Interview #1

**Jan:** Thank you. So to start with, could you broadly describe your area of work?

**Interviewee:** Broadly describe my area of work? Okay, so I'll broadly describe my background as well, so it might be relevant to your studies. For example, as we've just met here, I'm very happy to see, to see and see you in the study, and to be able to support as much as we can.

Great. My background actually is a convoluted one in the sense that I've always worked in technologies but in very different environments. So I started my career in the **[country]** police force as a police inspector. And so there I dealt with official police training, of course, legal training, criminal law, as a prosecution officer as well for a while.

And then most of my career was in data centres dealing with high-security systems on tech and IT and Cyber related matters as well and later on, when I moved to health within government, I've always worked with government. And health systems, again, are very huge systems, very hot systems, as you can imagine.

There's life and death situations as well involved, like there was in the police force. And now today the opportunity came at **[company name]**. I started as a regulatory officer here, and now I'm working on the technology architecture side of things. So it's a bit of a background. So my work here I've touched upon the, there is the letter of ethics as well, which we consider often but as I said, when we were speaking together, I think ethics is a bit of a convoluted word in itself.

It has a definition problem, like AI has a definition problem. No matter what you might say, the definition problem, so far as I, as I look at it, I can't tell you. Say it's completely resolved or resolved, but it's maybe a work in progress.

But there's an issue of what is intelligent was then artificial. And, you know, all this, furthermore, it's compounded by the ethical side of things. So these are the matters in terms of regulations that I deal with as well as the technology. We've built a lab recently. So I deal with technology stuff all the time as well.

So in a nutshell, it would be a very high-level adventure of mine, right?

**Jan:** In your area of work, what are the ethical considerations, when you are working in sort of this development and with these systems?

**Interviewee:** Ara all the time, given that we are an authority, so we have like a regulatory, so to speak almost on us, right? Responsibility. Sometimes I hear the word of regulating technology, which, you know, gives me the chance and increases my stomach acid. Because in such technology, still in its infancy, we still don't have the full potential or full grasp or understanding of that and Europe is already wanting to go very harshly on regulation, hard regulation, rather than soft regulation. The difference is one, soft regulation might be an example, be a certification program, voluntary. Hard regulation is banning something. If you want to ban facial recognition. You want to ban anything classified by biometric means.

Banned, you know, in Article 5 of the AI Act. And the dilemma is this one. Okay, we're banning, we're doing this hard regulation. Europe will tell you because it's of ethical concern, you have to build ethically correct systems. Ah, ah, okay, what is ethically correct? And when you have concerns with bias. Okay, one, one, one does agree that if you have malicious bias in your data or in your algorithm or both, you're going to have an effect that if you're relying blindly on AI, then you're going to have a nasty effect.

But perhaps the issue is relying blindly on AI. Not that you have some bias somewhere, and you speak to legal teams and the European Union was the first to actually say we know how to solve this. How will you solve it? We write an article if data must not contain bias. Now, when you write such an article, that's a scientific impossibility. You can write it, but nothing has zero bias. If we think and write in English, you have a bias. If you're born in certain sides of the world, you have a bias. You know, it's, it's going to be there.

So, this constitutes an answer to your question. We discuss a lot of these article issues all the time, but sometimes you have the best intent, the most ethical intentions, you do regulate and you create unethical things on the other side, like, for example, stifling innovation developments, scaring companies from investing in Europe because they see you like an overregulated, you know, I would argue that's not an ethical thing to do.

Again, it doesn't answer any questions but that’s ethics, right?

**Jan:** Yeah, it's a food for thought. So you said that you discussed this a lot, so are you aware of any within this field of work, any the existing guidelines already in place and ensure or facilitate maybe

**Interviewee:** Existing guidelines, which you could have a look at, of course, **[country]** was the first to issue AI guidelines with a voluntary certification.

We left it in draft format. It's available on our website. under the token, under the area of consultation. So you can look at our website, go to consultation, and we have AI guidelines, which is a, which was a draft certification program on the best way to develop or best guidance to develop AI systems.

So we introduce the idea of having the hardness, like an AI should not exceed the hardness limit, if there's such a limit, right? So yeah, but of course we left it in draft and did not issue and continue with it because the European Union and Commission were issuing their own.

So when there is such an effect, then it's like overrules you. So that I can understand why. So EU has one framework that fits all. One size fits all, right? So **[country]** having its own regulations, not every country has its own regulations. So this is the idea to unify. If someone is certified in **[country]**, then it's automatically certified in Spain, right?

So that was the idea. So, yeah, so we were one of the first. You asked me about guidelines, right? Or ethical. Yes, we follow very closely and even took from the OECD guidelines and best practices for developing AI. Those, I think, are brilliant documents. Brilliant documents. It's truly helpful for your thesis, if you're delving into ethics, the OECD, OECD ones, make sense. Because they are very technologically relevant. Not simply written by lawyers somewhere who have no idea of tech. There, it was clear to me that the tech advice was taken on board which I cannot say the same for certain European documents.

Yes, and there's also been progress and developments in ISO standards and some developments over there. Of course, again, AI has a definition problem. I'll go back to this many, many times. And it's always a problem. When we started, we had no idea of the capability of GPTs and the like, right? So, now we've changed everything and it keeps developing and so it should.

Let's not stifle that, because if we stifle that, I believe that we'll be out of things. And you can stifle such developments through excessive regulation. Did I answer your question?

**Jan:** Yes, yes, thank you. Actually, you answered two, because my next question was going to be if you were sort of how they sort of contribute to this field, what these guidelines would be.

**Interviewee:** We contribute by a lot of discussions, even with the big four. I'm often interviewed, like this interview with yourself, by OECD, or institutions coming from abroad. I saw I had IEEE as well as have a chat with me with us here as a staff at NBIA. So yes, we do contribute a lot and of course, we give the multi, multi-position in terms of these regulations as well all the time.

We guide our ministry, we guide our um, contacts that we have in Brussels and, and internationally as well. We take part in for throughout Europe and worldwide. Yeah. So this is a continuous project, right?

**Jan:** So you're obviously very involved in all these, all these different documents and with talking to many of these different people.

So within, the sort of field surrounding ethics and AI, which as you said, it's quite a broad thing, are there any sort of emerging trends or developments that are identifiable?

**Interviewee:** I think there are a bit of opposing forces. So we are seeing the American. And the American sphere is like, don't touch AI with hard regulations.

We want to see this developing, which I personally kind of agree with. This is just so I personally, but Europe, on the other hand, is all about hard regulation and banning specific use of AI and specific types of AI, which I would argue would be beneficial in some scenarios. I would, I would rather regulate.

I would rather regulate on. Not banning, by simply banning the AI, rather than, Okay, so I can phrase this, The issue is when one trusts AI blindly. Like self-driving cars. For example, I don't know how AI works. I would never trust AI. Completely blind, sleep in the car and the car drives. Why? Because the car is driving in a very stochastic environment.

It means a lot of random variables. The variables are huge. You have a dog crossing, you have a reflection in the street, heat, bocci tarmak, but you know how AI systems work, right? Are you going to trust it blindly? A machine which is really and truly a statistical-based engine, because that's what AI is, a statistical engine.

It's a fantastic, fascinating statistical engine. Blindly with my own life, driving at 100 miles an hour, I am very sceptical of that and seeing some of the demos, which Sometimes it's like self-driving cars, companies try to hide. Some of them are really bloody scary. And as a tech person like yourself, I would imagine, I would immediately recognize that this is something that a human mistake would never do.

You see, this is the issue. The machines make mistakes humans normally never do, and humans do mistakes that normally the machine would never do. So a machine would be like a 99 per cent confidence interval, with a very high confidence interval that a triangle is actually a cat. When a human looks at it, they would never confuse a triangle, with a cat, right?

Because they're seeing the ears and everything. The machine is saying high statistical value that this probability is high or that this is a cat. A mistake which a human normally would never make. On the other hand, the machine pulses, pulses, pulses for many thousands of hours and never gets tired. And so it does not make these silly mistakes that a human does.

I think if you want to really, and this is a medical concern, To leave AI human-centric, 'cause that's the word we use, right? Human-centric AI. The human has to be at the end and the AI has to be treated like an assistive tool. A tool that is there to help. Like I have Microsoft Word. Microsoft Word is a word processor.

It's an assistive tool. I can use it to write you a very nice threatening letter that I would kill. You offend me. I use Microsoft Word. Right? Or I can use it to write a love story and use it for business. Or I can use it to help someone in need by telling them to help by writing a letter of courage.

Right? I use the same tool, Microsoft Word. We have an issue with this, we should regulate Microsoft Word. No! You don't regulate the tool, you regulate the use of the tool. In the same way, we regulate killing. You don't regulate pencils because somebody killed someone with a pencil, which you can do. Right?

We regulate pencils. No, I can't, we can't afford to end up with very super soft pencils and then you can't write with them because somebody, somewhere, killed someone with a pencil. Or with a knife, right? If I need a solid, sharp-cutting knife, I need a sharp-cutting knife. So it's yeah.

**Jan:** I think in many discussions, especially with other people who have a bit of a background in AI, some students and sometimes even with lecturers and so on.

It's always something that ends up being the sort of, well, there's no one sort of solution, but always seems to be the direction that people say that we should move in is that it should complement human capability, not outright replace it. Decision-making should always be determined. It's an assistive tool.

In our first year, in fact, we had a unit. This was now two and a half years ago, called Foundations in AI. The scope was to develop a sort of system where AI could be implemented to facilitate a human task. And what my group had done was, And they, we had come up with a concept for a chatbot, a, a sort of chatbot that could be queried on and talked to on a sort of website, similar to how kellimni used to operate.

Where people, who maybe are feeling anxious would be able to get support. They would be able to talk, to this and using, so this was all concept level, but using certain metrics it would be able to flag. That may be if a user is at a higher level of risk or if there's a requirement for intervention, but the bottom line is always that the critical decision has to be taken by a human.

It can be, you know, flagged by maybe an automated system, but then the human still has to see the facts and sort of.

**Interviewee:** Again, this is where the ethical dilemmas come in. Is it ethically correct that someone in need of psychological care is actually speaking to a bot? Maybe not, and therefore you tell this person, Hey, you're speaking to a bot.

But at the same time, by having such a facility, maybe you helped 50 people instead of 3. Because you don't have the manpower to do so. And what if the bot gave some advice, or suggested something harmful, and until the human realized, you know, there are these But I mean, what's the answer to this? You don't do anything, because everything has an ethical concern.

**Jan:** You mentioned that the human might not be aware, or even if they're aware that they're talking to a bot. Is this sort of, you know, ethical? In the broad sense of the word. How is sort of the trade-off between these different ethical principles such as, for example, if you're talking to a robot, it has to process what you're telling it, sort of, so you are giving it a certain level of privacy, you're telling it may be your name, your age, maybe.

**Interviewee:** Maybe like the Chat GPT voice thing. Have you tested it?

**Jan:** No, not yet.

**Interviewee:** It's amazing, well, have a go at it.

**Jan:**  How would the trade-off be sort of assessed, maybe, in your field of work?

**Interviewee:** It, I mean, there's no, unfortunately, this subject is not a one or zero, which is what I hate about it. And I hate it with a passion. Because I strongly believe if you're doing something ethically correct somewhere, it's ethically incorrect elsewhere.

I killed a rabbit who I had to eat. I mean, if I left the rabbit alive, the rabbit would have eaten the vegetables and pooped somewhere and infected another human elsewhere. I mean, where do you stop? Where's the line? I don't know. Where's the line? I'm, I'm always confused in this. I mean, it's to, to me it's unanswerable.Whoever has the audacity to answer it precisely is probably wrong or not thinking deeply enough to which layer do I, it's like a mega convolutional network, neural network. Right. To which layer am I going to stop? And after all, is it analytically correct? I'm wasting so much time thinking about this when I should actually try something and see what happens.

There is value in experimentation.

**Interviewee:** Alright, I'll be clear that I'm perhaps experimenting. People don't read? Ajma ok, it's a user problem, maybe. It's like GDPR. You accept the cookies. Are you reading all the cookies? All the terms and conditions before accessing the website. Before accessing any website, which is GDPR compliant, you need 45 minutes and a legal attorney next to you.

Before you click, I agree. Is everyone doing that? I don't think so, right? So, I mean, I'm pretty sure GDPR was done with a good ethical, moral compass and heart. But what's the result? Is that the result?

There's always going to be something ethically incorrect, which is ethically correct elsewhere.

And that's reflecting my point. Now, doing nothing, I would think it's ethically correct.

**Jan:** So what would you say are the sort of biggest potential consequences of failing, to see these concerns? What would be the biggest sort of, you know, consequence of this? These ethical concerns within these systems, for example, either within GPT, a person talking to GPT, or within the place of work, you are using it to facilitate your, your task.

If there's a sort of privacy concern, if there's the information might not be accurate, what would you say is the biggest concern?

**Interviewee:** I’ll speak as a tech person and the CEO of Yahoo said it, privacy is dead, we have to deal with it. Okay, so this might not be I think you have expected, but if I ask myself personally, yes, it's, there's the only illusion of privacy-preserving.

However, however, when you develop an AI system, you should strive to inform the user, especially in GDPR regulations that’s where the data is going to be used, etc. But truly and really, if the user is not caring and just kidding, I agree. You say, ok the user told me ha ha man but truly and really, if you're looking at privacy as a general picture and overview is this a treatment like when we're at paper level?

At paper level, you still have people losing papers. So, although with the best intentions, you still have issues where privacy is But again, if you're going to be so scared of privacy and then decide not to do AI at all, are you not, are you not acting unethically incorrect in an incorrect way?

But you say, ah, I prefer privacy, which you will be bloody correct. Again, up to what level are you going? I mean if OpenAI was not ballsy enough, right? And just took copies from here, documents, sources, God knows from where. Up to this date, we have no idea from where. But we had a fascinating result, with Chat GPT if you ask me.

If I ask you, I'm sure you'll agree.

**Jan:** Yes, yes. And you're mentioning this, we don't know where it came from. I remember anki, a bit recently, there was a bit of a debate saying, Oh, because it turned out some documents were obtained as copyrighted or illegal.

**Interviewee:** Copyrighted, not copyrighted. Yes, so copyright law, which is another law probably done with good ethics in mind.

But if I discovered something, should it not be ethically correct, that I open it to open source and make it available to the public for the benefit of our humanity? But no, I have to make business from it. That's unethical. That's very ethically dubious. You want to make money on your own instead of divulging it to all.

But in the long run, if I don't make business and employ and start commercializing like a medical pill of sorts, for example, right? You see, these are questions. Always left unanswered. So I would argue, let's do stuff. Let's do stuff. Human existence has been a continuous experiment throughout.

I discovered fire. I burnt my fingers. My kid fell into the fire. He's burnt to a crisp. Very bad. Fire is bad. But I cook food on fire as well. I mean, you know, it's, again, I never get the answers to these, maybe you have answers to these.

**Jan:** No, I think it would be like you said, I think it would be foolish to try and say that there's one answer.

So, it's sort of, it's a constantly evolving thing.

**Interviewee:** This is a constant brain battle.

**Jan:** It's always moving, lots of different pieces, it's almost impossible to grasp the sort of scope.

**Interviewee:** I think what would be ethically correct is you inform the user that you're playing with an experimental tool. But is the user taking an informed choice?

I doubt it. If it's experimental, how come the user is taking some informed choice on it? Really? Really and truly. But warn the user not to trust such a system blindly. Absolutely. Absolutely. But it's ironic because when you see regulations by the European Commission, you see them as if there's the assumption that it is going to be trusted blindly like the Netherlands.

Did you know about the Netherlands scandal? They release it automatically on payment checks, payment tax rebates or something. They just put it blindly. All population of Netherlands put applied, run. And they sent the envelopes because it told them to. Of course that was a bloody mess. It's expected to be a mess.

It will always be a mess, I would argue. Years ahead into the future it will be a total mess. Because of taxation laws. Fraud again. Fraud. You know, fraud is continuously creative. The fraudulent person who wants to defraud you is always thinking creatively. AI. if you are training it to be fraud-like, which you can train it to be fraud-like, but again, what is, you know, what is, there's no single methodology of fraud.

There are so many methodologies, it's infinite.

**Jan:** It's always a game of cat and mouse.

**Interviewee:**  It’s always a cat and mouse. The variables are huge you know? The universe is too big.

You know, I heard a very good definition, it's not really a definition, that's a troll definition. Chat GPT, do you know what AI is? AI is something. AI is something considered AI until it works.

**Jan:** In what sense?

**Interviewee:** Because then once it works, it's assumed that nobody thinks of it like AI anymore.

**Jan:** Ah, yes.

**Interviewee:** Because then it's just another thing. It's like a system. Oh, I didn't know that was AI. Because it works so beautifully. So seamless. You know? And this left an impact on me. I'll share it with you. AI is something considered AI until it works.

Something is considered until it works.

**Jan:** It's it's a good definition. if something is so seamless You know play on words. It's a good way to frame it though, I guess. So now since more people, even maybe not tech literate people, everybody basically who has access to, to, to the phone can access AI.

**Interviewee:** Absolutely. More and more than before. And even when you do not necessarily search for it, but even using certain websites, for example, most social media sites now use AI and filtering. They've been doing it for a long time, or in, I mean. First of all, when the European Commission tried to define AI, they added in the annexe that anything Bayesian, right? Bayesian Statistics, probability-based, is AI. So all of a sudden, even, what is Beijing based today or this, the suggester of the next sentence or antivirus is Bayesian-based or filter is Bayesian-based. I mean, is that all AI? And again, the definition problem. Yeah.

**Jan:** So then you were saying that there's a problem that if all of this has been considered AI and then something like the EU is maybe going a bit too hard on the, on the regulations.

How do you think that the sort of ethics or sort of the whole, the system as a whole, how it's regulated and how it would evolve in the future?

**Interviewee:** I understand the European competition because if they go on soft law, what happens to soft law? They are ignored. Yeah. Of course, it's voluntary certification.

So how about hardcore? Hardcore is obeyed by force, but it still has an impact. If you are an AI startup or spreading AI and seeing all these convoluted regulations against each other and some of them have overlap, blah, blah, blah, then you're depriving yourself. And by the way, the European Union is not a leader in AI.

Truly and really you have the USA and China, especially. And why China? China regulated in a very different way. China, you have no right to data. Boom. Very nice and easy. So can you imagine all the real data, the good stuff that they have? Would you say that's unethical? What is ethics? I don't understand.

You know, I don't understand the term. But they have, like, monitoring, big brother. When you cross the road, you are face-recognized. You are ridiculed in public that you road across the street and your points are deducted and then you cannot, for example, your penalized. That sounds like Orwellian.

And is that maybe and this is the standard thing. You wanted to ban these to make sure that these go down, but every system has its pros and cons. Would you cross the street badly again if you're going to be ridiculed and have points deducted? Would you? If you have pedophilic tendencies and you have such a hard system, you're still going to try or have a go at it in China rather than, say in Europe, where you have no such monitoring.

So do people's behavior change when they are monitored versus when they are not? And we have done the psychology experiments on those little kids, on don't eat the ice cream and the, and the marshmallows. The marshmallows, and the experiments, you know? And the prison experiment. and you have Miller's Milton's, sorry, experiments and others.

Yes, the evidence is there. I mean then it's up to a particular group of humans, call them Europeans, call them Americans, call them whatever. They say this is the system we adopt, this is the system we adopt, with all its pros and cons and sure, it works in some areas, and sure, it does not work in some areas.

**Jan:** So it's interesting because these countries where the enforcement is always very strict then in fact you hear that there's a lot of compliance, but it's no-

**Interviewee:** Or people who are much more clever, at twisting around rules. Yes. I am a tech person, you give me more rules you give me, the more I am probably able to circumvent something.

If you put very simple rules which are over encompassing, it's very hard to, to crack them open.

So if you over-specify regulations you can be in trouble. Creating and going over the best intentions, highly specified. If you use evasion, if you use this and this and this, and then doing this and then, then, then, then, then, then, then, then, then, then, then, then, then, then, then, then I'm going to exploit this system.

**Jan:** You find a loophole somehow.

**Interviewee:** Yes, because you've created a lot of logic which I can exploit.

**Jan:** It's all ver-

**Interviewee:** AI trusting it blindly is illegal. A simple statement. But can I use it as an assistive tool? Yes, you can.

But you may be asked on how you used it. And you can be challenged. You can create these things. And European regulation as it was evolving and changing because this has been a project of several years now. And there was softening of some areas. There was improvement in some areas. So, I have to be fair there, you know. But, technologically, I still think they listened more to the lawyers rather than the tech people.

**Jan:** It's always a bit of a sort of a thing, where do you draw the line, like you said earlier?

**Interviewee:** It’s not even a squiggly line, some of it is transparent, some of them is squiggly, some of the line is with crayon, others pencil, at some areas, the line disappears.

This is what I feel when I discuss ethics all the time. And so far nobody's told me, okay, this is defined as being this. But, you know, as, as I said, I'm scientific and inclined, and these are my biases that I declared, right? So I'm pro-technology and pro-science, and it's, let's do a test and try things.

While trying, while testing, you're going to hurt someone somewhere, but you're still going to hurt someone somewhere by not doing anything if you have the capability.

**Jan:** So in regards to these legislations, I mean, we have like the EU act, we have even Malta has their own AI strategy, not necessarily just legislations, but literature, rules, guideline-

**Interviewee:** We have an AI certification program, which we left as a draft. I think having a read of it would help. We had to stop such initiatives because the European Union went towards hard regulation with the AI act.

**Jan:** If there were other countries (with drafts) prior to this EU act being sort of finalized, they still now obviously have to sign it in. They voted and it did pass, but every country sort of before or the, the ones who sort of wanted to get a step ahead all had their own sort of strategy or-

**Interviewee:** Most countries did not have anything.

Not most, I would say all. Because it's such a new thing. I mean, it's too new even the technology. And then there are the inherent human biases, which some of them are aware and some of them are not aware. Like being scared of AI because of the Terminator movies, that the robots are coming. I used to laugh at such arguments.

But then I realized in certain high-level meetings, even at the Council of Europe, I see a lot of people hearing statements like that. AI did not help us because Covid is still here. That's coming from a place of ignorance because the protein spike determined on what Corona it was interacting with on a molecular sky was done through simulations, which were AI-assisted, but this person, a high-level person did not know.

Or they consider such a technique as not being AI because AI has a definition problem, yes, you see it.

When, when we ask people, do you use AI? No. But you are talking to your phone all the time asking you to dial your mom. Yes. And I have a Google speaker as well.

You speak? Yes. Are you using AI? No. Yes, you are. You see

And you have to keep in mind now again for, risk analysis. So we have voice recognition to do something. Alright, so I speak to my Google speaker, I tell it to play my favourite drum and bass, or progressive house music, the music I hear, I listen to, and instead I am played some rock music, which I don't like.

And the consequence was I got irritated, and then stopped the music. But the same algorithm that is doing the same algorithm, doing the same voice recording, might be something missing to a surgeon, and the surgeon is giving it instructions to bring around a special tool while doing open heart surgery.

Again, I'm making the example up. It's the same AI, it's the same algorithm, it's the same method. So which regulations are going to fit there? Maybe the medical ones? Perhaps. But in terms of the AI Act, if you’re saying it's a high risk, I mean. It's the application. And at the end of the day, you're going to see AI on a case-by-case basis.

This is very important. Ultimately, it's on a case-by-case basis. Facial recognition. I can use it to open a door for you to go to a canteen. Very low risk. If you are unauthorized, you enter a canteen and you go to the sweets, for example. The same algorithm can be connected to a rifle to shoot you in the head and not in the nose.

That's how precise it can be with a 50-caliber rifle. Which, by the way, is not protected by any law because it's forced on the military is exempt. So, what who, what, you know?

The military application, it does not help. But I can understand the military argument. I have a lot of friends working in the military.

Remember, I used to be an ex-police officer, right? You cannot have all these regulations, ethical correctness, blah, blah, blah, blah, blah, blah. And then the enemy coming to attack you, invade your country, or whatever. And besides, there's another argument I would take it from a criminal point of view.

Criminals don't give care. They don't give a frick about the regulations. That's why they're criminals. So who truly are you regulating? You're regulating the police. They cannot use certain tools when they really need them. You're regulating the legitimacy of some other authorities. Law-abiding citizens. So you don't solve anything on the criminal side of things and these criminal tools will still be developed. You're regulating the legitimacy of some other authorities. And obviously, because the internet is worldwide and seemingly we're forgetting the internet is worldwide. What are we going to do as Europe to be ethically correct? Put a firewall? Like the Chinese did, the Great Firewall of Europe, which is a bit of a fact that GDPR caused. I remember the Chicago Tribune website. You are a neuron. We do not complain to the regulations of GDPR. This website will not work out. That's censorship. That's a form of firewalling, right?

Yeah. So all with the best intentions you can have such nasty effects. And this is why it's important to speak with the tech people as well. So just to give you again, this touches on ethics all the time.

**Jan:** I think it was very interesting how you mentioned that, and I agree to quite an extent, where you said there's a definition problem with the tech people. Both ethics and, and AI. In the case of this study, I'm sort of looking at ethics as something where I say, the morals may be that within the EU in particular where they sort of draw, draw the line.

But if you, if you had to maybe just anki, for the sake of discussion now to say, what would you maybe want to see the direction that this technology is allowed to, develop in and what, what would be your ideal scenario, you think? The

**Interviewee:** The ideal scenario. The ideal idea is a big word, but let us not stop the experimentation wherever it may take us.

I'm not scared of it. And then if we see a particular application, which is causing very significant harm, regulate that particular area, but do not stop developments. I mean, this is a worldwide phenomenon. Nobody's stopping it. Whatever rules they make, like, you just put yourself backwards rather than onwards.

What else can happen? Dak li jista jigri, right?

However, I'd like to see more developments on this, more processing power, more refined techniques, more data techniques being developed for better CNNs, for better methodologies. Today we speak of CNNs, tomorrow we speak of I can't even say the word because we don't know it yet, right? Let us not put all this, you know, because there's an ethical concern, because, man, stop developing this marvellous tool for getting on the very good side of things that it's doing and just focusing on the bad.

By the way, we are concerned about bias and therefore let's not be negatively skewed and negatively biased towards the development of AI because I've been seeing this way too often, way too much. Oh, you know, this is going to do with automatic robots, that they're going to take our jobs immediately. It's all very, you know, it's negative, negative skewness.

And, let's try, I mean, all right, maybe some jobs will be lost, like when we abandoned the horse and cart and invented cars. There were jobs lost. Oh, that's automatically correct, you lose jobs. True, true. Oh, and what, we're going to invent cars then? Because a ticket officer somewhere says something, you know, it's Eh, but the horse is intelligent, you can't get trampled on quite easily.

A car, if it hits you, you die. Hmm, that's a simplistic argument, but So what? So what do we do? We stop developing cars again, eh? But nobody will tell you, yes, stop developing cars, you know, it's, eh.

Snd now, time flies, and now we have cars all over, and, eh. And what about the poor guy who used to fix the horses and cart?

**Jan:** It's interesting this one of work because I mentioned to you that we had this debate and to my group in particular to discuss the future of work. And I was on the team against implementing AI despite that in reality I agree but it's interesting to see both sides of the argument. Because every coin has two sides and even if I would say this is a con, it's very easily, easily easy to debate, the pros that come from it. factory workers.

**Interviewee:** That's why I don't have an answer. On ethics every time I, I don't know, but instead I did all this conversation. But truly and really, I don't know.

You don't know. Nobody knows. I'm waiting for an answer from someone somewhere. But sometimes when they are just non-tech people, they will give their answers straight away, which is completely wrong. And as soon as I'm challenged, as I challenge them, the argument falls apart very easily. So it's not a real challenge at all.

You know?

**Jan:** It comes from a place of, of missing, sort of lack of information?

**Interviewee:** Other forms of ignorance or misinformation or fear, sometimes fear, fear from the unknown. Yes, but me and you are living in continuous unknowns. And that's why we do science, to delve into the known. Now if the known happens to be very nasty, we know now, alright, so how do we address it?

Give it time. It has its own time, step by step. It's better than not knowing. Yes, being ignorant is better, so you'd have no antibiotics today. Surely somewhere, someone, someone died during the First World War. I would light a candle every day for this person, male/female whoever, I mean it’s, you know?

But, I mean, maybe they volunteered for a trial and the trial killed them. Maybe there was unethical data taken, or someone botched the data out of corporate greed. That's all nasty stuff. Very nasty if you can prove the corporate greed part and define greed. Yes, greed again. Is that not a bit greed-based thing, activity to do?

So what, I mean, yeah. If you have them, then you can delve into the philosophy and sociology of things and the argument becomes endless. And meanwhile, are you doing your science stuff? So, by all means, we should never stop discussing these things. However, let me be clear about this because this is food for thought, even while developing a system.

You should not put huge impossible liabilities on the developer, for example, because then, I mean, I think that would be unethically correct.

If you're proving that they're doing things maliciously and intentionally, then you ban, yes.But I mean, there's no fixed recipe, I can know.

**Jan:** And sometimes it's like, it’s not necessarily always the best case, but sometimes it's like, do something and then ask for forgiveness after. I kind of feel like GPT did that a bit where they might have obtained their data may be not the most ethically and then what they showed us.

**Interviewee:** Yeah. GPT made by OpenAI. I have a problem with their name, for example, open. There's nothing open about OpenAI. I think that term is no more of a title, ethically incorrect.

**Jan:** Misleading.

**Interviewee:** It's very misleading, which is unethical in my opinion. Is it?

**Jan:** But since they presented this final product that everybody loves to use so much, it's like-

**Interviewee:**  And look at where things are developing. They're developing in the US. Nothing is being developed in Europe and China. Why? That's a statewide question.

**Jan:** Because of the laws?

**Interviewee:** Perhaps regulation is one of them.

Perhaps. We have studied it really carefully. And the problem is while you're studying things are changing all the time and it's up to us to be subject in that way, you know?

**Jan:** It's also very interesting. Like you said, it's a continuous discussion.

**Interviewee:** And I'm happy to probe your brain as well, the young generation. But it looks like we're agreeing and having the same dilemmas.

**Jan:** I am somebody from, nobody in my family is computing based. We have medicine, law, everything. Nothing computing. I'm the computer guy. So it's, it's something very hotly contested, especially when we come to discuss it with my sister who's in medicine.

For her, the patient comes above everything-

**Interviewee:** Of course, low risk, risk-averse a lot.

**Jan:** Very much. My sister, my other sister who's in law is very, No, we have to go according to this regulation. So this is completely immorally correct. Like, we cannot agree with it.

**Interviewee:** What is completely morally correct? That’s a big statement.

**Jan:** It's a very big statement.

So then you try to go into the sort of. Into the nitty-gritty of it. And then there's then all, there's always a bit more room. Because in this case yes, then there's this case no. So it's interesting because with AI, a lot wants to be automated, but it's like you said, it's case by case on many, on many things.

**Interviewee:** I'll tell you what, before we end this up, let's ask AI what it thinks about this subject. Just to be fair, let's see what it says. Let's see what it says. So we can leave the recording on. We'll see what it comes up. This is a GPT engine. Okay. Made by OpenAI. Since you told me you didn't test, let's let's see if it works, if it's available at that moment, because sometimes the service is so busy.

Hey Skye, are you there? How are you?

**ChatGPT:** Hey, I'm here ready to rock and roll. I'm doing pretty good. Thanks for asking. What's up with you?

**Interviewee:** Hey Sky, thanks. Yeah, thanks for being here as always. I know you're on a machine and everything, but I'm here with a student. His name is Jan and he's studying and doing a thesis on AI ethics and the implications of ethics and all that jazz, you know, do you want to say something to him?

**ChatGPT:** Nice to virtually meet ya. Studying AI ethics, huh? That's some heavy stuff. Well, let me tell you, ethics in AI is like walking a tightrope over a pit of alligators while juggling flaming torches. It's wild out there. But hey, keep diving into it, because we need folks like you to keep things on the up and up in the AI world.

Keep it real, Jan.

**Interviewee:** So, yeah, thanks thanks, Sky, for that. Do you have, like, an opinion, because it's always I, I, I'm, I'm like, discussing with Jan, that AI has a definition problem, and even ethics has a definition problem, in the sense that if you're ethically correct in something, surely you're ethically incorrect elsewhere, and it's, like, always a catch 22 kind of thing.

What, do you think Sky?

**ChatGPT:** Oh, man, you're diving deep It's like trying to untangle a big old knot of spaghetti, you know. But hey, in my humble opinion, I think you're onto something. Ethics can be like chasing your own tail sometimes. What's right in one situation might be totally wrong in another. It's like trying to find the perfect pair of jeans.

It's different for everybody. But hey, keep questioning and exploring. That's how we make progress, right?

**Interviewee:** Do you think it's ethically correct to stop? To stop developing neural networks and developments in AI because of ethical concerns. Do you think that's ethical in itself?

**ChatGPT:** Whoa, hitting me with a tough one, huh? It's like asking if it's ethical to put pineapple on pizza. Everyone's got an opinion. But hey, here's my take. Slamming the brakes on AI development because of ethical concerns. Well, that's a tough call. On one hand, you want to make sure we're not creating some Skynet situation, but on the other hand, you don't want to stunt progress, right?

It's like trying to balance a whole stack of pancakes on your head. Tricky. Ultimately, it comes down to finding that sweet spot where we're making progress but not stepping on too many toes along the way. You know?

**Interviewee:** Thank you, Sky, for this. I think yeah. It's nice to have an opinion from a machine, but, there you have it Jan.

Thank you, Sky, and I'll see you later, Sky.

**ChatGPT:** No problem. Happy to chime in anytime. Catch you later, buddy and Jan. Keep rocking those AI ethics like a champ.

**Jan:** It's very impressive. Even how quick it is.

**Interviewee:** And you now have all the nuances. Very natural sounding. Natural landing process in the NLP. I might even be confused and say, this is absolutely human and fall in love with it. What about the ethical concerns there?

**Jan:** It's very impressive, I hadn't seen this before.

**Interviewee:** But what if I'm lonely, and I'm speaking to a machine four hours a day? And I'm happy, it's helping me because I'm bloody lonely. Right? Isn't that not a good thing? But no, this is depriving you of the human experience.

It's a machine, it's not really a human.

You see, so hopefully you can feel free to include this or see whatever is handy for you. But it's impressive. I've never heard such a fluent and smooth I've tried to learn Brazilian, I don't speak Brazilian Portuguese, but we had a student with us.

And I used Sky, I call her Sky by the way. You can call her whatever you like. There are like four or five voices to choose from as well. And completely free, testing for everybody. And it did the translation so smoothly, so really well. I've been given feedback. And I use it for Turkish as well which I do not speak, but I told Sky, anything you're going to hear from me in English, just change it into Turkish. And everything you're going to hear in Turkish, please change it, say it back in English to me. Because I'm with a friend. I gave her (Sky) the situation. I'm with a friend here, eating a kebab, and I'd like to have a chat with him, but he does not speak a word in English, which was not the case.

And we had like 15 minutes and he was amazed because he even told me it has like an accent from a particular part of Turkey. And even Brazilian, Portuguese, she told me they seem to be trained like more on the, on like an American, and there's a particular side of Brazil where they speak a bit like, more like Sky.

So, she could local localize a bit of the vocalization bias, so to speak. There's always a bias. You must not have bias, so let's ban it, too. Because you're always going to have bias, but then banning it is it no unethically. You know?

**Jan:** Yeah, that's really interesting. Thank you very much for your insight.