this is the integration of broader controls of sort of controls and product management and security operations. So

150
00:23:28,000 --> 00:23:38,000
right now, we have all this, this very valuable information that's very much used to drive these compliance requirements and to measure our ability to meet

151
00:23:38,000 --> 00:23:48,000

these security standards. What's interesting to us is this this ability to kind of match these controls to threats and to other

152
00:23:48,000 --> 00:23:58,000
security incidents to understand how controls could potentially help mitigate some right, as well as using incidents in reverse to be able to assess the

153
00:23:58,000 --> 00:24:07,000
effectiveness of controls. Right. So the idea is, is that you design something you implement that you're constantly measuring it, but then you to go

154
00:24:07,000 --> 00:24:17,000
back and say, OK, it's been effective in mitigating these areas, these threats, these security events, preventing them, reducing risks, et cetera. But also what hasn't it been effective or

155
00:24:17,000 --> 00:24:27,000
how can it be matured? And how can it be addressed, adapted and tweaked to deliver more value, right? Kind of going back to that initial cycle in that slide where we

156
00:24:26,000 --> 00:24:36,000
have that maturity box near the bottom. So that's that's sort of the the area with a future opportunity. Again, we're still sort of connecting the dots and building out that infrastructure

157
00:24:37,000 --> 00:24:46,000
. But as we're adopting these things, we're noticing that a lot of these opportunity areas are arising that are kind of getting us to

158
00:24:46,000 --> 00:24:56,000
the see and to explore how we can leverage this framework and fundamentally what this framework helps us surface out of our own sort of internal practices and

159
00:24:56,000 --> 00:25:06,000
to apply to a broader or broader set of technologies and disciplines. If I'm not mistaken, this may be the first

160

00:25:05,000 --> 00:25:08,000
. That's it.

161
00:25:08,000 --> 00:25:14,000
All right. Thank you, everyone for taking some time out to hear our presentation

162
00:25:15,000 --> 00:25:24,000
. Thank you very much Vikram and Val, I will stop the recording and we can move to the open discussion. Thank you.