Nemko Digital Webinar Report - EU AI Act Update Webinar Transcript

1
00:00:02.009 --> 00:00:28.650
monica.fernandez@nemko.com: Welcome everybody. My name is Monica Fernandez, and I am here today to introduce you to the new obligations from the EU AI act that just came into force. These obligations are around general purpose. AI. And what providers have to do in order to to comply with the regulation in order to put these models into the market into the European market

2
00:00:29.210 --> 00:00:47.649
monica.fernandez@nemko.com: without further ado. We'll jump right in here. We have a summary of what the EU AI act framework is how it is structured. We can see that in that the EU AI act uses a risk-based approach, where, at the top of

3
00:00:47.750 --> 00:00:58.840
monica.fernandez@nemko.com: the pyramid, we could call it. We have prohibited AI systems that are completely prohibited from being placed in the European market in the EU market.

4
00:00:58.840 --> 00:01:21.639
monica.fernandez@nemko.com: Then we have high risk AI systems, which is what most of the regulation is about. There's a quite lengthy set of requirements that providers, deployers, importers, and distributors of high risk. AI systems must comply with. In order to put these systems in the market, then we have limited risk. AI systems which

5
00:01:22.158 --> 00:01:44.459
monica.fernandez@nemko.com: have to adhere to a more limited set of requirements, and they're mostly around transparency, low risk AI system fall completely outside the scope of the AI ad, and then general purpose. AI general purpose. AI models are normally around are

6
00:01:44.945 --> 00:01:57.569

monica.fernandez@nemko.com: around, or are subject to general purpose. AI. Specific requirements that providers have to adhere to in order to to place these models in the market.

7
00:01:59.740 --> 00:02:16.259
monica.fernandez@nemko.com: So without further ado, we can jump straight into the definitions of general purpose. AI. There's a simplified version of what the EU AI act defines a general purpose. AI model, as the Act describes it as models

8
00:02:16.260 --> 00:02:40.489
monica.fernandez@nemko.com: trained on large scale, data that can completely perform many distinct tasks, and can be integrated into various applications. Now, as we know, there are a set of rules for general purpose AI models, but these rules do not apply for models that are used for research that are in development or prototyping activities before they are placed

9
00:02:40.820 --> 00:03:05.190
monica.fernandez@nemko.com: in the market. So these set of models are excluded from the scope and indicative threshold for defining or for identifying models of general purpose. AI model are those that use a training compute greater than 10 to the power of 23 floating operation points

10
00:03:05.430 --> 00:03:10.400
monica.fernandez@nemko.com: and can generate text audio images or video.

11
00:03:10.660 --> 00:03:37.609
monica.fernandez@nemko.com: Now, an important thing to to understand from the AI act is that they also separately define general purpose. AI models with systemic risk, and providers of these models have to adhere to additional requirements on top of the requirements that we already have for general purpose. AI models. Now, how do we identify those with systemic risk?

12
00:03:37.610 --> 00:03:44.855
monica.fernandez@nemko.com: The act defines those as models that more more

13
00:03:46.130 --> 00:03:59.130

monica.fernandez@nemko.com: they're very likely to to require training, compute greater than 10 to the power of 25 flop, and they are likely to significantly affect the EU market.

14
00:03:59.470 --> 00:04:25.219
monica.fernandez@nemko.com: aside from providers having to identify for themselves that whether a general purpose AI model has a systemic risk, and they have to also notify the commission of it. The Commission may also designate general Purpose AI models as having systemic risk based on criteria that we can find in annexed

15
00:04:25.530 --> 00:04:27.719
monica.fernandez@nemko.com: 13 of the AI act.

16
00:04:30.370 --> 00:04:30.985
monica.fernandez@nemko.com: So

17
00:04:31.730 --> 00:04:41.272
monica.fernandez@nemko.com: the main question that people have is sometimes they're not sure if their model is a general purpose. AI model. So we'll go through some examples.

18
00:04:41.780 --> 00:05:00.860
monica.fernandez@nemko.com: to see if if these examples are or are not considered, general purpose. AI models under the AI act. So here we have an example of a model that is trained on natural language data, using 10 to the power of 22 flaw, and cannot perform many different tasks well.

19
00:05:00.860 --> 00:05:22.620
monica.fernandez@nemko.com: So based on the definitions that we just recently reviewed. We can say that because its training compute is below 10 to the power of 23 flops, and it lacks the ability to handle a wide range of tasks. We can consider this model as not a general purpose. AI model.

20
00:05:24.840 --> 00:05:49.260

monica.fernandez@nemko.com: Another example that we could use is this one a model that is trained using 10 to the power of 24 flop and a large diverse set of natural language data from Internet and other sources. Now, according to the definitions and the guidelines that the European Commission has provided for us recently.

21
00:05:49.260 --> 00:06:04.850
monica.fernandez@nemko.com: we can say that it, it is likely considered a general purpose. AI model, because it exceeds the 10 to the power, 23 flop threshold training threshold, and it can generate text and handle many different tasks

22
00:06:04.960 --> 00:06:07.490
monica.fernandez@nemko.com: due to its broad training data.

23
00:06:09.980 --> 00:06:29.189
monica.fernandez@nemko.com: We'll go through one last example. And this is an example where we have a model that is trained with 10 to the power, 24 flop. To transcribe speech to text. Now, while this model exists, 10 to the power, 23, flop threshold, and can generate text

24
00:06:29.190 --> 00:06:42.210
monica.fernandez@nemko.com: it only performs. And now test. So in this case, it's it's speech to text transcription. So according to the guidelines, it is not considered a general purpose. AI model.

25
00:06:45.640 --> 00:07:10.719
monica.fernandez@nemko.com: So what are the obligations? Exactly so. People want to know what the obligations are? After they they identify whether they're dealing with a general purpose. AI model or not, and and this is it so? On the left hand side we can see that there are 3 main obligations for general purpose. AI model providers. So these are aimed for the providers.

26
00:07:11.070 --> 00:07:21.470
monica.fernandez@nemko.com: and the 1st one is around transparency. So providers have to provide transparency over the content used to train such models.

27

00:07:21.540 --> 00:07:42.640
monica.fernandez@nemko.com: Second, they need to implement a policy to comply with EU copyright law, and, 3, rd they need to produce and maintain technical documentation for 2 purposes, one to be able to share it with the AI office and national authorities if they are requested, and 2 to make it available to downstream providers.

28
00:07:43.090 --> 00:08:01.319
monica.fernandez@nemko.com: Then, on the right hand side, we have the obligations for general purpose. AI models with systemic risk, and bear in mind these obligations are on top of the obligations that we have on the left hand side here. So these are additional obligations that providers of

29
00:08:01.380 --> 00:08:27.880
monica.fernandez@nemko.com: of these models have to adhere to in this case the obligations around performing model evaluation and adversarial testing to identify and mitigate the systemic risks that have been identified, or, as the second point here already says, we have to identify and address these risks that we find to be at the EU level.

30
00:08:27.950 --> 00:08:41.099
monica.fernandez@nemko.com: 3, rd they have providers of these models have to document and report serious incidents and their corrective actions to the AI office and national authorities without delay.

31
00:08:41.570 --> 00:08:52.340
monica.fernandez@nemko.com: and last, but not least, providers must ensure adequate cybersecurity protection to both the model and its supporting infrastructure.

32
00:08:56.700 --> 00:09:19.290
monica.fernandez@nemko.com: Then we have to talk about downstream modifiers as provider, as providers of general purpose. AI models. So as a lot of you may already know a lot of the times. We don't develop general purpose AI models from scratch, but we modify existing ones, or we use existing ones.

33
00:09:19.290 --> 00:09:41.860

monica.fernandez@nemko.com: And we have to understand the concept of downstream modifiers, because this has implications on whether we ourselves, as modifiers can become the provider of such models. So in very simple terms, downstream modifiers are actors who modify or fine tune an existing model, an existing general purpose. AI models.

34
00:09:41.860 --> 00:09:57.690
monica.fernandez@nemko.com: and, like I just said, they can become providers of such, mainly because these modifications can greatly make changes to the model's generality, capabilities and or its systemic risk.

35
00:09:58.210 --> 00:10:21.160
monica.fernandez@nemko.com: So we've we have known from the beginning that this is a thing that you could become a provider of the general purpose. AI model if if you fine tune the model. But until recently it wasn't too clear. What does what is this level of fine tuning that is needed in order to to become a that will make you a provider. So

36
00:10:21.160 --> 00:10:45.250
monica.fernandez@nemko.com: recently the the code of conduct, the guidelines that have been released around this topic has have given us more insight on this. And what they're saying is that if the fine tuning of these models, or the modification of these models uses more than 1 3rd of the compute

37
00:10:45.250 --> 00:10:59.009
monica.fernandez@nemko.com: that was used to train the original model, then that is considered a significant change, and would therefore mean that you are becoming the provider of the new model that you are.

38
00:10:59.700 --> 00:11:27.600
monica.fernandez@nemko.com: that you're making, that you're providing. And the reason for this, the reason for putting such threshold that it's just 1 3rd of the original compute is to make proportional to the size of the model, of course, and to make sure that it's fair across different model sizes, and also, most importantly, to to avoid the discouraging of improving smaller models.

39

00:11:28.217 --> 00:11:36.249
monica.fernandez@nemko.com: So yeah, however, it is not always known what was the the

40
00:11:36.630 --> 00:11:52.820
monica.fernandez@nemko.com: level of compute that was used for the original model. Sometimes we may not have information about that, or it's hard to get such information. So when it is impossible to get such information. The European Commission suggests that

41
00:11:53.270 --> 00:12:18.439
monica.fernandez@nemko.com: now we need to substitute this threshold by applying 1 3rd of the original threshold that we use to define a general purpose, AI model or a general purpose, AI model of systemic risk. So in this case, if you use 1 3rd of 10 to the power of 25 flop, which is the threshold that we use for systemic risk models.

42
00:12:18.500 --> 00:12:41.949
monica.fernandez@nemko.com: Then you are becoming a provider of a new general purpose. AI model with systemic risk. If you use 1 3rd of 10 to the power of 23 flop, which is the threshold with use for regular general purpose. AI models. Then you're becoming the provider of a new, regular, general purpose. AI model.

43
00:12:45.440 --> 00:13:15.129
monica.fernandez@nemko.com: So why does this matter? Why, why do we even have to identify new providers, or what? Why does someone have to become a new provider? If it makes significant changes? Well, as we can imagine. Some of you can already imagine these significant modifications can impact the transparency, copyright and systemic risk obligations that the original provider is adhering to. So

44
00:13:15.410 --> 00:13:39.680
monica.fernandez@nemko.com: yeah, but more more than anything, it's. It's also a forward looking approach. So, to be honest nowadays, it is kind of hard to to that. You will fall under under this specific provision, where you will use 1 3rd of the computational power to modify a model. Currently, this.

45
00:13:39.680 --> 00:14:02.419

monica.fernandez@nemko.com: like the modifications that are done to these models are not large enough to meet the threshold that is defined, but this may change over time, and as compute power becomes more available more downstream actors may qualify as as providers of as new providers of such models.

46
00:14:02.630 --> 00:14:25.359
monica.fernandez@nemko.com: And then, at the same time, the Commission is trying to or may update the rules over time to reflect changes in in the technology and and the market. So using these these thresholds are can be easily changed over time. As as things move forward

47
00:14:25.590 --> 00:14:26.820
monica.fernandez@nemko.com: with technology

48
00:14:28.730 --> 00:14:42.889
monica.fernandez@nemko.com: now, we cannot think of general purpose. AI models on their own. A lot of these models are often integrated into bigger systems. And we and these systems are integrated

49
00:14:42.890 --> 00:15:07.539
monica.fernandez@nemko.com: or implemented into a certain sector or domain with a certain intended use. So often, we need to not only think about the obligations that we have to adhere to from a general purpose, AI model perspective, but also from the perspective of the larger AI system that it is integrated into. So, for example, if we have a model like Gpt-four

50
00:15:07.540 --> 00:15:18.819
monica.fernandez@nemko.com: that you is used to summarize clinical guidelines. And this model is integrated or embedded into a system

51
00:15:19.140 --> 00:15:43.930
monica.fernandez@nemko.com: that is used in healthcare, and the system is a diagnostic decision support system for doctors, then we would consider the overall system as high risk. And what this would mean is that the provider, as such system would have to adhere to the full high risk, AI system, obligations, and the general purpose AI component

52
00:15:44.407 --> 00:15:56.809
monica.fernandez@nemko.com: or the general purpose AI. Obligations depending on whether this general purpose AI is a 3rd party model or not. So we need to think about.

53
00:15:56.810 --> 00:15:57.490
monica.fernandez@nemko.com: but

54
00:15:58.170 --> 00:16:10.139
monica.fernandez@nemko.com: about the entire environment, the environment in which the general purpose AI model is used in and how it is integrated into into systems. And

55
00:16:10.390 --> 00:16:11.090
monica.fernandez@nemko.com: so

56
00:16:13.350 --> 00:16:39.029
monica.fernandez@nemko.com: now important, the timeline for general purpose, AI model obligations. So we are releasing this news flash today because we have just gone past the the deadline of the second of August of of this year of 2025, where it is the date when obligations of general around general purpose. AI. Models have come into force.

57
00:16:39.030 --> 00:17:00.369
monica.fernandez@nemko.com: Now, what does this mean? There are other things to consider. 1st of all, these obligations apply to general purpose AI models that are introduced in the market after this date, so new general purpose, AI. Models that are coming into the market after the second of August, of 2025.

58
00:17:00.370 --> 00:17:10.120
monica.fernandez@nemko.com: Now there is a grace period for all the models that have been introduced before this date that ends

59
00:17:10.329 --> 00:17:13.579
monica.fernandez@nemko.com: on the second of August of 2027.

60
00:17:14.250 --> 00:17:25.290
monica.fernandez@nemko.com: Now, what does this mean for these models that have been introduced before today's dates before the second of August. It means they don't

61
00:17:25.290 --> 00:17:49.329
monica.fernandez@nemko.com: necessarily required to retrain or learn or unlearn the content. If if the retraining is is really impossible. It's practically impossible. And if the training data is unavailable or too burdensome to retrieve so yeah, you don't always have to retrain such models.

62
00:17:50.250 --> 00:17:52.350
monica.fernandez@nemko.com: But these limits.

63
00:17:52.550 --> 00:18:03.179
monica.fernandez@nemko.com: If, if, because of these 2 reasons, you cannot retrain the model, you must clearly explain this, explain these limitations in the

64
00:18:07.920 --> 00:18:16.460
monica.fernandez@nemko.com: either of these models. This is a very important thing to understand as it affects all. No matter

65
00:18:17.410 --> 00:18:21.210
monica.fernandez@nemko.com: if if your model was released before this date.

66
00:18:21.530 --> 00:18:46.030
monica.fernandez@nemko.com: now for new models. So maybe at this point in time that might be providers who have trained, or in the process of training or planning and on training a model with the view of placing it in the market. After the second of August of 2025, these providers must proactively engage with the AI office if they're facing any compliance difficulties.

67
00:18:46.030 --> 00:18:51.960

monica.fernandez@nemko.com: And if these models are identified to be

68
00:18:52.100 --> 00:19:09.460
monica.fernandez@nemko.com: of those that have systemic risk, they still have to notify the Commission within 2 weeks. So this is also important. You must notify the Commission within 2 weeks after the second of August. If you are releasing a model with systemic risk.

69
00:19:15.960 --> 00:19:41.180
monica.fernandez@nemko.com: Now I have been mentioning the general purpose AI code of practice here and there. But what is it? Exactly so. This code is of practice was published on July 10th of this year. It is a voluntary guidance framework, so it is voluntary, and this has been developed by over a thousand independent experts who have been working really hard in order to make this

70
00:19:41.180 --> 00:19:51.499
monica.fernandez@nemko.com: feasible, intagable, and easy for for providers to to to read and implement in order to

71
00:19:51.500 --> 00:20:06.980
monica.fernandez@nemko.com: to in practice adhere to the general purpose. AI obligations of the AI act. So yeah, like, I said, it supports a responsible implementation of the AI act around these models.

72
00:20:07.360 --> 00:20:37.099
monica.fernandez@nemko.com: The main thing to understand about this code of practice that you can find already online is that it includes 3 main chapters, which are the 3 main, like the main obligations, are around these models, and the chapters are on transparency and on copyright, and then a chapter on Safety and security that are only for models with systemic risk.

73
00:20:37.750 --> 00:21:03.120
monica.fernandez@nemko.com: However, it's not official, it hasn't been reviewed. This. This code is now currently under legal review by the AI office and the AI board, and what this means is that once it is approved, if, if it is approved, the Commission may issue an implementing act to give the code, general legal validity across the EU.

74
00:21:03.120 --> 00:21:17.889
monica.fernandez@nemko.com: Now, what does that mean? What that means is that we can use this code of practice to basically demonstrate compliance with this section of the AI, the chapter around general purpose. AI, so

75
00:21:18.438 --> 00:21:39.439
monica.fernandez@nemko.com: yeah. If if and once it is approved, providers can can show that they meet the obligations by following the code of practice. And what this means is that if you opt out of parts of the code, this will remove such benefits of of demonstrating the compliance in in those areas.

76
00:21:39.440 --> 00:21:53.929
monica.fernandez@nemko.com: And, like, I said, this is this code of practice at the end of the day. It is voluntary you don't have to use it. You can use other adequate methods, but it is quite a simple and straightforward path

77
00:21:53.930 --> 00:22:00.539
monica.fernandez@nemko.com: to to ensure that you do comply with the obligations of the AI. Act around general purpose. AI.

78
00:22:00.920 --> 00:22:21.579
monica.fernandez@nemko.com: And if you don't, noncompliance, as we have mentioned in other webinars and other sessions. Noncompliance with the general purpose. AI model obligations means that companies could be facing up to 15 million or up to 3% of global annual turnover in fights.

79
00:22:22.100 --> 00:22:22.760
monica.fernandez@nemko.com: Okay?

80
00:22:23.290 --> 00:22:47.200
monica.fernandez@nemko.com: And that is it. We just wanted to provide a quick update on what is going on. But if you want to keep track of more webinars or news updates. Please feel

free to follow us on Linkedin. We keep posting things on there to make sure that everybody is up to date with the news on regulations

81
00:22:47.200 --> 00:22:55.970
monica.fernandez@nemko.com: around the world, and we have a special focus on the AI act these days. So so just follow us if you haven't already.

82
00:22:56.750 --> 00:23:26.500
monica.fernandez@nemko.com: and if you have some questions, if you are part of a company and you were thinking of, you have a product or general purpose model in mind or or you have a use case in mind that you were thinking about while you were watching this webinar, and you have some questions, or simply want to discuss it for a bit feel free to to book a 15 min consultation with us. It's completely for free, of course, and you

83
00:23:26.500 --> 00:23:40.050
monica.fernandez@nemko.com: we can. We can discuss it. We can have a 1 to one discussion about your use case. And we can guide you on. What would be the most appropriate next steps for your for your use case.

84
00:23:40.650 --> 00:23:51.519
monica.fernandez@nemko.com: Right? Thank you for being here today for listening in, and I hope everybody's having a wonderful, wonderful summer holiday or summer period.

85
00:23:52.790 --> 00:23:53.819
monica.fernandez@nemko.com: Thank you.

86
00:23:54.610 --> 00:23:55.310
monica.fernandez@nemko.com: Thank you.