Bird & Bird

The political agreement on the AI Act - Part 2

Podcast transcript – Feyo Sickinghe (Of Counsel, Netherlands), Dr. Simon Hembt (Associate, Germany) & Oliver Belitz (Associate, Germany)

March 2024

Feyo Sickinghe:

Welcome to Part 2 of Bird & Bird's podcast on key insights from the political agreement on the AI Act. I am Feyo Sickinghe your host and I am joined by my colleagues Simon Hembt and Oliver Belitz, both AI legal and regulatory specialists. Many thanks for being with us here. In this episode, we will discuss general-purpose AI, biometric identification, and open-source models. To start off, Oliver can you explain how general-purpose AI will be regulated?

Oliver Belitz:

Yes, of course. So general-purpose AI models were one of the most discussed topics in the legislative process. Allow me to spend a few sentences on the context here. When the first draft of the AI Act was published back in April 2021, the AI landscape was a different one. We did not know about generative AI, and the AI systems we talked about back then were predictive or analytical AI. Fast forward somewhere in November 2022 ChatGPT entered the stage and we started talking about generative AI. So, AI that produces content for us, texts, pictures, or video, was not foreseen in the first draft of the AI Act. Back then, we thought about AI systems that were built for single purpose. To predict, for example fraudulent payments, and now we have to deal with a new sort of AI that was built for several purposes. We can do a lot of things with ChatGPT, and OpenAI did not foresee all those things that we can do with the tool. And so, the law maker had to wrap his head around what we will do with general-purpose AI models because they did not fit the risk base approach. They introduced first a new category that was called "foundation models" and back in December it was changed to "general-purpose AI" and, as I already mentioned, it was one of the biggest discussion points in the negotiations.

So, where did those discussions end up? Currently, we have five articles in the AI Act for general-purpose AI. Just a quick comparison, for high-risk AI systems, that we talked about last time, those are 45 articles, so it is a lot less. What GP AI is? I have already talked about it; it is an AI system that is not built for a single purpose but rather a wide array of purposes. As for example, the GPT model from OpenAI or Google's Gemini or Meta's Llama, those are all GP AI models. Those fall in two categories under the current AI Act. The first one is what I would call basic GP AI obligations and the second one, I think we will come to that one in a minute, is GP AI with systemic risk.

So, we will start with the first one. If you fall under that category, you have certain obligations that are addressed at you as a provider and this is, for example, to have technical documentation for your GP AI model to provide downstream information for other providers that will use your model. For example, if you put GPT on the market then other providers can take your model and build a new tool with GPT at its core and then you have to provide information to those providers. And then, two additional obligations, and I think Simon will touch on them, the first one is you have to have a copyright policy and you have to provide a summary of the training data.

Feyo Sickinghe:

So, talking about copyright Simon it's actually very interesting. The results of what the outcome of the data is, there are questions as to whether those could be copyrighted, and how to deal with any data that has been copyrighted for use as an input for the model. So, two sides of the discussion, maybe you could inform us a bit more about this Simon?

Dr. Simon Hembt:

Yes, thanks Feyo. The discussions about the impact of general AI and copyright is huge, of course. We all know the discussion about whether AI generated output is copyright protected or not. But also, the AI Act has some provisions for GP AI here and Oliver mentioned these two.

The first one, and I think it is quite an interesting one, is to put in place a policy to respect union copyright law and here we have one example. One example is that the providers of GP AI models have to indemnify and respect reservations of rights expressed by the right holders. What does it mean? If you think of a person who is uploading some pictures on his or her own website and now is declaring a reservation and saying, "*well I am not consenting into the use of these pictures for training purposes*". So now, the providers of GP AI are obliged to indemnify and respect these reservations here. But we will see in the future how the construction of this provision will look like, because this provision can open up a door to more intense content moderation rules as well. Like the notice and action mechanism which must be in place to make sure that infringing content, which is generated by these models, can be reported to the system, and potentially be taken down if it is posted on a publicly available server.

Feyo Sickinghe:

It all remains a bit unclear I would say at this stage Simon. How can a provider of the foundation model know that this specific picture in your example is limited for use as an input in AI? How can he know?

Dr. Simon Hembt:

I think the only option could be that the right holders as the copyright direct of claims put in these reservations in a machine-readable format so that the AI providers rely on crawlers who can identify this.

Feyo Sickinghe:

It is a whole new terrain becoming here. A reference database in which you need to match the data you have generated through scraping, and then filter out the data that will be protected by copyright. That is quite something. We are entering open terrain, I guess.

Dr. Simon Hembt:

Yes, we have some examples in the platform regulation like Audible Magic or for any other tools that allows us to filter content online, but when speaking about texts or pictures there is no such big database, yes sure.

Feyo Sickinghe:

That is on the input side but on the output side if you create something to work with the use of generative AI there is still a question as to whether that could be copyrighted. In the US it is pretty clear. How is that in Europe?

Dr. Simon Hembt:

Yeah, we got some decisions in the US for sure. They are saying even if you are creating something with 600 prompts it is still not capable for copyright protection. Here, for instance in general we have no case law on that - on output which is generated with AI. But we have some guiding principles which states that every work has to be produced by a human author for instance. Or at least the human author can use technology to produce something but, in that case, the human author has to use it like a tool, or a craft and here using for instance the text to text, or text to image generators the tool is kind of dominating the process. So, here it is rather difficult to get a copyright protection for the work here.

Feyo Sickinghe:

So, in essence that means that if the tool is the dominant creator there will be no means for copyright, but if the human person is the dominant creator which uses AI for a part it might still be eligible for copyright. I think that the template is something that needs to boil down into test cases that will be brought for the courts.

A very interesting discussion, a bit of a sidestep in the whole thing but very important because it is very open and good for everybody to be aware. Back to you Oliver to guide us through the part of the regulation which deals with general-purpose AI and systemic risks. How does that work? Is that a bit similar to being a very large online platform or a gatekeeper under the Digital Markets Act?

Oliver Belitz:

Yes, to some extent. We have metrics here. First, the definition when you have a systemic risk in your GP AI it is rather abstract. So, the AI Act says, your model has to have high impact capabilities. So that is rather broad. Then in the next section under that article we have a presumption. There, it says your model is presumed to have high impact capabilities and therefore be a GP AI model with systemic risk, if the compute power used to train that model is above 10 ^ 25 FLOPS. FLOPS is a metric for compute power, it is Floating Point Operations Per Second. To make it short, that is a lot of compute power that you need to put into your model to be above that threshold. Currently that would only aim at the biggest models that are available. For example, OpenAI's GPT-4 or Google's Gemini, those would be above the threshold, especially all of the European models would currently be below that threshold.

Feyo Sickinghe:

ChatGPT at 3.5 was trained on the 10 x 24 so would be just below the threshold, anything above, Chat GPT-4 would be in?

Oliver Belitz:

That is right yes, and maybe one additional remark here - so we have this rather technical threshold in the AI Act right now, but the commission can adopt delegated acts to amend that threshold. It could increase the FLOP threshold for example, or it could introduce a completely new threshold that would not be based on the compute power but, for example on energy consumption, or efficiency, or training data, or number of parameters of the model. So, this is kind of future proof as the commission can amend it according to the current technical landscape.

Feyo Sickinghe:

I would say that is quite remarkable that such an important threshold is in the regulation itself and then can be changed by the commission to a delegated Act.

Oliver Belitz:

Yes.

Feyo Sickinghe:

Which gives commission huge discretionary powers in how to deal with systemic risks without any democratic control behind it whether that will be justified.

Oliver Belitz:

Good point. Yes. It is exactly that and maybe to close the point on GP AI with systemic risk, I have not talked about the obligations that come with this categorisation. On top of the basic obligations that I have already talked about, as a provider you have to do some model evaluation, you have to assess and mitigate the risk that comes with your model, you have a reporting obligation for incidents that happen in your model, and you have to maintain a certain level of cyber security measures so that comes on top.

Feyo Sickinghe:

Then may the commission add other obligations to it as well?

Oliver Belitz:

Yes, that is right, so the commission can adapt those obligations.

Feyo Sickinghe:

So, it would be interesting to take a look at the way the commission approaches systemic risk in the context of the AI Act. Systemic risks for very large online platforms are a digital service act obligation, where there is a presumption of systemic risks when you are very large and have more than 25 million users in the EU. Clearly this is not the approach that the commission has taken for this, instead they measure systemic risk on the threshold of computer power, but that still may change, and it will be very interesting how that develops in the future - how those two systemic risk approaches would relate to each other. Thank you for that Oliver. Let's talk about biometric identification and categorisation and how that is being dealt with Oliver.

Oliver Belitz:

Okay so, we have a lot of provisions in the AI that deal with biometric identification. I will try to break it down and make it rather short. First, we have to distinguish between real time biometric identification and post biometric identification. So, I will start with the post category. This is in general high-risk, and we have many specific requirements for that high-risk system. If you use an AI system for post biometric identification, for example you have to get authorisation by a judicial authority and the use has to be strictly necessary for your current use case. For the real time biometric identification, the requirements are even higher so in general that is prohibited but we have three exceptions. I am going to mention one, for example if you use real time biometric identification in the targeted search for a victim of abduction. In this case you might be able to use this AI system but in general it is prohibited and even then, there are very high requirements. The same as with GP AI, the biometric identification has been one of the most discussed points in the legislative process. We have all the prohibitions in article five and this biometric identification takes up the most space by large, so we have I think four or five subsections that deal with the requirements that you have to comply with if you use real time biometric identification in one of the three exceptions that I mentioned.

Feyo Sickinghe:

I would say especially if you are in the industry and want to use it you have to be really careful to comply with those and it would be good to consult at least a lawyer to look over your shoulder, because when things go wrong there will be a lot of focus on supervisory authorities whether you do this well so, particular point to be aware of. Thanks.

On that topic, let's also touch upon open-source AI systems. They are to some extent exempt from the regulation but nonetheless are in. Can you explain this a bit more? At first sight it seems to be a bit confusing how it has been dealt with.

Oliver Belitz:

It is a bit tricky so, again, I'll try to break it down. First, we have in the article dealing with the AI Act scope we have an exception for open-source software, but we have at the same time several exclusions to that exception. For example, your open-source model cannot be prohibited, it cannot be high-risk AI and it cannot be AI with specific transparency obligations. If you provide an open-source model and it falls under one of those three categories, the exception does not apply. That is the first step.

The second step, what about GP AI? Most open-source models will be general-purpose AI models I guess so what do we do here? So, we have rather high requirements whether you fall under the exception for GP AI so your parameters that means the weight in biases, the model architecture all those really important information about your model have to be publicly available. And you are not allowed to use your OSS model commercially, even if you sell support services to that model then you lose the exception. So, currently when I read the AI Act, I think even, for example one of the most prominent, current open-source models Meta's Llama will not fall under this exception because we have some slight commercial use on the side of Meta. Therefore, it will not fall under this exception. And as a closing remark on that point even if you fall under the exception, it does not free you from the obligation to provide a summary of the training data and to have this copyright policy that Simon talked about, so those two obligations will nevertheless apply. Of course, if your open-source model bears a systemic risk that means it is above the compute power threshold that we talked about then your model cannot benefit from the exception either.

Feyo Sickinghe:

To conclude, one can say even though you might be formally exempt that is only in very, very rare cases so you always have to be very much aware of the provisions that potentially would apply to your open-source system and the way you place it into the market. It is more likely that some points will be regulated than it would stay fully in the unregulated arena.

Oliver Belitz:

Yes, it is a very rare and light exception for open source that we have here.

Feyo Sickinghe:

With this we have reached the end of Part 2 of our podcast on key insights from the political agreement of the AI Act.

In the next episode we will dive into bias in AI systems and how to deal with that, enforcements of new rules, penalties, what individuals can do and some key takeaways that might be useful to you, including the time for the rest of the regulatory process and what to expect. We hope you have enjoyed this episode. Stay safe with AI and stay tuned for Part 3.

twobirds.com

Abu Dhabi • Amsterdam • Beijing • Bratislava • Brussels • Budapest • Casablanca • Copenhagen • Dubai • Dublin • Dusseldorf • Frankfurt • The Hague • Hamburg • Helsinki • Hong Kong • London • Luxembourg • Lyon • Madrid • Milan • Munich • Paris • Prague • Rome • San Francisco • Shanghai • Shenzhen • Singapore • Stockholm • Sydney • Warsaw

The information given in this document concerning technical legal or professional subject matter is for guidance only and does not constitute legal or professional advice. Always consult a suitably qualified lawyer on any specific legal problem or matter. Bird & Bird assumes no responsibility for such information contained in this document and disclaims all liability in respect of such information.

Bird & Bird is an international legal practice comprising Bird & Bird LLP and its affiliated and associated businesses.

Bird & Bird LLP is a limited liability partnership, registered in England and Wales with registered number OC340318 and is authorised and regulated by the Solicitors Regulation Authority (SRA) with SRA ID497264. Its registered office and principal place of business is at 12 New Fetter Lane, London EC4A 1JP. A list of members of Bird & Bird LLP and of any non-members who are designated as partners, and of their respective professional qualifications, is open to inspection at that address.