Nemko Digital

Scaling AI with a Risk-Based Control Framework - Transcript

1
00:00:20.140 --> 00:00:21.879
Pep Van Der Laan: Hello, everyone!

2
00:00:22.580 --> 00:00:25.090
Pep Van Der Laan: I think we have some people joining.

3
00:00:25.340 --> 00:00:26.540
Pep Van Der Laan: Very nice.

4
00:00:34.240 --> 00:00:41.000
Pep Van Der Laan: So let me wait a little bit until the counter… It's roughly stabilizing.

5
00:00:42.670 --> 00:00:45.010
Pep Van Der Laan: Before we kick off…

6
00:00:49.560 --> 00:01:09.230
Pep Van Der Laan: So… Perfect, yeah, so maybe now is the, is the moment to slowly, slowly get, get started. So, welcome everyone at this, at this webinar, Nemko Digital. Today's topic, Scaling AI with a Risk-Based Control Framework. So…

7
00:01:10.570 --> 00:01:16.270
Pep Van Der Laan: Before we dive into that topic, let me…

8
00:01:16.920 --> 00:01:30.079
Pep Van Der Laan: briefly flash the agenda, what we'll be doing today. First, briefly talking on Nemko Digital, just a couple of minutes for everyone who is here for the first.

9
00:01:30.250 --> 00:01:43.580
Pep Van Der Laan: time, then diving into different perspectives on AI risk, which Alicia will, will lead, and then following into a case study, which we will both share on.

10
00:01:43.640 --> 00:02:02.220
Pep Van Der Laan: And then also giving you a lot of time for Q&A, depending on what the asks of the audience is. So, before we dive into that, I think good to introduce ourselves. So, Alicia, can I give you the first honors?

11
00:02:02.820 --> 00:02:20.949
Alicja Halbryt: Yes, hi everyone. Thanks, Pep. My name is Alizia, I have been with Nemko Digital, since last year, November. I am the AI and Digital Trust Consultant, with a background in human-centered design, as well as, philosophy of technology, so…

12
00:02:21.000 --> 00:02:38.740
Alicja Halbryt: as you can tell from that background, I am quite focused in the ethical side of things, and AI specifically. And I have also some experience working with, the Dutch government here,

13
00:02:38.750 --> 00:02:41.839
Alicja Halbryt: In the AI standardization, team.

14
00:02:43.370 --> 00:02:53.710
Pep Van Der Laan: Perfect. So, thanks, Alicia, great to have you here today. So, I myself, I'm Pep, I'm the Technical Director at, at Nemko Digital.

15
00:02:53.820 --> 00:03:10.320
Pep Van Der Laan: doing this from a background where, for years and years, I have been focused on building AI solutions and scaling them, and now at Nemko Digital, really focusing on, yeah, how do we make sure that from an AI trust perspective, from a governance, from a compliance perspective, those

16
00:03:10.320 --> 00:03:17.700
Pep Van Der Laan: solutions are good to go, because basically, what I saw in my experience

17
00:03:17.700 --> 00:03:34.300
Pep Van Der Laan: sitting on the other side of the table, working on implementing and scaling the solutions, was that, in the end, the bottleneck of that scaling, yeah, is typically not in the technology, it's typically in how do you make sure that those systems are trusted and,

18
00:03:34.770 --> 00:03:44.080
Pep Van Der Laan: and reliable in such a way that you can… that you can scale with confidence. So that is what we also will be talking about today.

19
00:03:46.500 --> 00:04:05.450
Pep Van Der Laan: So, before we do that, briefly introducing, Nemko Digital, for those of you who don't, know yet. So, Nemko Digital, part of Nemko Group, Nemko Group, strong in technical compliance and,

20
00:04:05.450 --> 00:04:13.629
Pep Van Der Laan: testing, inspection, and certification already for, for a long time, over 90 years. Of course, with

21

00:04:13.650 --> 00:04:17.380
Pep Van Der Laan: the changing world of AI, that

22
00:04:17.440 --> 00:04:30.199
Pep Van Der Laan: testing and certification of products increasingly included, digital products, so cybersecurity became, important. More recent also, data and AI trust.

23
00:04:30.200 --> 00:04:39.009
Pep Van Der Laan: So, that is what we're building as an expert center out of, out of Amsterdam, basically servicing,

24
00:04:39.380 --> 00:04:56.499
Pep Van Der Laan: the whole of the Nemko network, covering over 28 locations on 3 continents. We do that in the spirit of Nemko, so compliance without complexity, and really with a strong emphasis on trust.

25
00:04:57.930 --> 00:05:08.519
Pep Van Der Laan: We help our clients, of course, in different, in different settings across industries, focusing on making compliance and governance

26
00:05:08.520 --> 00:05:25.640
Pep Van Der Laan: actionable, so in that sense, we really have a maker's mindset that we bring to the compliance game. Next to the work for our clients, we also continue to build our thought leaderships in collaborating with our partners.

27
00:05:25.940 --> 00:05:42.210
Pep Van Der Laan: And if you look at what we do on a day-to-day basis, what are the topics that we help our clients with? There's, of course, regulatory compliance, the AI Data Act when it comes to AI, but also other acts like,

28

00:05:42.480 --> 00:05:55.389
Pep Van Der Laan: Like, Cyber Resilience Act is now a whole topic. We support our clients on being ready for ISO, in particular ISO 42001, on

29
00:05:55.670 --> 00:06:10.829
Pep Van Der Laan: on AI management system. In the end, often this is also about global market access, so how do I make sure that my products and services are ready to come onto the local markets?

30
00:06:10.830 --> 00:06:21.460
Pep Van Der Laan: often that's also focused in particular for that digital domain, getting onto the European market, where regulations are typically maybe a bit more strict than in some other,

31
00:06:21.460 --> 00:06:22.880
Pep Van Der Laan: areas.

32
00:06:23.100 --> 00:06:38.220
Pep Van Der Laan: And we have our Nemko AI Trustmark, which is basically a way to make sure that you can certify, in this voluntary scheme, your products, your software solutions as

33
00:06:38.370 --> 00:06:50.379
Pep Van Der Laan: being up to standard when it comes to AI trust. Trust marks strongly aligned also with the principles of, of both ISO and,

34
00:06:50.840 --> 00:06:54.520
Pep Van Der Laan: and NIST, as well as the AI Act framework.

35
00:06:54.680 --> 00:07:07.139

Pep Van Der Laan: So, those are all, on the left-hand side of the… of the slide, the things that are very important when it comes to more the focus on regulatory compliance and,

36
00:07:07.190 --> 00:07:19.689
Pep Van Der Laan: And standardization, often in our collaboration with our clients, that branches out into other questions around technical assurance, about governance maturity.

37
00:07:19.690 --> 00:07:36.610
Pep Van Der Laan: In those, in those domains, as well as, yeah, shaping the, shaping the strategy for, for AI, cyber, and data. So, broad set of, set of services, that we support our, our clients with, but with that, yeah, let's focus on…

38
00:07:36.620 --> 00:07:55.540
Pep Van Der Laan: the specifics of AI risk, because you already saw a lot that we talk about in the… with our clients is about regulation and standardization, but there's more to the picture than that. So for that part, Alicia, can I hand over to you?

39
00:07:56.690 --> 00:07:59.479
Alicja Halbryt: Thanks, Bob, for the introduction.

40
00:07:59.720 --> 00:08:18.869
Alicja Halbryt: So, there are two main approaches to handling AI risk globally, and one of the risk controls is, of course, in a way, regulations. We can see many developments happening around the world when it comes to building AI strategies.

41
00:08:18.870 --> 00:08:21.580
Alicja Halbryt: And here are just a few.

42

00:08:21.700 --> 00:08:39.890
Alicja Halbryt: most advanced examples outside of the European Union, to which I will come back in a minute. You can see here, for instance, that the US wants to focus mostly on allowing innovators to experiment and be more or less free with their ideas.

43
00:08:39.890 --> 00:08:43.429
Alicja Halbryt: Similarly to the UK, where there's…

44
00:08:43.429 --> 00:08:49.080
Alicja Halbryt: no standalone regulation for AI, and there's a pro-innovation approach.

45
00:08:49.450 --> 00:09:08.599
Alicja Halbryt: Asian markets are also far in their AI regulatory developments, with South Korea, for example, emphasizing the human-centric protections, and on the other hand, China, where, state oversight plays, the major role.

46
00:09:08.600 --> 00:09:20.160
Alicja Halbryt: And if you are curious about, Asian AI regulation landscape, you can have a look at our recent Insight article, which will be shared in the chat in a second.

47
00:09:21.600 --> 00:09:25.229
Alicja Halbryt: So if you go to the next slide…

48
00:09:25.530 --> 00:09:32.109
Alicja Halbryt: Most regulations, if not all of them, focus on societal risks.

49
00:09:32.370 --> 00:09:40.279
Alicja Halbryt: Because this is their main goal, right? To protect the citizens and make technology work in our favor.

50
00:09:40.590 --> 00:09:56.879
Alicja Halbryt: These regulations often take a risk-based approach, and the EUA Act here is a perfect example of this, because one of its aims is to protect health, safety, and fundamental rights.

51
00:09:57.010 --> 00:10:07.920
Alicja Halbryt: of individuals, and I'm sure many of you are familiar with how the AI activates AI systems, and most importantly, their use cases.

52
00:10:08.190 --> 00:10:21.460
Alicja Halbryt: into risk categories. So we have the prohibited AI level, which are, for example, social scoring, systems or emotion recognition in the workplace.

53
00:10:21.650 --> 00:10:31.480
Alicja Halbryt: Then we have systems that are high risk, but still allowed, only with multiple strict controls needed to be put in place.

54
00:10:31.610 --> 00:10:38.020
Alicja Halbryt: It would be, for example, biometric identification of people or credit scoring.

55
00:10:38.870 --> 00:10:47.380
Alicja Halbryt: And then AI-generated media or, simple chatbots are considered… are considered a limited risk.

56
00:10:48.040 --> 00:10:56.130
Alicja Halbryt: So, as you can see, all these examples can and do pose societal risks.

57
00:10:56.410 --> 00:11:00.279

Alicja Halbryt: Which the law is trying to manage and control.

58
00:11:00.740 --> 00:11:06.789
Alicja Halbryt: And this is one of the perspectives, to look at the AI risk.

59
00:11:07.110 --> 00:11:08.840
Alicja Halbryt: Next slide.

60
00:11:10.420 --> 00:11:27.400
Alicja Halbryt: Now, to support the development of AI systems that, that pose these, risks to individuals and society, the EU Commission sent a standardization request to the European Standardization Body, specifically SENS and Alec.

61
00:11:27.880 --> 00:11:34.890
Alicja Halbryt: They requested experts to write standards in these, 10 areas that you can see.

62
00:11:35.080 --> 00:11:50.249
Alicja Halbryt: And as you can also see from this standard architecture, and as you can imagine, it is a very complex process to put it all together, to write all these rules, to get the hundreds of experts working together.

63
00:11:50.540 --> 00:11:56.880
Alicja Halbryt: And agree. Which, in case of the AI standards, this…

64
00:11:57.070 --> 00:12:05.569
Alicja Halbryt: of course, results in a very big delay in delivering these documents. And if we go to the next slide…

65

00:12:08.310 --> 00:12:17.710
Alicja Halbryt: Most of these standards are still in drafting stage, even though the original deadline to deliver them has long passed.

66
00:12:17.890 --> 00:12:31.649
Alicja Halbryt: Three of them are in final, or almost final stage at the moment, but anyway, the whole package of harmonized standards is not expected to be completed before, 2027.

67
00:12:31.830 --> 00:12:33.680
Alicja Halbryt: And on the next slide.

68
00:12:34.250 --> 00:12:43.119
Alicja Halbryt: I provide a more in-depth overview of what these standards that we will see published the soonest entail.

69
00:12:43.120 --> 00:12:56.039
Alicja Halbryt: The most relevant, to us today, of course, is the first one, the draft harmonized, standard 18228 on AI Risk Management.

70
00:12:56.390 --> 00:13:02.770
Alicja Halbryt: It describes how risks, can be managed during the whole AI system lifecycle.

71
00:13:03.480 --> 00:13:22.440
Alicja Halbryt: It's now in public inquiry stage, which means that anyone can read the draft and provide comments, which then will be taken back to the committee, discussed, and once changes are implemented into the draft, then the standard will go into the formal vote stage.

72

00:13:23.220 --> 00:13:35.489
Alicja Halbryt: Now, all this, so regulations and standards, are there to make sure society is benefiting from the technology, and that risks to people are mitigated and controlled.

73
00:13:35.860 --> 00:13:42.430
Alicja Halbryt: However, For businesses, focusing only on these governmental rules.

74
00:13:42.590 --> 00:13:46.449
Alicja Halbryt: Alone is far not enough.

75
00:13:47.130 --> 00:13:48.689
Alicja Halbryt: Next slide.

76
00:13:50.240 --> 00:14:05.499
Alicja Halbryt: So, as a business implementing or using AI within its operations, there's a large number of organization-specific risks as well, next to the societal risks that need to be managed.

77
00:14:05.700 --> 00:14:17.950
Alicja Halbryt: Here the ISO 42005 standard on AI system impact assessment tries to explain this correlation and says that

78
00:14:17.950 --> 00:14:27.529
Alicja Halbryt: A successful and complete AI impact assessment consists of considering both of these, both social and organizational aspects.

79
00:14:27.540 --> 00:14:32.699
Alicja Halbryt: So, yes, the next slide, puts it into a bit simpler words.

80
00:14:34.280 --> 00:14:45.909
Alicja Halbryt: if we think of a potential event or threat coming out of using the AI system, the first thing most people think of is the blue box.

81
00:14:46.370 --> 00:14:57.140
Alicja Halbryt: That is how it would impact people and society as a whole. Which is, of course, good, and which is the main concern, also, to regulators.

82
00:14:57.690 --> 00:15:05.570
Alicja Halbryt: What organizations need to consider in addition to that is the impact on their business, of course.

83
00:15:05.690 --> 00:15:21.760
Alicja Halbryt: And when Pep later talks about a case study we had at Nemko Digital, it will become clearer, I think, what an impact on an organization can be, and how to implement the risk management strategy.

84
00:15:22.410 --> 00:15:31.679
Alicja Halbryt: So if you take the impact on organization and on society together, you also need to assess the likelihood.

85
00:15:32.070 --> 00:15:35.070
Alicja Halbryt: Of these potential events occurring.

86
00:15:35.270 --> 00:15:50.510
Alicja Halbryt: And all this taken together can indicate to you the severity of the risk and suggest where to prioritize your efforts, which again will become more clear once we dive deeper into the case.

87

00:15:51.280 --> 00:15:53.780
Alicja Halbryt: So next slide…

88
00:15:54.870 --> 00:15:56.019
Pep Van Der Laan: Yeah, I think not… I think…

89
00:15:56.020 --> 00:15:56.500
Alicja Halbryt: It's…

90
00:15:56.500 --> 00:16:06.499
Pep Van Der Laan: for, for a little, for a little intermitto, a quiz, because as, as Aditya already mentioned, there's the two ways of looking into the.

91
00:16:06.500 --> 00:16:16.939
Pep Van Der Laan: into the risk of an AI system. Either you can look at the thing that's being regulated, so the impact at system level, on

92
00:16:17.130 --> 00:16:29.809
Pep Van Der Laan: society, on people. On the other hand, yeah, what does it do for your organization? So, very curious to hear from the audience here. Yeah, where are you focusing,

93
00:16:29.810 --> 00:16:37.650
Pep Van Der Laan: at this moment, your effort, is this almost entirely at the organization level, or are you mostly looking at the system level?

94
00:16:37.650 --> 00:16:51.580
Pep Van Der Laan: So, are you either focusing on, yeah, how does this affect your business, or are you focusing on how does this affect people and society, with a strong focus often on regulation?

95
00:17:00.920 --> 00:17:02.810
Pep Van Der Laan: So…

96
00:17:03.240 --> 00:17:06.579
Alicja Halbryt: Yep, so I think you can perhaps now, go into…

97
00:17:06.589 --> 00:17:07.379
Pep Van Der Laan: Yeah, so…

98
00:17:07.380 --> 00:17:08.030
Alicja Halbryt: Ace.

99
00:17:08.369 --> 00:17:17.189
Pep Van Der Laan: Awesome. Yeah, so with that, we'll come back to the answers in a bit, but let's first focus on…

100
00:17:17.209 --> 00:17:32.389
Pep Van Der Laan: on a case study to basically bring this concept, to life, and this is in a government, context. So let me explain, the case, that we worked on, recently.

101
00:17:32.389 --> 00:17:39.769
Pep Van Der Laan: to give you a flavor of, yeah, how does this… how does this play out for… for organizations? So…

102
00:17:40.759 --> 00:17:57.869
Pep Van Der Laan: this is a case where large organizations and large corporates can ask, basically, a government bully for a pre-assessment to understand, yeah, how do

laws and regulations actually apply to them in specific scenarios? So you can imagine that this is,

103
00:17:57.869 --> 00:18:07.779
Pep Van Der Laan: a process that's very heavy in terms of the dependence on jurisprudence, earlier,

104
00:18:08.249 --> 00:18:18.819
Pep Van Der Laan: earlier, statements in similar cases, as well as a lot of formal requirements that need to be met. So, heavily,

105
00:18:18.989 --> 00:18:24.149
Pep Van Der Laan: let's say, administrative and text,

106
00:18:24.329 --> 00:18:29.239
Pep Van Der Laan: Yeah, in-text, heavy, case, where…

107
00:18:29.269 --> 00:18:44.909
Pep Van Der Laan: Of course, this is typically something where, yeah, generative AI can provide a lot of value in those slow and time-consuming processes. So, in this case, the organization had built a solution to really

108
00:18:45.529 --> 00:18:53.079
Pep Van Der Laan: take the formal requirements of a case, make an assessment with AI already, kind of write a report.

109
00:18:53.079 --> 00:19:15.519
Pep Van Der Laan: not making the decisions, but just kind of summarizing what is there. And the second thing that I used it for was really looking, okay, in this case, where do

we find in the jurisprudence, in the regulations, in the supporting, documentation, the things that human assessors have to look at?

110
00:19:15.519 --> 00:19:16.749
Pep Van Der Laan: So that that whole…

111
00:19:16.749 --> 00:19:35.949
Pep Van Der Laan: evaluation process can be smoothened and accelerated. So, typical generative AI use cases, as we see them in many organizations, in this case, of course, within the government body, but also in many commercial organizations, we see similar use cases.

112
00:19:37.199 --> 00:19:54.599
Pep Van Der Laan: So, this started out with a very successful and promising pilot with enthusiastic frontline workers who really wanted that type of solutions, but there was not really a path to production for this solution.

113
00:19:54.599 --> 00:20:10.249
Pep Van Der Laan: And that had a lot of causes. So, first of all, there was insufficient AI experience across the organization to even identify and delimit the AI-related risk, and to understand, yeah, what would be the potential impact on those.

114
00:20:10.609 --> 00:20:27.189
Pep Van Der Laan: It was also a very risk-averse organization, where, yeah, decision makers were very hesitant to greenlight any novel applications, and in particular also when they had the AI label on them, irrespective of what the

115
00:20:27.189 --> 00:20:30.619
Pep Van Der Laan: AI solution, in fact, did and was.

116

00:20:32.059 --> 00:20:51.209
Pep Van Der Laan: also not a shared perspective on, yeah, how should you even treat those AI risks, and how should you judge them? So, missing formal ownership for those type of, for those type of decisions. Who actually is going to make the decision if something is good enough?

117
00:20:51.209 --> 00:20:53.649
Pep Van Der Laan: In this orgas… in your organization.

118
00:20:53.699 --> 00:21:04.329
Pep Van Der Laan: And all that led to basically stalling the go-live decisions, making… yeah, getting to a situation where the organization and its clients could not

119
00:21:04.859 --> 00:21:12.019
Pep Van Der Laan: yeah, basically profit from the benefits of generative AI in terms of, the supporting…

120
00:21:12.349 --> 00:21:24.099
Pep Van Der Laan: in, completing the process, speeding up the turnaround and operational efficiency. So, in that deadlock, yeah, there was really, yeah, a need to…

121
00:21:24.559 --> 00:21:38.609
Pep Van Der Laan: basically take steps and work together towards a solution. So that is where we stepped in to really help the organization towards getting ready for production. So.

122
00:21:39.299 --> 00:21:41.549
Pep Van Der Laan: What we did there was…

123
00:21:42.209 --> 00:21:58.529

Pep Van Der Laan: Of course, starting with the kickoff to really understand, yeah, where is the organization at this moment, diving into the documents, but then in the end, really shaping the solution, first in terms of, yeah, what does a good risk assessment look like?

124
00:21:58.569 --> 00:22:11.169
Pep Van Der Laan: So starting with that, that risk analysis, to also defining, yeah, how do you then put controls against that to make sure that you have your risk under control.

125
00:22:11.169 --> 00:22:22.609
Pep Van Der Laan: as well as, in the end, when it comes to completing that framework, yeah, how are you going to make sure that you operationalize that in an organization? How do you

126
00:22:22.609 --> 00:22:41.729
Pep Van Der Laan: make sure that people take the right responsibilities so that you can apply it to this use case, as well as many, many others. So, what we'll do in this, this, webinar is take you through a couple of the steps in that, in that process, hoping to give you a bit of

127
00:22:41.859 --> 00:22:49.449
Pep Van Der Laan: inspiration and, and understanding for how you can do that in your own, in your own organization as well.

128
00:22:50.589 --> 00:22:55.089
Pep Van Der Laan: So, first step there, understanding the risks,

129
00:22:55.389 --> 00:22:59.479
Pep Van Der Laan: What we find in many organizations is that, yeah, they…

130

00:22:59.669 --> 00:23:05.529
Pep Van Der Laan: don't have a complete picture of, yeah, what are AI risks? So, if you have

131
00:23:05.689 --> 00:23:21.939
Pep Van Der Laan: a risk assessment of AI, it often ends in the development team working from one stakeholder to another, and every stakeholder starts to pilot more risks that they see from their perspective on top of AI. So, that can be…

132
00:23:22.159 --> 00:23:38.179
Pep Van Der Laan: very time-consuming, that can be very demotivating for the development teams, and in the end, it doesn't lead to any constructive approach often. So, what we introduced in this organization was

133
00:23:38.189 --> 00:23:48.419
Pep Van Der Laan: a risk atlas, which we based on the… on the thinking of IBM, but adapted to the client-specific situation. So, the nice thing about,

134
00:23:48.559 --> 00:23:52.679
Pep Van Der Laan: AI-specific risk atlas is that you map out

135
00:23:52.839 --> 00:24:10.569
Pep Van Der Laan: all the risks that you… that are related to AI as a technology on a conceptual level even, so that you have a framework also for completeness, that you can really validate against that, yeah, did I look at all the aspects of the technology?

136
00:24:11.049 --> 00:24:28.709
Pep Van Der Laan: And based on that, we started to map the… map the risk. So, yeah, maybe interesting also to hear from you, Alicia, from… from your perspective, working with the… with the client. Yeah, how did they perceive this way of working? What's… yeah…

137
00:24:29.109 --> 00:24:31.649
Pep Van Der Laan: Yeah, what maybe surprised you also in that…

138
00:24:38.119 --> 00:24:41.479
Pep Van Der Laan: I… don't hear the… your sound.

139
00:24:42.449 --> 00:24:43.939
Pep Van Der Laan: Is that,

140
00:24:51.699 --> 00:24:55.179
Pep Van Der Laan: For me, the sound doesn't, doesn't work.

141
00:24:56.059 --> 00:25:05.009
Pep Van Der Laan: So, okay, yeah, so then… then I'll just continue, never mind. I hope that I'm the one who's audible in the end in the…

142
00:25:05.199 --> 00:25:18.379
Pep Van Der Laan: In the webinar as well. No sound for Alicia, I see. Okay, so, yeah, we'll just continue. We'll make it work, although it would have been nice to hear Alicia's perspective on this.

143
00:25:19.089 --> 00:25:24.859
Pep Van Der Laan: So… The first thing that's good to note is kind of at that big…

144
00:25:24.979 --> 00:25:29.059
Pep Van Der Laan: slide with all those risks, yeah, it can be quite…

145

00:25:29.159 --> 00:25:45.109
Pep Van Der Laan: overwhelming for people if they foresee it. But one of the reasons to start from this is to also say, look, these are the things that are applicable, and these are the things that you don't have to worry about. So in the cases that we just talked about, yeah, out of those.

146
00:25:45.109 --> 00:25:51.509
Pep Van Der Laan: 87, risks that are there on the, on the slide. Yeah, only,

147
00:25:51.629 --> 00:26:05.269
Pep Van Der Laan: a little bit less than half was actually relevant for those use cases. So, already bringing that into perspective based on what type of use case do you have, what is even relevant here.

148
00:26:05.269 --> 00:26:13.939
Pep Van Der Laan: Does help a lot in basically bringing down the complexity and making this manageable for the organization and for the teams involved.

149
00:26:15.649 --> 00:26:32.659
Pep Van Der Laan: So, next step after identifying those risks is basically making that… that assessment of the risk. So, how do we estimate the risk based on that, based on that specific use case? And this is also a good point to, to briefly pause,

150
00:26:32.769 --> 00:26:33.869
Pep Van Der Laan: Because…

151
00:26:34.129 --> 00:26:51.419
Pep Van Der Laan: What you often see in AI, in particular, if you have people who are less close to the topic, every risk seems gigantic, because people have heard about missiles, and have heard about all the scary things that can happen.

152
00:26:51.419 --> 00:26:59.519
Pep Van Der Laan: when it comes to cybersecurity, or have heard about AI taking over the world, you name it. But

153
00:26:59.549 --> 00:27:09.729
Pep Van Der Laan: In a lot of use cases, this is only a theoretical risk, and not something that's actually, in any case, tied to the way the technology

154
00:27:09.729 --> 00:27:26.849
Pep Van Der Laan: is being applied in this specific use case. So that is always, and it's here on the vertical axis, the first distinction that we try to make looking at this kind of risk assessment, yeah, bring it back to actually the case that we are working on.

155
00:27:26.849 --> 00:27:29.969
Pep Van Der Laan: And any societal concerns

156
00:27:29.969 --> 00:27:45.399
Pep Van Der Laan: about AI. Maybe very relevant, maybe very important, but maybe not for this case. So bringing it down in that dimension is super important. The second thing is that when you look at,

157
00:27:45.759 --> 00:27:50.389
Pep Van Der Laan: At the gross risk, of that use case.

158
00:27:51.239 --> 00:28:03.519
Pep Van Der Laan: you often see that people start thinking about, yeah, what could theoretically go wrong without taking into account the context of the organization. So…

159
00:28:04.089 --> 00:28:12.989

Pep Van Der Laan: Often, in many organizations, around data security, data privacy, there are already a lot of policies in place. There are

160
00:28:13.199 --> 00:28:19.949
Pep Van Der Laan: Yeah, there are already a lot of limitations to what can actually go wrong in terms of…

161
00:28:20.899 --> 00:28:27.389
Pep Van Der Laan: Is the system, hey, is the system, contained? Is the… is the data,

162
00:28:27.669 --> 00:28:41.269
Pep Van Der Laan: being cleaned, etc, etc. So that means that the gross risk, if you just look at the use case, is usually not the one that you want to look at when you do a risk evaluation. In that risk evaluation, you want also

163
00:28:41.339 --> 00:28:42.979
Pep Van Der Laan: To take into account.

164
00:28:42.999 --> 00:28:48.649
Pep Van Der Laan: the controls that are already there. So, in the end, the thing that you want to look at

165
00:28:48.649 --> 00:29:06.659
Pep Van Der Laan: In order to make it actionable and make it, make it operational, really the net risk based on the current scope of your use case and the future plans, features, of course, as well as based on the controls that you already have in your organization in place.

166
00:29:06.659 --> 00:29:24.649

Pep Van Der Laan: And only the remaining risk is the thing that still needs to be managed, because that is the point of risk management, to manage the remaining risks, not to do any theoretical exercises. So, in this organization, we saw that,

167
00:29:24.989 --> 00:29:28.339
Pep Van Der Laan: Yeah, there, if you map them out, those risks.

168
00:29:28.379 --> 00:29:35.629
Pep Van Der Laan: along the impact on the horizontal axis and the likelihood on the vertical axis. We saw that in the…

169
00:29:35.659 --> 00:29:50.959
Pep Van Der Laan: Yeah, in the gray area are the low risks, there's the light blue with the medium risks, and there's a darker blue with the… with the high risks, where, based on the risk appetite of the organization.

170
00:29:51.319 --> 00:30:05.189
Pep Van Der Laan: The focus was on, okay, those high risks are really the ones that we want to manage, that we want to bring down, at least to the light blue level, in order to make sure that we, bring only,

171
00:30:06.159 --> 00:30:14.349
Pep Van Der Laan: Use cases into production that are good enough, that are secure enough, that are reliable enough and trustworthy enough.

172
00:30:14.529 --> 00:30:19.399
Pep Van Der Laan: So bringing those ahead, so that already brings a lot of focus.

173
00:30:19.489 --> 00:30:36.859

Pep Van Der Laan: in terms of where you have to, to take action, right? Rather than looking at all the risks and, starting to run around, it's really focused on those high risks and think, how can we bring those risks, those risks down to a manageable level.

174
00:30:37.409 --> 00:30:39.539
Pep Van Der Laan: And how do you do that? Yeah, of course.

175
00:30:40.709 --> 00:30:52.829
Pep Van Der Laan: you need to define the controls against those risks, and that's in four pillars, so basically on the strategy level, when it comes to implementing governance, prevention, which is often

176
00:30:52.829 --> 00:31:01.409
Pep Van Der Laan: In the head design of the system, and how you manage and clean the ingoing data.

177
00:31:02.279 --> 00:31:13.779
Pep Van Der Laan: The detection, yeah, often the technical defense of the system, and correction, which is in… which is often the human oversight at the end.

178
00:31:14.009 --> 00:31:21.899
Pep Van Der Laan: So, in the end, we use this to build up a taxonomy of, of controls, so that you can really build out a kind of

179
00:31:22.019 --> 00:31:39.049
Pep Van Der Laan: Defense in-depth, strategy there, to basically put multiple, have multiple, controls against a single risk, and make sure that in the end you have a system where you can, can accept that, that risk,

180
00:31:39.259 --> 00:31:41.089

Pep Van Der Laan: Had a risk profile.

181
00:31:41.999 --> 00:31:42.979
Pep Van Der Laan: So…

182
00:31:44.419 --> 00:31:57.779
Pep Van Der Laan: what did all this bring us? Clarity is, I think, the key word there. So, we already saw, getting back from, yeah, from, the 80… the 87,

183
00:31:58.799 --> 00:32:17.879
Pep Van Der Laan: generally described risks to 40 that are available for the use case. We saw that we mapped out a lot of controls, over 150 controls that were identified as potentially helping in bringing down that risk level.

184
00:32:18.129 --> 00:32:31.289
Pep Van Der Laan: but also that only one-third of the risks was actually classified as significant enough to take additional measures. So, in the end, we identified 39 controls.

185
00:32:31.589 --> 00:32:50.209
Pep Van Der Laan: either new controls or updates of existing controls that needed additional effort, in order to make sure that the risk level of the, of the AI use case was brought back to the level that, yeah, that you could… was… yeah, that the organization was comfortable to bring it into production.

186
00:32:52.049 --> 00:33:08.279
Pep Van Der Laan: Because that is, in the end, the, also the objective, yeah, how do… yeah, bringing that, bringing that more into production. And in order to do that, you need to make sure that this kind of risk management is not only the, part of the…

187
00:33:08.519 --> 00:33:19.779

Pep Van Der Laan: onboarding and the building of the model. Now, it needs to basically also be continued in the deployment and the serving and monitoring

188
00:33:20.219 --> 00:33:38.659
Pep Van Der Laan: phases of the development flow, and that means that basically next to the AI development flow, you also need to have a strong AI governance flow, which means that you need to have the roles and responsibilities mapped out.

189
00:33:38.659 --> 00:33:47.039
Pep Van Der Laan: And the process is clear in order to, to integrate this type of risk management, in your, in your life cycle.

190
00:33:49.289 --> 00:34:04.659
Pep Van Der Laan: So, that's also what we see then here, in terms of how we made that explicit for this client. So, after the risk analysis and the control specification, really going into, okay, let's implement those controls.

191
00:34:04.749 --> 00:34:24.089
Pep Van Der Laan: do a proper risk acceptance, because always, when you do this type of risk mitigation, there's always a risk risk. That risk risk is not just something that's AI-specific, it's something that's generic for any type of risk management.

192
00:34:24.229 --> 00:34:35.229
Pep Van Der Laan: But that risk needs to be accepted. So you need to also create the ownership for both the implementation of the controls and for the acceptance of the risks.

193
00:34:35.369 --> 00:34:43.939
Pep Van Der Laan: Where the… Yeah, and in the end, that leads into a monitoring and feedback type situation. So, that…

194
00:34:44.069 --> 00:34:52.609
Pep Van Der Laan: Action plan is what we developed, so really mapping out those roles and responsibilities. Good to mention that for the individual controls.

195
00:34:53.169 --> 00:34:55.689
Pep Van Der Laan: It can be a very, yeah…

196
00:34:56.179 --> 00:35:14.019
Pep Van Der Laan: a differentiated group of people who are responsible for those, yeah, because there are technical model monitoring controls, which will be a responsibility of a tech lead, but also AI literacy of the frontline workers who work with the technology.

197
00:35:14.019 --> 00:35:22.959
Pep Van Der Laan: Is, in the end, part of the controls, and that will be the responsibility of the team lead on the business side.

198
00:35:23.529 --> 00:35:33.219
Pep Van Der Laan: So, accountability typically lies with the, yeah, with the owner of the, of the business, of the business process where this is, this is implemented.

199
00:35:34.299 --> 00:35:47.399
Pep Van Der Laan: And that is actually an interesting bridge into, yeah, where does the responsibility lie? And that is both responsibility within the process as responsibility for the norms that we apply. Because what you…

200
00:35:47.509 --> 00:35:57.959
Pep Van Der Laan: often see in organizations that are starting to implement AI at scale is that the responsibility for deciding when is good, good enough.

201

00:35:58.029 --> 00:36:13.569
Pep Van Der Laan: is not properly, defined yet. So this is an example of one of the decision flows that we developed for this client to also say, yeah, what do we do if we basically don't have

202
00:36:14.249 --> 00:36:31.349
Pep Van Der Laan: an owner for the… for a control who has the mandate to also decide when good is good enough. So, I'm not going to talk through this in detail. Maybe one thing to highlight, what you see in many organizations as a… as a solution to

203
00:36:31.939 --> 00:36:36.769
Pep Van Der Laan: speed up this type of processes when you don't have

204
00:36:37.179 --> 00:36:43.189
Pep Van Der Laan: Everything figured out yet, and when you have a lot of questions you still need to answer about,

205
00:36:43.489 --> 00:36:54.899
Pep Van Der Laan: what is actually good enough in our organization is to have an AI board that can, in the end, make decisions about AI technology, bringing together

206
00:36:54.899 --> 00:37:08.349
Pep Van Der Laan: basically a cross-functional set of people representing different aspects of the organization. A very important aspect of making sure that this is… and not just a technological,

207
00:37:08.579 --> 00:37:15.199
Pep Van Der Laan: Exercise, but really something that… Covers all the angles of the… of the organization.

208

00:37:16.739 --> 00:37:17.669
Pep Van Der Laan: So…

209
00:37:18.149 --> 00:37:30.929
Pep Van Der Laan: after doing that, where did we bring this organization? What was the value that we delivered through this use case? Yeah, first of all, for this use case, of course, the framework that we created

210
00:37:30.929 --> 00:37:40.769
Pep Van Der Laan: that clear path bringing that generative AI use case to production in a safe and controlled and responsible way, and they successfully did that.

211
00:37:40.819 --> 00:37:45.129
Pep Van Der Laan: But also, point two, yeah, this is…

212
00:37:45.369 --> 00:37:58.829
Pep Van Der Laan: not just something that you need to do, kind of, in a one-off way, this is something that then becomes the template for also applying this to other generative AI use cases. And you really see that that is also starting to,

213
00:37:58.829 --> 00:38:11.079
Pep Van Der Laan: And starting to gain traction in that organization as a role model and a template for, yeah, for doing this on a broader scale for many more use cases.

214
00:38:13.169 --> 00:38:19.479
Pep Van Der Laan: Great benefit of, yeah, doing this in a structured and,

215
00:38:19.759 --> 00:38:22.639
Pep Van Der Laan: Yeah, and, and also quite…

216
00:38:23.339 --> 00:38:40.589
Pep Van Der Laan: that intensive process with a lot of stakeholders is that it shifts the internal discussions within the organization, from only talking about abstract concerns and theoretical objections, to, okay, how do we actually make sure that we

217
00:38:41.099 --> 00:38:56.739
Pep Van Der Laan: understand what this… the impact of this system is really on a practical level, and what are the controls that we can put into place, so that in the end, you can work towards, really scaling… scaling your AI solutions.

218
00:38:56.859 --> 00:39:07.929
Pep Van Der Laan: And that is also, in the end, what's… yeah, the base… yeah, the basis is that we build with this, with this use case, really making sure that we have that

219
00:39:08.039 --> 00:39:11.809
Pep Van Der Laan: Foundation on which you can scale.

220
00:39:12.569 --> 00:39:29.829
Pep Van Der Laan: And that is scaling from both regulatory readiness as from a business readiness perspective. Because to round that off, if you look at the risks that we identified, the mitigations that we took, yeah.

221
00:39:30.019 --> 00:39:33.449
Pep Van Der Laan: I guess that in the end, maybe…

222
00:39:33.599 --> 00:39:44.329
Pep Van Der Laan: 20% was directly related to regulatory concerns. Most of the other concerns were about, okay, what do we want towards

223

00:39:44.559 --> 00:39:56.639
Pep Van Der Laan: the people that we serve as an organization? What is the risk for us as an organization? What are the, yeah, so maybe… so basically, kind of, this organization, I think, in the end.

224
00:39:57.969 --> 00:40:15.459
Pep Van Der Laan: when we had the discussions that mattered, I think it was around 80-20 in terms of 80% business, 20% regulation, where the value was. So, yeah, with that, Alicia, I don't know if you have the mic again?

225
00:40:18.489 --> 00:40:20.309
Pep Van Der Laan: I think the…

226
00:40:21.690 --> 00:40:22.080
Alicja Halbryt: Nope.

227
00:40:22.080 --> 00:40:24.409
Pep Van Der Laan: The old… Yes, there you are! Awesome.

228
00:40:24.860 --> 00:40:25.850
Alicja Halbryt: Perfect.

229
00:40:25.960 --> 00:40:29.690
Pep Van Der Laan: Yeah, perfect. So, with that, I wanted to,

230
00:40:30.330 --> 00:40:38.679
Pep Van Der Laan: Yeah, to basically, round, round off. Do you already have the results of the, of the quick, quiz that we did?

231

00:40:38.970 --> 00:40:47.239
Alicja Halbryt: Yes. So, 21% said they are almost entirely at organization level.

232
00:40:47.580 --> 00:40:54.080
Alicja Halbryt: Then most of the answers, we got, said mostly at organization level.

233
00:40:55.280 --> 00:41:07.919
Alicja Halbryt: And then, 5% mostly at system level, so not, not much, 11% almost entirely at system level, and then 26, balanced.

234
00:41:09.000 --> 00:41:17.369
Pep Van Der Laan: Yeah, so that's… that reflects quite neatly, I think, then, the balance that we also saw in this case study, I think.

235
00:41:17.370 --> 00:41:19.209
Alicja Halbryt: Yeah. Yeah, exactly.

236
00:41:19.210 --> 00:41:23.230
Pep Van Der Laan: Yeah. Nice, nice to hear, and

237
00:41:23.610 --> 00:41:35.330
Pep Van Der Laan: Yeah, with that, so I think the key, the key message is don't focus on regulation alone. It is often the thing that gets, gets a lot of the headlines. Also.

238
00:41:35.560 --> 00:41:43.069
Pep Van Der Laan: From, yeah, from a compliance side, but, the organizational lens is at least as important.

239
00:41:43.600 --> 00:41:48.300

Nemko Digital

Pep Van Der Laan: So, with that, shall we see if there are some questions in the…

240
00:41:48.850 --> 00:41:51.119
Pep Van Der Laan: In the chat already, maybe?

241
00:41:55.810 --> 00:41:57.630
Pep Van Der Laan: Let's look at that, huh?

242
00:42:01.870 --> 00:42:02.900
Pep Van Der Laan: I don't…

243
00:42:03.690 --> 00:42:12.259
Pep Van Der Laan: see, and yet, but maybe, kind of, that is a good, good opportunity, Alicia, now that we have you with voice,

244
00:42:12.280 --> 00:42:31.370
Pep Van Der Laan: already, kind of, to, to ask you a couple of questions also about how you experienced this, this collaboration, and what you saw in… in working with this organization on, on the risk, risk management. Maybe starting with the question, yeah, how, what was the early, yeah, what was basically the most.

245
00:42:32.810 --> 00:42:42.849
Pep Van Der Laan: Yeah, striking thing when you stepped into the organization, when you saw what is the current state, and how are people also, in their mindset towards, towards AI.

246
00:42:44.060 --> 00:42:58.689
Alicja Halbryt: Yeah, sure, so, I think one of the main things, or maybe the biggest benefit coming from this risk management approach and the IBM Risk Atlas as well, kind of showing this breakdown to people.

247
00:42:58.690 --> 00:43:06.560
Alicja Halbryt: It was definitely beneficial for them, for the organization, to see the amount of risks, which

248
00:43:06.640 --> 00:43:24.100
Alicja Halbryt: part of them they were not even aware, existed, so it was a real, like, a good reality check for them, for sure. But also, on the other hand, breakdowns like this, or overviews like that, reassure

249
00:43:24.160 --> 00:43:34.809
Alicja Halbryt: The client or the organization that the list is not infinite, and that you can take control over the risks.

250
00:43:34.990 --> 00:43:42.360
Alicja Halbryt: And another thing, that stood out to me in that project was… this…

251
00:43:42.550 --> 00:43:46.000
Alicja Halbryt: Huge necessity of being able to communicate

252
00:43:46.600 --> 00:44:05.450
Alicja Halbryt: the ethicists and the engineers, it was quite a challenge, that I saw, during, during, this project that, the, the ethicists present at the organization and the engineers, they, they had to be able to, kind of,

253
00:44:05.510 --> 00:44:16.240
Alicja Halbryt: You know, agree or disagree, but to have a conversation, and that sometimes is very difficult, for example, about data bias or data privacy.

254

00:44:16.320 --> 00:44:18.730
Alicja Halbryt: Yeah, so…

255
00:44:18.730 --> 00:44:30.330
Pep Van Der Laan: Yeah, I think that's a nice, and indeed kind of had those different, yeah, perspectives in the organization. I think that the third one to mention, there's also the, the people

256
00:44:30.410 --> 00:44:36.770
Pep Van Der Laan: on the business side, who were also very eager then to kind of, yeah, just use the technology, right? Because…

257
00:44:36.770 --> 00:44:55.620
Pep Van Der Laan: as soon as you see the benefit, it's also very, yeah, very tempting to say just, what can go wrong? Just, let's use it. But on the other hand, then, kind of, yeah, you have the balance from more the ethics and compliance people, and you have the, yeah, the engineers who, in the end, have to make it work.

258
00:44:56.730 --> 00:45:07.100
Pep Van Der Laan: But also cannot solve everything in technology, and then need, again, the business people to also weigh in in making

259
00:45:07.510 --> 00:45:22.480
Pep Van Der Laan: So, yeah, I think that's a nice, nice perspective. Yeah, maybe also your perspective on the, on the different controls, yeah, that we, that we identified. If you look at those controls, and yeah, what's…

260
00:45:23.310 --> 00:45:43.189
Pep Van Der Laan: Yeah, was it the controls that we identified? Was that a bit in line with what you had expected? Were there things where you say, well, this was in this

organization's particularly kind of an interesting lens of what was needed there in order to become ready for production?

261
00:45:44.750 --> 00:46:03.970
Alicja Halbryt: I think what might have been surprising is that, given that this was a governmental institution, you might think that, privacy would be the biggest issue, and controls connected to that, should be, kind of prioritized.

262
00:46:04.090 --> 00:46:08.009
Alicja Halbryt: But actually what turned out, was that the…

263
00:46:08.140 --> 00:46:17.570
Alicja Halbryt: The risks coming from governance and value alignment were the ones that had to be focused on the most, because that…

264
00:46:17.650 --> 00:46:31.490
Alicja Halbryt: The likelihood and impact of these, would have the… well, what was, was, kind of graded as, as the highest and as the most, significant and burning, to control.

265
00:46:31.710 --> 00:46:40.139
Alicja Halbryt: So I think, yes, this was, I think the most, kind of, surprising and, like.

266
00:46:40.140 --> 00:47:03.050
Alicja Halbryt: an interesting, learning point, that privacy is not always, the, the most important thing. If you're a governmental institution or a public service institution, then, actually, yeah, value alignment, being sure, that transparency is, you know, intact, and, that you do the correct testing.

267
00:47:03.050 --> 00:47:10.550

Alicja Halbryt: That these, sort of things, are, kind of in control, and also…

268
00:47:10.960 --> 00:47:15.919
Alicja Halbryt: because this was a generative AI case,

269
00:47:16.480 --> 00:47:22.710
Alicja Halbryt: A problem that the client also could see was the over- and under-reliance.

270
00:47:22.890 --> 00:47:31.370
Alicja Halbryt: On the output of the generative AI, and that was, I think, quite interesting as well.

271
00:47:31.760 --> 00:47:34.000
Alicja Halbryt: And it needed to be controlled.

272
00:47:34.440 --> 00:47:35.370
Alicja Halbryt: Yeah.

273
00:47:35.370 --> 00:47:52.270
Pep Van Der Laan: Nice. So, yeah, I think, kind of, yeah, maybe it, we, we slowly go towards the, yeah, towards the end of the, of the webinar, so let me see if I can, can one more time, share the screen for, all of you.

274
00:47:54.180 --> 00:47:56.919
Pep Van Der Laan: There, so basically,

275
00:47:57.940 --> 00:48:12.240

Pep Van Der Laan: to offer you the opportunity, if based on this conversation, you think, yeah, this is something that I need in my organization, how do we approach this? Yeah, please, call, call our AI experts.

276
00:48:12.250 --> 00:48:20.149
Pep Van Der Laan: There's a link, here with, with a QR code, so, feel free and,

277
00:48:20.290 --> 00:48:28.960
Pep Van Der Laan: complete the application form, and, hope to speak you… speak to you soon. As well as, then.

278
00:48:29.310 --> 00:48:42.650
Pep Van Der Laan: on the next slide, if you want to stay updated on what we are doing, please follow us, particularly on LinkedIn, we are quite, quite active, and always feel free to reach out.

279
00:48:43.740 --> 00:48:48.130
Pep Van Der Laan: With that, thanks a lot for the attention, and see you in the next webinar.

280
00:48:48.670 --> 00:48:49.910
Alicja Halbryt: Thanks, everyone.