Interview #3

**Jan:** So, kif spjeggajt qed nghamel dawn l- interviews to compare the results with the results as well of the systematic review and to see any maybe emerging trends or any identifiable sort of points or if there's any overlap or sort of just to see what's the most key, key concerns raised are so I have a few questions I'd like to ask you. So to start, could you please broadly describe your area of work?

**Interviewee:** So my area of work would be explicitly philosophy. Philosophy of mind, philosophy of language, and so the relationship between language and logic and the objects in the world.

So things related to metasemantics, related to reference and because of this, therefore a significant part of the work I work in is related to AI, but more along the lines of when we speak of cognition, understanding, and reasoning. Whether what we're observing in certain AI systems is more along the lines of, for example, one of the very typical questions is whether we're talking about semantic understanding or just syntactic manipulation.

Pero, given my other background, in theology and in wide areas of philosophy, I also am beginning to work in areas related to transhumanism and kind of philosophical questions that arise on a moral level related to AI both in regards to consciousness and personhood, with regards to societal factors and societal changes that are being affected. Kind of changes that arise due to the implementation, the wider implementation and more widespread implementation of AI systems.

**Jan:** Alright, thank you. So, in your area of work, or in your areas, what are the ethical considerations, sort of, maybe more than in the development and deployment of AI systems, in the use of AI systems?

**Interviewee:** Okay, so I think I would divide them into three. So the first one I suppose is the more abstract and maybe the least pressing at the moment, which are ethical concerns related to the status of AI systems.

So especially with discussion, again, this is more theoretical, right? So discussion related to AGI. If we're going to talk about real cognition, real reasoning, a real understanding that is going on within a particular system, whether therefore certain AI systems should be accorded certain rights.

We speak about the consciousness arising through a system. These are still questions that are still very far away. But obviously, this is one concern which shouldn't be neglected because of the issues there. So what do we mean when a system is conscious?

How's it conscious and if it is conscious, what does this mean? Should it be afforded any rights, et cetera. The other ethical concerns I think relate to, or the way I see it relate, to humanity, to us, but it would be divided into two.

One is the more individualistic concerns. So ethical concerns that concern myself as an individual, as a user, for example, of an AI system. These are generally kind of concerns related to privacy, for example, related to data, use of data related to, for example, certain decisions that are made by AI systems, for example, if certain important financial decisions or health decisions are taken which influence me, which affect me very crucially, which are taken, for example, by AI systems as opposed to by a human operator.

So there are kind of a group of concerns on that level. For example, the question of deception of whether I'm engaging with an AI system or not, issues related to deepfakes, right?

So these are what I would call more singular concerns, meaning concerns that affect the individual as an individual. But then there are also, which at times are a little bit more overlooked, more wider ethical concerns which are more societal in nature. So the introduction of AI may or may not lead to societal changes which might affect individuals in an adverse way.

So, for example, the lack of community, the lack of human contact, right? So, for example, the introduction, hypothetically, right? If we were to introduce AI systems as a replacement for teachers or as a replacement for doctors or as a replacement for care workers in old people's homes. This would influence and affect not only the individuals themselves, but it would affect, I suppose, in a wider sense, society, in the way that we educate the young, in the way that we view elderly people, in the way that we treat these kinds of specific cohorts of society.

In the same way that we might speak about de-skilling in terms of work, there is now also discussion in regard to moral de-skilling.

So, if we offload, there is a risk. The argument goes that if we offload certain functions, certain kinds of practices that we are already undergoing as humans, for example, educating the young. There is a risk that the young will be educated by AI systems, not by human beings and therefore will not have those human interactions with a real human being with a real teacher and will fail to grasp certain values such as solidarity, empathy, respect, and compromise as opposed to dealing with a system that is always, for example, acquiescent, or that is inflexible, or that is unable to show real emotion. So this would be, I think, a very, very quick survey of the ethical concerns.

**Jan:** Thank you. So do you know of any existing frameworks or guidelines in your field that ensure or facilitate the ethical deployment of AI?

**Interviewee:** So we're talking about philosophical?

**Jan:** Yes, yes. Are there any sort of guidelines in that regard when dealing with AI in a philosophical sense to make sure it's done in an ethical way?

**Interviewee:** So there are different approaches.

There are different approaches. There isn't a single monolithic document kind of. Obviously, this is not law, right? So we don't have a law that is being proposed given that certain realities are still new and given that the reality of AI kind of intersects with many other kinds of philosophical trends and also cultural trends.

So for example, I'll give you one example. A lot of the literature on AI is also intersecting with transhumanism. So transhumanism is not explicitly AI. But obviously, it's almost to some degree impossible to speak about one without the other. So there are approaches. There are approaches, for example, being offered through the Catholic Church via, for example, the Catholic Church has a number of documents.

For example, one is Encountering AI. So a recent publication. There are publications, for example, A publication called the ITEC, I T E C Handbook, which offers an operational roadmap for businesses which make use of AI systems in order to do so in an ethical way. So offering a specific way.

That is more, I suppose from an institutional level, as a religious institution, a religious body, which is making use of researchers within philosophy and theology in order to provide in some cases, very practical, pragmatic roadmaps. In other cases, more reflections on what might make sense to do and what it would not make sense to do.

The encountering AI document, for example, touches on the engagement with seemingly personal AI systems, engagement with the use of AI for example in the military or in healthcare and the way we engage with one another. Another risk, for example, I suppose, going back to the previous question would be that we now cease to look at each other as human beings and the totality of the human being, but more as a set of data points.

You are a set of data. Where you've travelled, the data I can scrape off Google, your data I can scrape off from social security, for example, your ID card details, and I reduce you to a set of data. Therefore, you are no longer a human being with a certain mystery, a certain sacredness, a certain importance, moral importance, and moral weight and moral value but you become a set of data points.

So this is a risk that is being challenged over there. There is this literature, as I was mentioning, on moral de-skilling which is very much within the realm of AI. More specifically, there have been attempts, to kind of propose rather than be specific, because this Philosophy. Philosophy is more dealing with the principles and with the overarching direction and the reasoning of things as opposed to the legal nitty gritty.

**Jan:** Okay, thank you. So, you mentioned that there's been some recent literature published even theologic, from the theological side by the church and that there's this new moral de-skilling as a new sort of thing that's being seen and considered. Are there any emerging trends or developments surrounding the ethics in AI that you would identify in particular in these fields?

**Interviewee:** Rather than emerging trends, it is the application of existing ethical frameworks that are now being applied to these new realities. So for example, another line of work that is being pushed within the philosophy is a lot of research with regards to LLMs.

So we already speak of testimonial knowledge. So forget AI for a moment, in epistemology, we speak of testimonial knowledge. Meaning the information I gain from things that you might tell me. Maybe something that you saw yesterday or something that you know. So I'm gaining this information but this information is gained in virtue of, for example, a certain trust that must exist between me and you, right?

It seems that a lot of people, at least people who don't really know what an LLM really is and how it functions, are trusting LLMs in the same way they would trust a human agent. So, what should we say there? How should we value the information a subject gains via LLMs? Should we treat that as testimony or as evidence?

These are two very different things, right? Evidence would be, for example, reading off of a thermometer. Whereas, if you tell me something, I'm not taking it as evidence, right? I'm taking this particular knowledge. You're presenting it to me in such a way, as to tell me this is knowledge. I'm giving you information about the world.

So I suppose there's existing literature and existing kinds of issues and topics which are now being rehashed and seen in a new light. With the emergence of AI.

For example, another area is as I mentioned that we have kind of the whole debate going on, what is consciousness within human agents? Where does it arise from? What gives rise to consciousness? What is the relation between consciousness and thought and my physical body? This whole mind-body problem.

These questions arise again when speaking about AI systems. What is it for a machine to be actually thinking? What is the difference between a simple firing of, in the case of human beings, the firing of neurons in the electrical circuit, the ones and zeros of an electrical circuit and thinking? So there are questions that have already been going on but are now being applied to AI systems.

So across the board and a good number of areas, be it in mind, language, be it in epistemology, where you speak about, for example, testimony, in ethics where we talk about, for example, let's say this whole discussion on emotions or on creativity, what is it to be creative? Is it just making something new that has never existed before or must the process involve something else?

Therefore when something like Dall-E or Midourney provides, you know, a new image that has never existed before. Is that an example of real creativity? So I have all these questions which are rising with respect to AI, which in a sense are old questions that are now being posed anew, afresh under the light of these kinds of new realities that we're experiencing now.

**Jan:** Thank you. So it's like with the technology the questions are always there, but with the technology now reaching a certain level, it's like they're actually being forced to the forefront, kind of?

**Interviewee:** Exactly, I suppose. And, I mean, we have now a piece of technology that at least in very evident ways is becoming more and more human-like. At least on an exterior level. If it's a simple chatbot that speaks to you or something that creates an image. The capacities and the functions that these systems can now perform more and more functions that we have generally thought are functions reserved only for human beings.

And the argument, at least one of the very simple arguments in the past used to be, well, what makes us different from animals or from machines is that we can do these things that these animals or machines cannot do. Now that we're seeing these machines performing these functions, these questions are being posed. So what is it exactly that makes us different or is there nothing else? Is there nothing at all that makes us different? So I suppose we're being now forced in a good way to be more explicit about the delineation between humanity and non-human agents. If it's animal agents, non-human animals or non-human agents in terms of machines.

Yeah, so I suppose emerging technology is not just posing new questions, but it's also making us, forcing us to kind of reformulate old answers we had to old questions. Which is always leading to progress.

**Jan:** So, when using AI or sort of considering the use and development of AI within the field, how is the trade-off between these different ethical principles such as privacy and transparency or accountability, how are they assessed within your field?

**Interviewee:** So this is a very thorny question because I suppose the answer depends on the moral or ethical framework that one adopts, and this is once again a discussion that has been going on for decades or centuries within philosophy. So, you have utilitarian, right? Or what we call consequentialist frames of reasoning which would argue that the way that we should reason between tradeoffs is according to some utility principle, right? In order to maximize benefit or maximize the good or maximize happiness within a society, and even there the question becomes more complicated whether you're going to calculate my own goodness or my own kind of benefit or the benefit of my society or the benefit even of future generations. Who do we include? And if you're going to run with that reasoning certain kinds of trade-offs can be made. So, for example, you can easily see there where it makes sense to trade off privacy in order to safeguard the public safety of the community, for example, from terrorism.

There could be other views, for example, from a more deontological, more Kantian which are more based on kind of certain, certain norms. These norms we should follow in any circumstances, in any case. Act in such a way that you would like everyone else to act, right? So if you don't want anybody to do this to you, very simply put, then you are not permitted to do that yourself. If you don't want anyone prying into your information, then you shouldn't do that at all, at no cost whatsoever. Again, so, a number of people who follow this kind of view, and people who maybe, say, right? And these are debates that are still ongoing. A more preferred view would be what is called kind of a virtuetics kind of, a virtue view.

So virtuetics would be, it's also quite a known view, right, coming back to Aristotle, and the idea here is what is good or what is bad, the way we should act should be in such a way to achieve the flourishing. Both of the individual and of society. So on this particular view, flourishing is understood in terms of.. let me rephrase, the flourishing of the individual and the flourishing of society, the wider society, are never at odds.

They always go together, hand in hand. If society is flourishing, I am flourishing. If I am flourishing, I am also helping society to flourish as well and therefore there is what is considered good or bad or virtue or vice. Something good to do or something bad to do is treated according to different principles, different means according to the end that we hope to achieve. Obviously, this leads to maybe a bigger question, which is what is it to flourish in the first place, in the beginning?

So there is an institute called Humanity 2.0 and they seek to answer this very question. There's a group of philosophers and theologians, along with technical experts and AI experts, who have gathered together, and they explore what is it to flourish in the future. As we're moving forward in different societies, in more multicultural societies, in more, you know technology advanced societies, societies that kind of, you know, where AI plays a greater role, what is it to flourish here?

And when we speak about flourishing, we ask kind of what is the aim, right? When we speak about education, what is the aim of education? For us really to gain a set of, you know, a corpus of knowledge, a body of knowledge? Or is it to form a character? What is the aim of healthcare? What is the aim of the government, of the state? Is it to police? Is it to help us become rich? All these different questions. But when they’re viewed from a flourishing perspective, from this kind of virtue perspective these questions are then situationalized, right? Depending on what, what is it to be good, what is it to flourish, to be better, right? To reach our fulfilment in these new situations, premised on the view and the understanding that, the human beings that existed in ancient Greece or ancient India or ancient China, the human beings that exist today in Malta, England Sri Lanka or Mexico and the human beings that will come tomorrow or in century's time all have the same goal of flourishing. Because in a sense we're the same human beings. Society changes, things around us change, technological advancements change around us, but ultimately what makes us happy, what we seek to be fulfilled remains the same. We must kind of then do this whole process of trying to find what is the best expression of this fulfilment in the particular society that we live in today.

**Jan:** Okay, thank you. So this next question might be a bit difficult because of your sector, but in your sector, how could policymakers or regulatory bodies maybe effectively address these challenges?

**Interviewee:** So I suppose the easy way out to answer that question, and maybe a bit cheeky as well, would be to say that policymakers and lawmakers must consult with philosophers. The point is, and we already see this in other non-AI-related questions, right? From the, the concept of not killing anybody, for example.

Or I don't know, think of any law you might think of, right? The rationale behind the law is because lawmakers or society believes that this thing is good and this is not good, right? Now, we can either leave that up to a majority to decide, or else we might want to say that there is a certain good which even if maybe the majority doesn't accept as right we still think that this should be done or this should not be done.

The majority of the population thinks that we should just kill everyone who's not Maltese. This does not make it a good thing to do. So what is a good thing to do? What is the best thing to do? So putting into law, in order to put something into law, there must first be a certain philosophical exploration, a philosophical understanding.

What is the value we're giving to human life, for example? What is the value that we're giving to society? Are there human beings who have different values, right? For example, should you have less value if you're unable to contribute to the economy? Once you're a pensioner, you have less of a moral value? Should you be treated as a second-class citizen?

In the past women couldn't vote because they were viewed as being less important or having less dignity than men. Slowly there was this realisation that this is actually incorrect and despite the fact that maybe a number of people at the time didn't agree with this, the realisation was that men and women are human beings, and both of them have the same dignity.

This was a philosophical point. There is no, there is no experiment you can do to realize this. So policymakers and lawmakers, in a sense, must engage with the philosophical discussions that are going on in these areas. to ensure that whatever laws, whatever policies are passed, truly aim for the good of society, truly aim for the liberation of society.

We want to live in a society that is free. Now, it could be free in the sense that we're not in prison. It's free also in another sense in that I can speak against the government and I'm not thrown into jail or killed. That's also a type of freedom. It could also be free in a sense by ensuring that I'm not hindered by maybe certain technological realities that are put in place that, for example, tie me down and force me to do certain things or disallow me from doing other things. So if you want us to be truly free, we want the laws to be there to truly protect that freedom in a deep sense, right? Not just in doing whatever you want, and in order to do that, that must be accompanied by philosophical reflection. Because if we don't, then laws will be passed based on maybe for example, purely economic growth, and we already see this in [country] where in different areas we've looked, we've sought exclusively this economic growth and now we're reaping the problems that have arisen from that with regards, for example, the environment and buildings and construction.

So we have to understand kind of what does true freedom, what does true fulfilment look like in a society? Is it just economic growth? Is it just, I don't know security, for example? What is it? And these questions can all be answered if they are accompanied, or answered properly if they're accompanied with the appropriate philosophical exploration.

**Jan:** Thank you. It's interesting that you mentioned this consultation, because in one of the earlier interviews coincidentally I had mentioned the AI Act, which obviously has been passed, and now each country has to adopt and one of the points that the person I was interviewing brought up quite frequently was, he believed there wasn't enough consultation with the relevant sectors.

He felt that it was a bit, written too much of it just by lawyers and regulators, rather than consulting, in his case, the technological aspect. So I think there could even be a point where maybe the, maybe a psychological aspect was also not considered?

**Interviewee:** Yes, I mean so normally when we pass a law, right, we want to safeguard the whole of society, right?

So therefore a law in a sense should reflect, respect and protect to some degree, or at least safeguard the interests of all the stakeholders. So if we're passing a law and we're only consulting lawyers themselves, that's already a problem, right? Because lawyers are supposed to be the instrument of writing the law.

For society, if you are then only consulting technologists, the people who own these companies, or people who run these companies, then we're only kind of protecting and looking after their own interests. They have an interest to make money. They have an interest to sell their product. They have an interest for technological advancement, which isnot necessarily bad, but them in end of themselves alone would be would be bad.

We also therefore have an interest in safeguarding the other stakeholder, the wider society. People who make use of such AI systems. Who are consumers in a sense or users of certain AI systems? Such people might not be as well informed or as, you know, kind of technologically articulate as, you know, people who work in AI companies, but their interests should be safeguarded.

In one of the lectures I gave, right? Maybe you were there or not, I don't know. I kind of used the example of organ donation. We don't allow the sale of organs. It is illegal to sell an organ. It is illegal to sell an organ, even if I have an organ I want to sell, like an extra kidney and you need a kidney and you're willing to pay for it. We agree on a price. So everything is above board. You're not drugging me or, I don't know, right?

Even though both of us agree it is still illegal to sell an organ, and the reason is because if we allow the sale of organs, we'll be commodifying the body, and since this would have severe repercussions for individuals, for example, who are experiencing poverty. Will they be now forced to sell an organ themselves as well? Individuals who have debt. Can the state, can courts, can people who are collecting debt, you know, we already coerce people to, for example, if they're bankrupt, to sell their car or sell their house, right? That's collateral. Could you have a kidney as collateral, right? So, this makes no sense and because of our law that protects that, now we're protecting this kind of tranche of society, the wider society as a whole. Not just poor people, so to speak, but all of us.

So the law reflects in a sense these values and also upholds them and protects them. So we would like any act, any law, any bill that we pass that will reflect the values that we would like to instil in society and in the process of doing that, will uphold them and maintain them. So yeah, there should be a consultation with any law, right?

If it's an AI law, we should consult people who are in the field of AI working on AI. We should consult with people who or at least have representatives of people who make use of AI systems or whether they want to or not. But we should also have kind of consideration with, once again, kind of, I speak of philosophers in a sense because we're doing kind of this our very role, our very work, is this work of reflecting on these principles, on what is it to be good, what is it, what is it to be truly free, what is it to be a truly just society, right?

Speak of equality, for example, in AI. And we already see how much, how fraught the subject is, how, you know, because nobody can really agree on what equality is, right? This philosophy has been doing this process for decades and centuries, right? So yeah, I suppose this consultation is needed, of course, yeah.

**Jan:** Okay, thank you. So, it's important to consult before passing the laws, obviously, and to ensure that the laws address all of these concerns. What would you say are the potential consequences of failing to address these ethical concerns relating to AI in your area of work?

**Interviewee:** So I think a phrase that was, that was used, it was quite interesting, was algocracy, meaning kind of like a dictatorship, so no longer a democracy. So no longer the power is in the people, but the power is in the algorithm. So if we're no longer kind of reflective and questioning on which areas of our life we would like to be kind of decided by AI systems, we would like to offload our own decisions to make life easier, onto AI systems and onto algorithms. The reason there would be that these very vital decisions are no longer taken by human beings who have certain, you know, kind of, the fact that we are all human beings, the fact that there is a certain way that we understand each other, there's a certain common humanity. This might have consequences, of rather than truly liberating people, of enslaving them, of shackling them. There's a very simple example, which doesn't really use AI, but I think it's an example of what, you know, worse things can come about, right?

A few weeks ago, an eight-year-old individual won a court case against a bank in [country], right? Because the bank went through the procedure of, well, you have to fill in this form, you have to fill in this form, you have to fill in this form. They weren't flexible with the realities, with the tailor-made specific realities of that individual, of an 80-year-old who had a bank account that was closed, right?

And all the problems that individual had to face in the process of having his bank account reopened, right? So imagine now, How worse these problems might be, right, potential problems might be, if we are at the mercy of an algorithm. So this could be one problem, a related problem would be this problem of responsibility and tolerability, right?

For example, who's responsible for problems that arise if people get hurt or people die as a result of the proper or improper use of an AI system? Should we offload, for example, should we let autonomous weapons that make use of AI decide by themselves without a human agent ultimately taking the final decision whether to kill or not to kill an individual?

So this is one aspect. The respect is kind of the common, I suppose, I don't know what phrase to use, but maybe a wider dehumanization, right? So if we cease to, if we offload more and more kinds of engagements that we have amongst ourselves onto AI systems we will have this kind of dehumanizing effect of, of no longer being committed, no longer having this kind of social fabric, no longer having the sense of brotherhood, of humanity, of me caring for other people, of being with them and their problems and the troubles of helping others and having others help me, right?

I mentioned this a bit before, two very clear examples would be education and healthcare. In a sense, we have to recognize that healing or education is more than just a question of an output or a final result, but of a process. So these are significant risks that we might have.

**Jan:** Okay, so, in light of, sort of What you've mentioned so far, so the discussions being held...

**Interviewee:** If I might, maybe at a small part, maybe another risk is that we're already seeing, I mean, the use of AI, that there's, there's a great, great, great risk of increasing, severely increasing the inequalities that already exist in society.

This is a very powerful tool. Therefore, if it is in the hands of a few people, they are now wielding this great power. That will as I'm saying increase these and make even more inequalities between those who have and those who do not. So more, a question of justice is there kind of introduced. Sorry.

**Jan:** No worries, no worries thank you ansi. What I want to ask is sort of in light of all of this the concerns, the developments, the discussions, the old questions, the nuances and so on. How would you say that the relationship between ethics and AI is evolving in your industry or will evolve in your industry within the coming years?

**Interviewee:** I suppose with the greater kind of, with the more novel applications of AI right, the new areas or new situations where before AI was on the news. It's kind of, we're now posed with these questions of is this a good idea? Is this not a good idea? Where should we or where should we not make use of AI systems?

That's a question that's already arising, but obviously with the technological advancement, kind of new situations arise. That's already going on, and it will continue going on. I think then the greatest, the greatest I think jump would be kind of the questions I began at the very beginning, which I said at the moment are very abstract and very ahead in the future, like questions related to rights, dignity, personhood that should they, should they not be afforded to AI systems.

The more AI systems become advanced, the more we now ask ourselves, kind of, what do we do now? How do we distinguish? Do we draw a line? Is there a line to draw between what makes us human and in which ways are we different as human subjects, human agents, and in which ways are we not?

**Jan:** All right. Thank you. So last question or the question before the last. So one of the outcomes of this project is that apart from the systematic review, And these interviews with the findings from the systematic review in particular, I sort of, I'm going to be constructing a very simple website that contains this literature or references to this literature.

The scope of it is to present data in different fields in an easy user-friendly way because there are existing resources, but they are not very user-friendly. So that is one of the things I to provide. Are you aware of any existing resources similar to this, which group together literature and legislation regarding ethics and AI?

**Interviewee:** More institutes that work in this area, right? So you could look up the work that they do, right? So Oxford has the Institute for AI Ethics. The University of Edinburgh has the Center for Technomoral Futures. Cambridge has the I think it's called the Future of Intelligence, I think.

Leverhulme, Leverhulme, Future of Intelligence Institute, something like that. So those are there.

**Jan:** So aktar the people sort of working in the fields?

**Interviewee:** Working in the field, hekk hu. The church has the AI research group within the centre of digital culture. The website is not very good, but they have a number of publications that have been published. At least three. And again, which engage kind of on different levels.

Someone to look at in terms of publications would be Paolo Benanti. So Paolo Benanti is a moral theologian but working with the UN committee on AI, or stakeholder committee, high-level stakeholder committee on AI, something like that. Paolo Benanti even spoke at G20 summits and stuff like that.

So that might be interesting, to look at. Yes, I think that would be more or less. There are individuals who have written, Philip Larry, Michael Baggett yeah.

**Jan:** Thank you. So, to conclude, this wasn't something originally that I had thought of, but it came up in the first couple of interviews that I've done, and with your background in psychology, in philosophy, sorry I thought it might be an interesting question as well to see what you would think on it.

So, a point that was raised was that people tend to believe that there is a definition problem, both when we mention AI and when we mention ethics. When we mention, so for instance, AI, what is AI? Or ethics, what do we constitute as ethics? What do we consider as ethics? What would your opinion be on this?

**Interviewee:** I think regarding AI, I don't believe that there is a definition problem.

I think it's very clear, right? So, obviously, then, I mean, what counts as a system that makes use of an AI system, it's clear. Now, obviously, there are different complexities and ever-evolving complexities. At least people who are more in the technical field would tell us the difference between if it is a neural network, or if it's using this process. So I think, regarding AI, there is no definition problem.

Whether the people out there, right, people who are not very well-read in this area, are a bit confused about what it is that… people have told me is Google AI? What does it mean? Or, what is it? Is AI something like Windows?

So you have these very weird questions. People who don't know. So, whether it's a lack of education, that's one thing but I don't think there is a definition problem with AI, and the same, I don't think there is a definition problem with ethics.

The difference, I suppose, with ethics is given that we're no longer in the realm of the empirical, right, but we're in the realm, of the rational.

There's not a question that we can kind of perform a case study, or a scientific experiment or some sort of sociological experiment and discover what is the right ethics or the best ethics, right? This is a kind of continual discussion that goes on. It's not a discussion, kind of, with no progress, there is a lot of progress that goes on, and ultimately, the question that we are trying to answer is, what is it to be truly good? What is the good? What is it to live a good life? What is it for society to be a good society? What is it for society to as I said before, flourish? So, obviously, we'd all agree, right? Do you want to do good? Of course, I want to do good.

Everyone would tell you, of course, I want to do good. But I might do this and say, wait, hold on, that's not really good, right? And we argue on, is this good or is this better? Or is this good and this not good? And these discussions, in a sense, are ever-evolving. Given the new situations, and the new evidence, and the new realities that we find ourselves in.

So it's not that ethics, I think, has a definition problem. We know ethics. What ethics is, is the study of how to act and behave in a moral way, in a way that is good. That's a very good definition. You get 10 points. But what does that mean? And how does that reveal itself in a practical way if I am writing an AI program?

Or if I am running a business? Or if I am driving, right? What is it to be good here? How to do this in a good way? To raise a child, for example, right? How do I raise a child in a good way? These are very profound questions that depend a lot on kind of these, this, this exploration and reflection on what these values, these individual values are.

Which is only a philosophical work. In the sense that no amount of study of neuroscience or physics or computer science can give you an answer to these questions.

**Jan:** Okay. Thank you very much. This was my last question. Thank you for all of the insight.