## **Craig Smith Host** 00:00

But artificial intelligence is a set of technologies developing at an astonishing pace and then it's very difficult to say at one point in time there is this one piece of regulation and that's it. We do need artificial intelligence that serves humanity, that serves the people, and it's our task as elected legislators to make sure that that happens. The council came up with its position, which is called the general approach, earlier than parliament, and something happened in the meantime, before parliament approved its position, and that was called CHET GPT. And that was the moment, like a week or two after CHET GPT was rolled out in Europe and everybody was using it, when the majority in the European Parliament completely changed.

## **Alexandra Geese Guest** 00:53

Hi, I'm Craig Smith and this is my AI. This week I talked to Alexandra Geese, a member of the European Parliament. She's worked on the Digital Services Act in the EU and is now one of the members drafting the AI Act. Parliament has just passed an initial draft of the Act and it's now under negotiation with the European Council and the European Commission. This is the first significant, comprehensive attempt to legislate AI and everyone in the world is watching what the EU is doing. Certainly, the United States is expected to follow a lot of what the Europeans come up with And open AI. Microsoft, amazon, all the big AI players Alphabet of course have been lobbying the EU to try and influence the Act, try and soften the Act. So I wanted to hear from Alexandra how the process has gone and I think anyone interested in AI legislation will want to hear what she has to say.

## 02:18

Before we begin, I want to give a shout out to another podcast. This one's by TechCrunch. It's called Found. Each week, Techcrunch Plus reporters Speckus, gutark and Dom and Dory Davis talk with a founder about what it's really like to build and run a company from ideation to launch. I'm interested in listening to the podcast because I have my own project. They talk to founders across many industries and their conversations often lead back to AI, as many startups start to implement AI in what they do. New episodes are found or published every Tuesday, and you can find them wherever you listen to podcasts, so I encourage you to do so. Now. Here's Alexandra.

## **Craig Smith Host** 03:10

I'm a member of the European Parliament, i'm from Germany and I'm with the Greens EFA group, so that's the Green Group in Parliament and I'm a member of the Green Party in Germany, which is currently in government as well, and I've been working on digital politics since I started in 2019. I was what we call a shadow rupture, so the person who, for my political group, worked on the Digital Services Act, which is a major piece of legislation for digital platforms, and I have always been interested in artificial intelligence and they have been following the AI Act very closely.

## **Alexandra Geese Guest** 03:57

Okay, so the AI Act has. The first draft has been passed by the Parliament. Now it's in a three-party sort of negotiation. Is that right to nail down a final version? I guess, before we begin, can you give the headlines the sort of top line explanation of the acts, of various clauses or the main points of the Act.

## **Craig Smith Host** 04:37

I know, i mean, i know what they are, but for listeners that haven't been following it, Well, yes, the Act is trying to be a very dynamic piece of legislation, because we are quite aware of the fact that artificial intelligence is a set of technologies developing at a astonishing pace, and then it's very difficult to say at one point in time, there is this one piece of regulation and that's it. So what we are trying to do is to introduce different categories and to look at applications of artificial intelligence rather than at the technology itself, and the categories are basically bands, applications which are banned, which is, for example, social scoring in Chinese style. That's one of the most important bands. Another band is the one on remote biometric identifications in the public sphere, therefore cameras in the public sphere that, in real time, can tell the authorities where you are, where you have been, what you have been doing. This is something we know very well from the Hong Kong protests, where people could only protest completely masked and look very fearfully at these huge cameras installed everywhere, and this is a situation we would like to avoid in Europe. So those are examples.

## 06:11

The parliament introduced further bans, for example, on applications that claim to be able to recognize your sexual orientation, just to give an example.

## 06:24

Then there's a second category, which is called high-risk, which is a long series of combinations of applications in different fields, like recruiting, for example, border management, education, where applications need to go through an authorization or licensing procedure.

## 06:48

So companies producing them have to decide whether their tool, the AI tool, fits into that category And then go through this procedure and need to get an authorization by a national authority in one of the countries, the member states of the European Union. They also have to have documentation for what they have been doing, what kind of data they have been using for the modelling, and so on, and they need to make sure that at every time of the process there can be a human person, a human in charge. And then there is basically the biggest category, which is low risk, which just has just a few documentation obligations but doesn't need to be authorised but can be marketed right away. The parliament introduced the fourth category, which are foundation models, so basically large generative models, which has become famous with chat, gpt, and therefore we foresee a few obligations also for producers of these large foundation models.

## **Alexandra Geese Guest** 08:00

Yeah, actually, can we talk about that? because that's what initially spurred my desire to talk to you were comments later retracted by Sam Altman saying that open AI may have to withdraw from Europe. I mean, he later said that they were not going to withdraw, but can you talk about that fourth category and what obligations it puts on people that build these foundational models?

## **Craig Smith Host** 08:34

Well, providers of foundation models have to. I'm going to read this out to get it right, because its law demonstrates, through appropriate design, testing and analysis, the identification, reduction and mitigation of reasonable foreseeable risks to health, safety, fundamental rights, the environment and democracy and the rule of law.

## 08:57

Try and throughout development with appropriate methods, such as with the involvement of independent experts, as well as documentation of remaining non-mitigable risks after development, which I think is interesting because it sort of breaks up this narrative Well, we just provide the technology and then we don't really know what happens, which, you know, for us Europeans is a little bit shocking come the coming from the US, that people provide this enormously interesting and powerful technology and then the developers themselves say well, we can't really control it, we don't really know what's going to happen.

## 09:32

You know, in every other industry sector you would expect the provider of a product to have a vague kind of idea what that product could do in terms of harm, and I understand that for artificial intelligence this is quite difficult.

## 09:47

But at least we require them to make an effort to figure this out. And I think another important obligation is the one on data governance. So it says they have to process and incorporate only data sets that are subject to appropriate data governance measures for foundation models in particular, measures to examine the suitability of the data sources and possible biases, and appropriate mitigation, which is extremely important, because one thing we really want to avoid in Europe from the beginning we're very concerned about was was bias, which we know most artificial intelligence and tail, and it's very difficult to put the responsibility on companies using AI tools that are based on foundation models and then tell them well, you need to get rid of the bias. I mean, how they serve the purpose of the AI tools is very important. That is, especially the providers of these big models really reflect on how to avoid bias seriously. Yeah, and the other issue that I've seen raised is that of copyrighted materials.

## **Alexandra Geese Guest** 10:57

Being used in the training of these models. I imagine there's been quite a bit of discussion in the drawing up of this act About the foundational models that are in the market now GPT for or, or, you know, Google's barred. How, what do you do now? I mean because those models are already trained and they're trained on, you know, virtually the entire English language Internet, Or maybe the Internet itself. I'm not sure how that was segregated.

## **Craig Smith Host** 11:55

Well, in German as well, Interestingly yeah.

## **Alexandra Geese Guest** 11:58

So I mean you can't, once the model is trained, you can't, you know, take out that information.

## **Craig Smith Host** 12:10

So does that mean that people providing foundational models will have to train from scratch, start over, and if they want to operate in Europe, and provided that the act passes in its current form, Well, to be very frank, I think it's a question that is currently very difficult to answer, but I think a good start would be obliged the providers of foundational models using that material to to be open and frank about it and to be able to distinguish between copyrighted material and other kind of material. I mean copyright may be a little bit outdated, but it's still a foundation, the economic basis for human creativity. Basically, you know, journalism, arts. That all depends to some extent on copyright legislation And we need to find a way forward there without just completely abolishing it.

## 13:16

You know, I've been talking to artists who have been what we're very, very worried about, about what they're going to live on, and I think there's a, there's a component really of that goes beyond just the economic reward or the legislation. In terms of what? What does this kind of technology mean for quiet creativity? Because, you know, this looks, the AI looks very creative, but then its work is based on the work of humans and if humans are not rewarded for it anymore in any way, at a certain point creativity will end. So what does that mean? Well, I'm not a copyright expert, so where this leaves us currently in terms of enforcement of copyright legislation, frankly I don't know.

## 14:06

The AI Act is not legislation yet. You explained that correctly. It's the parliament's report And now we enter negotiations with the European, with the council, and then we will have a final legislation and I think in the meantime we will reflect on what that means for copyright legislation. I don't see open AI or bright retraining completely their models, but maybe, you know, looking into paying more attention to what kind of data is used to train these models. I think it's sort of a learning process for us as legislators to understand how artificial intelligence works and what it does to the world, but also to the developers and the companies providing it to abide by, to comply with existing legislation and to understand the needs of society that are expressed by that legislation as well. I see it was a process, frankly.

## **Alexandra Geese Guest** 15:12

Yeah, I mean, is there a concern in Europe that if the act is too restrictive, that Europe will fall behind in the implementation of AI and consequently fall behind economically?

## **Craig Smith Host** 15:38

There is a concern definitely from the conservative side, I think, then, from the more progressive part that I represent, but there is a concern. There is a concern that those products might not be rolled out in Europe, as Sam Altmont said Well, he would just leave. There is a concern of European companies not being able to build on these foundation models and therefore to fall behind. That is true And I think we do need to take this seriously. On the other hand, the European Union is a huge market. The digital world is basically a more western world, so the Chinese market doesn't count that much because, in digital terms, it's a different universe. So Europe is still quite important, i think, for companies like Google and Meta and Microsoft rolling out AI, and therefore I think there will be some arm wrestling, but I think they will listen to us And I know that many colleagues from the US as well are looking with great interest to what we are doing in Europe and hope we might be a model, and, at the same time, i think we do need.

## 17:05

We do need artificial intelligence that serves humanity, that serves the people, and it's our task as elected legislators to make sure that that happens. I mean, corporations have a different purpose. Their purpose is to make money for their shareholders. So I think we need to tell them what society wants, and I try to listen to my voters and to the citizens and figure out how they want artificial intelligence and what it should be used for. And there's quite a lot of fear, and I think that doesn't help the companies either. People have to have the feeling they understand these instruments, they are in control, and if they want some more control, the providers of AI tools can provide it, and I think that would help us all, that would help the corporations and it would help society.

## **Alexandra Geese Guest** 18:00

Yeah. So right now there's a dialogue or a trial log I guess they're calling it in between the commission, the council and the parliament to finalise this. What are they? does the council tend to be more conservative and more pro AI, particularly with regard to the foundational models, than the parliament? I mean, are we likely to see this draft become more restrictive or less restrictive as it goes through this negotiation?

## **Craig Smith Host** 18:49

It depends. Normally the council is a little more conservative than the parliament And another government might change from progressive to conservative, far right in the process. In Spain That's a risk, but there is another and conservative or far right companies countries tend to be more open to what the corporations want. This is definitely the dynamics we have observed in parliament as well. But there's also a second issue which is, or a second aspect, which is time, because the council came up with its position, which is called the general approach, earlier than parliament And something happened in the meantime, before parliament approved its position, and that was called chat GPT.

## 19:52

And that was the moment, like a week or two after chat GPT was rolled out in Europe and everybody was using it, when the majority in the European parliament completely changed And the conservatives went along with a few of the bands, for example the one on recognition of sexual orientation or emotional recognition in education or in water management, that kind of thing and went along with the idea to include the foundation models, which they didn't wanna do before, and at that moment I think they realised how crucial and how influential these models are, and that is a good idea for a democracy to at least be able to understand what the risks are and where the data comes from.

## 20:43

So there is this aspect of time and of really understanding what a foundation model is, And I think this might entail the fact that council might be a lot more open towards parliament's position as well, and the commission also had a very strong learning process.

## 21:04

The general approach by council is very close to the original commission proposal And that was also set out before CHET GPT. I think if the commission had to redo it, they would do it differently and they would definitely include foundation models, because in the beginning the idea was just to cover high risk applications and foundation models at the time were not considered high risk. Now, when you say, well, you know, if my eye didn't do a lot of moderation, you could just ask a CHET GPT to explain how to build a bomb or something like that. They realised there is a high risk aspect to it and I think the council will realise that too. So in terms of dynamics in council, let's see, and I think I'm very optimistic in terms of keeping the foundation models. I think the other big issue with council is the biometric recognition in public spaces, because they are the ministers of the interior that would like to be able to survey people to a much larger instant than members of parliament think that is useful.

## 22:17

So I think that's going to be one of the big issues In terms of foundation models. I'm quite optimistic.

## **Alexandra Geese Guest** 22:25

Yeah, on the foundational models, you say you're optimistic or that there won't be major changes required on behalf of the providers of these models. I mean that the existing models won't be banned based on the inability to identify copyrighted materials in their training data, for example, and that companies providing those models will have to start from scratch. You're optimistic in that way, or you're optimistic that those restrictions will be waived for the current generation and it'll be applied going forward?

## **Craig Smith Host** 23:17

I'm optimistic that the obligations to test and analyse, identify reasonably foreseeable risks and risk medification measures will stay in, and the obligations on data governments. That's what I'm optimistic about. What exactly that means for copyright, I can't say at the moment. I think we will need a system I don't know. some people have been saying watermark or the copyrighted material or material created by artificial intelligence. I'm very frankly, I don't think we have the solution yet, but I'm confident that the idea of providers of foundation models having obligations for these models and for the use of copyrighted material will stay in. What exactly this is going to look like, I'm not sure yet.

## **Alexandra Geese Guest** 24:22

How much have the big American AI companies been involved in this process? I mean, have you consulted with Meta and OpenAI and Google, and how involved have they been in consulting on some of these issues?

## **Craig Smith Host** 24:48

Well, they're quite involved. They talk to the Commission on a regular basis and different formats. They talk to the members of the European Parliament and, more involved, they talk to the national governments. They are quite involved.

## **Alexandra Geese Guest** 25:04

Yeah, do you think that the negotiation of this act will have, until it's settled, will have a chilling effect on? because right now we're in a period where everybody you know across the economy is trying to figure out ways to integrate foundational models into their products and services, and in fact there are. There's a sort of wave coming wave of new startups and applications built from the ground up on these foundational models. But if you're an investor or a company in Europe, it seems to me you would be cautious about investing too much in building products and services based on foundational models as they exist today, without knowing what the regulation is going to be in a year or two.

## **Craig Smith Host** 26:21

Well, I would see it the other way around. I think what we are basically doing is saying the providers of the foundational models themselves are responsible for fulfilling these obligations, because otherwise it would mean that the smaller companies implementing the foundation models in their products would be liable for fulfilling the whole legislation. And the legislation already says if you are in a high risk environment, you have to make sure, for example, that the data is representative. How are you supposed to make sure that the data is representative if the provider of the foundation model just says, well, I have no idea where this data comes from? So I think what we are trying to do is to make sure that the burden of fulfilling with the legislation is with the legal subject that can actually do, that has actually access to the data, and not with somebody just buying it or implementing it, that doesn't really have the possibility to influence what happens, and usually what the companies want is legal certainty and not a completely open space.

## 27:42

I mean, in the US you have the same problem because if there's no legislation, companies are still liable in case of discrimination in lawsuits. So we have a different system in Europe. We say, okay, we give you the legislation, but then if you fulfil the legislation, you don't have these enormously high risks in lawsuits which companies in the US have. So it's a little bit different. But legal certainty is certainly something that companies in Europe do want. And making the people liable who actually develop the tool and have access to the data and can have the possibility to know which data has been used. I think that makes sense. I don't think there's a chilling effect. I think there's more of a chilling effect if you know, well, probably there's something coming, but we don't know what it is. It's better to know what it is.

## **Alexandra Geese Guest** 28:42

Yeah, and are you consulting? I mean imagine you are with the US Congress, because everyone's following this very closely and the US Congress is looking at legislation itself to set a legal framework around some of these things. How much influence do you think the parliament's having on what's happening in the US Congress?

## **Craig Smith Host** 29:21

I haven't been consulting personally directly. I have done that on the Digital Services Act, where I was really very, very directly involved, and what I heard then was well, go ahead. In Europe we are super interested in what you're doing and we try to follow your lead, but we have difficulties in the US putting together a majority on these issues. It's actually more difficult to do if the companies are in your country and they are extremely powerful corporations. So in a certain sense, we have that advantage in Europe. I mean, we're sorry not to have the technology produced in our in an hour content, obviously, but it does give us some leeway to reflect on the legislation and to pass on the legislation. What I'm hearing from my colleagues who have been to the US recently is that colleagues in the US are still looking very much at Europe and what we are doing and that they appreciate the fact that it's a very dynamic kind of legislation. And then let's see what happens.

## **Alexandra Geese Guest** 30:35

Yeah, when you say dynamic you've used that word a few times what do you mean? that it can change over time?

## **Craig Smith Host** 30:43

Exactly, yeah, that, for example, on the foundation models. What we are basically saying is not they're banned, they're not banned or they have to do this, and that We are just saying providers take responsibility for what you're doing. So we ask that's why I read out the paragraph on the risk assessments. They basically should themselves try to identify the risks and propose risk mitigations. So we're saying, well, you're rolling out this product, great, really interesting innovation, looking forward to using it, but please tell us what it entails in terms of risk, which is exactly what you would tell pharmaceutical companies. So you're helping people get cancer therapy, great stuff. We're supporting you, but tell us what the side effects are, and if you don't do that, you get oxy content or something like that.

## 31:40

We are just asking these companies to do what in each other, each and every other industrial sector would be completely obvious. Give us the product very interesting, but tell us what the risks are. I mean, we do drive and we know it's dangerous, but we do have traffic lights as well. So what we are doing here is just Tell us what kind of traffic lines we need, what do you think? And then let's discuss it. So I think this legislation, especially on the foundation, will be evolving. But when you have an industry that's telling you, well, these are very interesting products, but we don't really know what they could do, i think we need to get involved with this legislator, say, well, try to figure it out and let's discuss it together and figure out how we can protect our societies together and use this product in a way not to harm people.

## **Alexandra Geese Guest** 32:36

Yeah, has the so-called threat debate that's cropped up in the last couple of months affected? I mean, imagine it has the direction of the act. And have you had? because a lot of the people who are raising these concerns are originally European Yanlacun from France and Joshua Benio from France and Jeff Hinton from the UK I'm not sure what their current nationalities are, they've lived in Canada for quite a while But have those people been involved in talking to the parliament? And the reason I ask is there's a discussion. You know that debate has split the community, the research community, but there is an effort right now to come up with a consensus within the research community, so that it's not just people sort of yelling at each other on Twitter that there's some unified view of what the risks are and how they can be addressed.

## **Craig Smith Host** 34:06

I don't think the fact that they are Europeans makes a difference. I mean, a lot of the people developing and also financing those are Europeans. So the fact that those who point out the threat are of European origin, i don't think it made a difference absolutely.

## **Alexandra Geese Guest** 34:27

Right, well, go ahead, I mean, but have you consulted with those senior researchers, people like Lucun and Benio and Hinton?

## **Craig Smith Host** 34:39

I personally haven't. I don't know whether my colleagues did So. I haven't seen that as a formal process in the European parliament And I haven't personally, but I cannot exclude that other colleagues have done that. I don't know. I imagine they might have spoken to the Commission as well, but I'm not sure about that.

## 35:01

Yeah, yeah, what I do think is that that more the discussion on, you know, the singularity issue and AI wiping out humanity has not been extremely powerful in Europe as far as this piece of legislation is concerned, because the first draft of this was written two years ago and is really trying to address the current very real risks, like bias, for example. Now, I think something I care a lot about is disinformation also, which is not that well addressed in the AI Act actually. But I think bias, bias, what the big thing? security issues, that kind of thing, So more the things that are already on the table rather than what could happen in 10, 20, 50 years from now, that's more an intellectual kind of discussion also going on in Europe, But it hasn't had influence on this legislation, I think.

## **Alexandra Geese Guest** 36:03

Yeah, and the? Where is the? How long is the process expected to take before this is formalised as law?

## **Craig Smith Host** 36:17

That is very difficult to foresee. Could be six months, could be two years.

## 36:23

Right, I think there is some ambition to conclude this within this mandate, which ends in June 2024. So we have a year, basically, but that would mean that the negotiations would need to end by something like the end of the year, yet January, February in order to be able to go back to parliament and vote on the final result and counsel the same. It's a complicated process, so only in the couple of months at the end of the negotiations can I put it through. I think that's the ambition everyone has, but it is a long piece of legislation with a lot of content and in an environment that is constantly changing at an amazing pace, so I can't promise that this is going to happen. I'm sure the commission and parliament would like this to happen.

## **Alexandra Geese Guest** 37:14

Yeah, And on the enforcement. so the law gets passed. Who is charged with enforcement? And I understand that it's up to the companies providing models to ensure that they adhere to the act, But who is reviewing models or will be reviewing models and deciding on which risk category they should?

## **Craig Smith Host** 37:54

be National authorities, so it's the governments of the member states who decide which authority will do what Might be one single authority. There might be several. I'm hearing that the Netherlands is going to be the data protection authority, for example. In Germany, i think it's more what we call the agreed organism, so the authorities that do audits on technical specifications, like in the automotive sector, that kind of thing It's. I think there's going to be a variety of models of implementation, which is going to be very interesting. This is why we also have the board where these authorities will exchange on a European level their experiences.

## **Alexandra Geese Guest** 38:44

Yeah, and what do you think would be the penalties? I mean, of course it's too broad a question to ask and answer very specifically, but what kind of penalties do you envision? I mean, is it the sort of thing, as with pollution, industrial pollution, that the very large companies may be content to pay fines and continue to operate in the way that they do? Or would they be banned entirely from the union? And in that case, how do you police that? How do you know if someone is using a foundational model that breaches the act?

## **Craig Smith Host** 39:39

That's a pretty good question. I think we start out with fines And I do have doubts about the efficacy of fines. We see that and data protection. When Facebook, for example, was fined with a huge fine by the FTC, actually stocks went up the day after. So I do have my doubts about this. Yeah, it's also true that the data protection legislation has not been enforced very well in Europe. So we have a few fines, but more are coming. Then the Digital Services Act also works with fines. So this is going on full speed next year And I think we will see some fines there. Maybe it's a mechanism that is not at full speed yet And at a certain point it will become a mechanism which does become not just cost of business but a threat for a company, especially when you have shareholders just watching this and saying, okay, these fines are accumulating, we are constantly violating important legislation, let's change. So this is what I hope for. I hope I'm not too optimistic.

## **Alexandra Geese Guest** 41:15

And what about government uses of this technology? Does the Act constrain governments and militaries of member countries?

## **Craig Smith Host** 41:35

Military issues are not covered explicitly, so military tools are not covered. It does constrain governments. Yes, in many cases the rules for governments are even stricter than for private operators because clearly governments have full control over people, so people don't have a choice. In some cases the European Parliament has made stricter rules also for private entities, whether the European Commission only foresaw those strict rules for public entities. Cases like, for example, data scraping on the internet with image recognition like Clearview, for example.

## 42:30

the Commission only prohibited that for public actors and we added a ban for private actors too because, i mean, it's absolutely, especially for women. it's a super huge risk. People you know see you in the street and can figure out where you live and what you have done in the past 20 years of your life. You know almost every woman has had a stalker at some point in your life and it's just a total nightmare the idea that there is a commercial software out there that allows everybody to know who you are just from looking at you or taking a picture or something like that. So, but normally the rules for public entities, for governments, are even stricter than for private ones.

## **Alexandra Geese Guest** 43:10

yes, Yeah, and I mean there's so many directions to take this. There's a view that, as this legislation takes hold and the technology continues to develop within those regulatory frameworks, that different zones, ai zones, will emerge, that you'll have the areas of Chinese influence, where AI uses in surveillance will be much greater and will have a much bigger impact on governing society. You'll have Europe, where it's less restrictive but yet more restrictive than the US, and then you'll have the US, which will be much more restrictive and more restrictive than the US. So I think that's a good point. You'll have the US, which will be kind of a bit of a wild west, depending on what Congress comes up with. Do you think that that's the case?

## **Craig Smith Host** 44:35

Might be. Yes, It's a very probable scenario. There could also be a scenario of the US and Europe merging and the US maybe following at least the spirit of European standards. That depends on the US. I think, Yeah, as has been the case with GDPR?

## **Alexandra Geese Guest** 45:02

Yes, yeah. And going back to the foundational models, I mean this is a real problem for regulators everywhere, but since that came along late in the process, what's your best guess of how that's going to evolve, that you're going to evolve in the future? I'm going to evolve that existing models, like GPT-4, even though they cannot take out the data that the models were trained on, I mean it's too late. The models are trained that those will be grandfathered in and these standards will be applied to models going forward, where those original models will eventually be banned because they were not designed. With all of these questions in mind. I mean, I know not, no one can answer these definitively, but I'm just curious what you think.

## **Craig Smith Host** 46:27

Well, it would be great to have a glass for your words, wouldn't it? Yeah, I think that's an interesting idea, like some kind of grandfathering and then going further. It depends a little bit on the choices the companies and the developers make, i think, whether they say well, the Europeans have a point there and let's try to build models where we have more control over the data, which would make sense maybe in a US environment as well, in order to avoid lawsuits because of discrimination, for example. I mean, it's pretty probable that when the foundation models gets implemented in some kind of other application, there will be a strong form of bias and discrimination, because we know this is historical data which is basically racist and sexist, and now they are sort of controlling it because you can't ask it to make jokes about men but not about women. But it requires an enormous moderation effort And I think that is very interesting. That's something we don't have in the legislation because it sort of came up too late. But who is doing this moderation and based on what So deciding?

## 47:44

What are the questions that chat GPT can answer and which ones are not allowed to answer? I think that's going forward with the really interesting question. Can we get some transparency on that? Can we have a global standard, and what is that global standard? Is it human rights? I mean, this is where it really becomes interesting, where it becomes a cultural issue, a political issue, and where the US and Europe differ quite a lot. So I think that's the more interesting question And they're saying the answer might be well. We explain what the data is, we've been using and what sort of the guardrails are, in order also to have some kind of legal certainty not to be liable for every form of discrimination. I think that could be an idea. And how to get there, whether it is to start a wholly new model or whether it is rather modifying the data, because the models get fed with new data and trained on new data as we speak, as opposed, and whether there's a gradual shift or a wholly new start. Frankly, I'm not able to say.

## 49:05

We would not need to do this. but I think even the developers and the companies can't say at this moment that they will wait for the legislation actually to enter into force, probably before they change something and try to avoid it as long as it's not, and then we see what the game looks like. But I can envisage different kinds of avenues there. But as I said, I mean this is developing so quickly that it's difficult to foresee from here.

## **Alexandra Geese Guest** 49:33

Yeah, I'm curious. Have you used chat GPT or GPT-4 in your work or personal life?

## **Craig Smith Host** 49:44

I have Experimenting more than really using it, and I used it to do different things, which was interesting because I had the feeling that the hardest tasks performed best and the very easy ones performed extremely poorly. Like I had it, you know, just write about myself, who am I, and it came out like with the wrong birth date and the wrong commissions I serve on, like you know, all the kind of information you can very easily see on the website. It was all wrong. And then I asked to write very complicated, you know, argumentations for something I was negotiating, a lot of pieces of legislation that are really complicated and very few people know about, and the arguments were great. You know it would have taken me two days to put that all down and Google and chat GPT did it in two minutes And it was amazingly good. It wasn't perfect, but it was amazingly good. So I'm a little puzzled as to that.

## **Alexandra Geese Guest** 50:51

Yeah, this is the hallucination problem, but there's a lot of the products coming out. There's a way to solve that. I mean already chat, gpt and Bing. you know there's a model that browsers the internet, so it gives you footnotes. But the reasoning power of the model is what people are seeing as a promise, not using it as sort of a knowledge system, but as a reasoning system, and I'm curious.

## **Craig Smith Host** 51:35

See it as a knowledge system as well, that we use it instead of Wikipedia, because it's so much quicker and it picks up knowledge. This is something that worries me, because it's going to be the tool that gives access to the knowledge of the world, and who decides what kind of knowledge? I think that's something that I think is really interesting.

## **Alexandra Geese Guest** 51:52

Yeah.

## **Craig Smith Host** 51:53

The experts are more interested in the reasoning and it's really on intelligence and so on, but normal people just use it in their daily life, i think they're more interested in this.

## 52:01

You know, just abbreviating the task, basically, and having access to that, to all kinds of knowledge, and that it completely makes up the commissions, the committees I serve on, is bizarre, but I think that that will be fixed.

## 52:17

What I've found more worrying and you probably know that example that Tristan Harris came up with, that he shows how it's being implemented in Snapchat, and there's this example of a 13 year old that basically tells the artificial intelligence implemented in Snapchat that it's going to be he or she is going to date an adult or 30 something year old in a weekend away somewhere And the tool just says, oh well, that's great and have fun, you know, rather than saying, well, i should call the police So, and that's the kind of risk I think foundational providers of foundation models should foresee somehow before we go out and implement it everywhere. So that's the reason. Just to come back to a question you asked me before. I mean, even if we like a little bit of implementation in Europe, but if that means we make sure that people are safe. Our kids are safe. For me, it's okay to like behind a little bit.

## **Alexandra Geese Guest** 53:22

Yeah, so you're optimistic that these questions will be answered. And you think that within two years there will be an act, but it may not be implemented strictly at the beginning because it's going to take time for providers of models to adjust, in sort of looking forward, do you do, you see? I mean, you know the AI Act has gotten a lot of attention within the AI community, but this technology is really going to alter, it appears it's going to alter every aspect of society, if not, and certainly of the economy. I imagine that you guys have thought about what, what the world, or certainly what Europe, is going to look like in five or 10 years. Are you? Are you optimistic, hopeful or are you terrified?

## **Craig Smith Host** 54:46

I think the three at a time Yeah that's on how I wake up in the morning.

## 54:56

I'm optimistic because, you know, when I started professional politics in 2019. And people told me well, you can't regulate digital because it's such a new technology and we know it's too late anyway and you can't regulate it, it's all you know. Don't, don't even start. And then we did the digital services act. We started at least enforcing GDPR and a little bit more serious. Well, still a long way to go. But the digital services act was basically regulating social networks. And I had been told for you well, you can't do that And, yes, we can, let's see what comes out of it next year, but it's possible. And now it's AI And you know, the thing started over when you can't regulate it. You can't regulate it without stifling innovation, without having your blank behind, and now we are engaging in a process and we are regulating it. So I think that's the optimistic part. And then you know, everybody tells them.

## 56:02

When you look back into history, there were always times you know industrialization when it seemed that the world was just going completely rogue and terrible. Then, in the end, something good came out of it. We dealt with it as humanity. So that's the more optimistic part, saying well now we're not at war, we're not fighting, we are just debating. We are in a very civilised process of passing legislation And let's see.

## 56:32

I think the companies will, at a certain point, adapt to it, maybe not, and the legislation is not perfect yet. There's not everything in there I would have seen, would have wanted to see, and things are coming up. New things are coming up We are not dealing with yet. That's absolutely true. So that's the more optimistic part. The more terrifying part, you know, if you look at the kind of surveillance this allows, the power concentrations in the hands of really a handful of companies out of Europe, what the Chinese totalitarian government is doing with it, the kind of environmental impact that it could have on top of the climate crisis we are already living and not really facing, yeah, the moment when I am terrified and I think it's justified to be both terrified and keep working, and in order to keep working, you need to have the belief that you can do it without being certain.

## **Alexandra Geese Guest** 57:36

That's it for this episode. If you want to read a transcript of the conversation with Alexandra, you can find one on our website, that's EYE-ONai. And don't forget to check out TechCrunch's podcast if you're interested in how to build a startup. In the meantime, remember I don't believe the singularity is near, but AI is about to change your world, so pay attention.