Yale Talk: Conversations with Peter Salovey
Episode 46: *Navigating the AI Revolution Responsibly*
Guest: *Luciano Floridi, founding director of Yale's Digital Ethics Center and professor in the practice in the Cognitive Science program*
Description: *Professor Luciano Floridi and President Peter Salovey discuss Yale's new Digital Ethics Center to research the governance, ethical, legal, and social implications of digital innovation and technologies like artificial intelligence.*
Publish Date: *March 27, 2024*

FULL TRANSCRIPT

Peter: Hello everyone. I'm Peter Salovey and welcome to Yale Talk. In the past few years, we've seen advances in artificial intelligence, AI, which promise to transform all sectors of society and every aspect of our lives. As a result, Yale faculty members are not only conducting research to advance the technology, but also to inform the responsible use of AI. Here on campus, we are convening multiple disciplines to address the governance, ethical, legal, and social implications of digital innovation. And my guest today is positioning Yale at the forefront of this critical area of knowledge. Professor Luciano Floridi, one of the world's leading ethicists in AI, first gained international renown as a formulator of the philosophy of information. His areas of research also include digital ethics and the philosophy of technology. And last year we formed the Digital Ethics Center around his work. Professor Floridi joined the Yale community from Oxford University, and at the center, he leads a team of researchers in identifying the benefits of digital innovation and in mitigating their risks. Luciano, it gives me great pleasure to welcome you to today's program and to Yale.

Luciano: Thank you. Peter. I'm delighted to be here.

Peter: Great. Thanks so much for joining me on Yale Talk. Before we delve into your work at the Digital Ethics Center, let's start by the contributions you have made to the philosophy of information. For our listeners who may be unfamiliar with the field, maybe you could tell us a little bit about it and what sparked your interest? This is the philosophy of information.

Luciano: Yeah, it's never easy to talk about philosophy briefly, but we don't have much time. So let me start with a bit of an anecdote. I was in Oxford, early-stage career, also known as 'young guy,' and there was this ongoing revolution around us. I mean, internet was academic, military, computer revolution, digital society, information society. And it became clear that this required a lot of expertise from many, many fields, including philosophy, meaning that some of the conceptual, ethical, foundational challenges had to be tackled with the tools that Greek philosophy had provided us. So at some point there was a conference in London and someone asked, can you give us a paper on what you're working on? And I thought, you know what? I can call it, should there be a philosophy of information? And that was how I coined the phrase, and I was successful. There is a need for this conceptual analysis as we move towards a future where this digital revolution is going to make a huge difference, increasingly.

Peter: So was this the paper in 1996?

Luciano: It was even before that, it was '89, so very early. But it was obvious. I mean, anyone with eyes to see, internet was changing our lives. And we were just a few million people online, but you could tell that we had turned a page, and some deeper understanding of the cultural, social transformations was required. Philosophy works like that. Every now and then there is a big boom, and philosophers wake up and say, oh, maybe we should take a look.

Peter: You were prescient because in the 1996 paper that I was referring to, you talk about the ways in which people would use the internet to disinform, to pass along misinformation. And certainly that's a problem we're grappling with now.

Luciano: That is what I recommend to also students, that is what logic does to you. You're going to go in terms of 'if, then.' If this is a new medium of information, then certainly we'll have the same problems that other medium of information have had in the past. So we have this information through newspapers. Surely when this thing called internet will grow, I'm afraid there will be probably misinformation, disinformation problems, and boom. So I wrote that paper, and I have to say that if you read that now, some of the examples are ridiculous. But the analysis, I'm proud to say, is pretty solid.

Peter: Oh very good. Did you predict the political use?

Luciano: What I did not predict, I'm afraid, and shame on me, was the commercial side of the web. Those were years when we thought that this was going to be very academic, perhaps military. There was a lot of applications there. Socially, politically relevant, but it wasn't yet visible. In fact, there was no web. There was no Amazon, no Google, no nothing. So there were no search engines, for anyone is listening here. I mean, there is...

Peter: Nothing to search.

Luciano: Nothing to search. And so it was a bit difficult, and I failed to predict that it would have been also a remarkable commercial enterprise, for good and bad, with all the elements that we find today.

Peter: So now you bring that background, thirty years ago, on the philosophy of information, on the use of the internet to inform, but also to misinform, you bring that all to Yale's new Digital Ethics Center. And how are you anticipating the problems of the future in your work now at the center here at Yale?

Luciano: I think that there are a couple of ways of doing that. I think one is not very exciting, and is the futurologist who comes up and says, oh, in thirty years, in forty years, the world would

be any imagination that you want. I remind students that a century ago, in 1924, it wouldn't be that easy to predict what the world would be like in 1954. Then there is another way, what we do at the center, which is to look at the present and see what actually is already there that is growing and might make a difference. And this reminds me of a phrase I've used more than once, but it's Shakespeare, so it's a classic. When Banquo meets the witches in Macbeth and asks the witches to predict the future, and he has this beautiful phrase which stayed with me, "you can tell which seeds will grow and which will not." And this 'seeds of time,' this is beautiful. So one thing, if I may add, is that we do at the center, is not just to try to anticipate what seeds today will actually grow or not, but also, which one should and should not. And that's the normative side. The ethical side is not just simply looking at the world and thinking, oh, it's like a TV series, see how it ends. But it's more like, are we doing the right thing and are we doing it rightly? Is it the kind of direction that we want to take? That is the more socially-involved, politically-informed kind of work that we do.

Peter: Very good. And back when you were first writing thirty years ago, and fast forwarding to today, it's not just the internet, it's all the social media that uses the internet. Was that something you saw coming? You know, Facebook, and TikTok, and Instagram?

Luciano: It became clear once the web appeared. So once this extra layer on top of the internet started becoming so huge, so difficult to navigate. At the time there was something called AltaVista, which predates Google. I remember the day when I was back again in Oxford. A friend and student with me in Rome sent me a link saying, there's this new thing called Google. I said, oh, that's a strange name. Okay. Let me see, oh, this is powerful, because of course we had Yahoo! So it became clear at that stage that this was going to be the space where we're going to spend increasingly more and more of our lives. And one thing that really turned the page was something that I did for the European Commission. They asked me to work on this project, and they needed a name for this project. And it was the impact of digital technology internet on citizens in Europe in the nineties. And I came up with this word. I said, well, I need something about online, offline, analogue, digital--how about 'onlife' as a single word, and it has become a word now. I'm so surprised because of course, philosophers love to invent new words, but you're not always successful. So this one has taken roots. And I think that that is the experience we have today. This mix, where you can simply say, I am online, I am offline, even just the two of us now. In what way are we enjoying the digital podcast? It's important to understand that there is not either or, but there's a lot of mix.

Peter: Yeah. You know, I have a very good friend who studied cognitive science, essentially with Herb Simon at Carnegie Mellon in his younger days. And I remember in the 1990s, he said to me something like, you know, in another ten years, twenty years, we're going to live our entire lives online. And I remember thinking, this is nuts. Like, the computer is a tool, and I'll use it when I have a task to solve, right? Pay my bills, or maybe we anticipated send somebody a message, or make a spreadsheet, word process a document. And it would be a tool like my calculator, right? We don't talk about living in our calculator. But I didn't see it. I should have, as a psychologist. It was so much more because it's so encompassing and involving.

Luciano: You are completely right. And it's one of the things that was still digesting. I started saying, oh, there's a new medium, a new way of communicating, but through this is really a new space. So instead of having a mass media approach, it's much better to look at it from an environmental perspective. So the kind of issues that we have, for example, pollution these days, now this is the year of elections. Everybody is voting somewhere on this planet for some people. Polluting their environment with AI and bad information, misinformation, well, you understand that much better if you are from an environmental perspective, which means that we're spending quality time in that space. So normally I remind people to say you don't live on a TV, on a radio, on a newspaper, but you live online. So it is important to have clear ideas so that you can also take measures and understand what needs to be done.

Peter: Does virtual reality take it to yet another level?

Luciano: I think it would be complementary. It's always difficult to predict, of course, but if I were to put my money somewhere, it would be more in terms of instead of vertical transformation, say, oh, next revolution is virtual reality, more in terms of horizontal, what is more pieces of this digital revolution, big data, mobile phones, internet, then the web, then artificial intelligence, and one day there will be more computational power. Maybe with quantum computing. These are all pieces of a new era, essentially.

Peter: At the center, what are the trends you're working on now, or the issues, the problems, the challenges?

Luciano: So of course, we don't want to divulge some wonderful secrets before doing the research, but I mean, a huge blocks of impact, and they're quite obvious. One is the whole world of wellbeing, health. There, the digital is making huge, huge impact. So there's a lot going on there, concretely, anticipating collaboration with the computer scientists, with the neuroscientists, with the medical school, and brain surgeons to look at a new design for a very flexible chip that we are producing here at Yale, which would be increasingly flexible in terms of what you can do with this brain implant to tackle some cognitive issues. Now, there is the future. And of course, it generates an enormous amount of philosophical promise. Just to give you one sense, if you put something in my brain, do I own it? And therefore, do I have full control? I mean, after all, it's my brain, my chip. But would you really allow everybody to control and say, reformat, a chip in their brain? Not so sure. Who has access to the data? Should the police, for example? And this is not science fiction, it is already happening. So should the police have access to the data that been collected by that chip? So zero science fiction. I love science fiction, but not when I do philosophy. But very, very concrete and pressing issues. Final example, a company has already gone bust producing these chips. We never thought about what would happen to the chips. And now the judge said, well, you have to remove the chips from the brains because there is no company that can actually maintain those chips. And of course, these are people who have been living with those chips and therefore having a better life because of that. Now, this is now, it's not tomorrow. So one thing that philosophy should help us to do is to prevent, anticipate, and tackle

these kind of problems. I'm not convinced that philosophy comes at the end of the day to make sense of the world. I think it comes at the beginning of the day to produce a better world. That is my philosophy.

Peter: Yeah, yeah. So these are interesting questions because, 'do you own the implanted chip that is in your brain?' I remember philosophical debates about 'do you own your organs?', right? Can you remove your own kidney?

Luciano: And sell them?

Peter: And sell it?

Luciano: Yeah. Yep.

Peter: And I have a very libertarian friend who argues, yes, you should.

Luciano: Yeah. Should I be in charge of my body, and do whatever I want with it?

Peter: There's a philosophy to build on here.

Luciano: Precisely.

They're not precisely the same problems, but there's a continuity here.

Luciano: And that's what philosophy also, the good kind, in my view, which is scientifically informed and rational in its arguments, can inherit from the past good lessons, but without making the mistake of thinking, all I need to do is apply old lessons to new problems. That is a mistake. So the analogy I use here is like, wonderful chapter. Now we need to write our own chapter for the 21st century. We built that chapter on previous chapters, so better learn those lessons, but don't get into the opposite problem of thinking nothing new, more of the same, we have the solutions, all you need to do is apply old philosophy to new problems. There is a big mistake.

Peter: Does the new philosophy build on old philosophy, or is it discretely different?

Luciano: It does. So, you know, for example, the ethical debate in the case of the brain implant, we have all the medical ethics or the bioethics in place. We also have a lot of social understanding. So it really is a matter of growing and building your own perspective. But by using the best, and as much as possible, what the past has refined. After all, we're talking about thousands of years of headaches and head scratching, and by some of the most amazing human beings ever. So surely there are lessons there to be learned. But the opposite is do not believe that this digital evolution is not as radical as the agriculture, as the industrial, and therefore something

that brings up new problems. So synthesizing old problems. And we have also new issues, new problems. More problems, better life for philosophers.

Peter: Ah, right. Full employment.

Luciano: Absolutely.

Peter: That's great. Let's talk a little bit about AI, since everybody's talking about AI these days, particularly generative AI. What are the trends that you're seeing in AI and what are you working on there?

Luciano: I think there's a distraction at the moment, which I like to sort of put aside, which is all the hype, all this sort of new form of intelligence is coming. We are creating new forms of call it agency, a reservoir of abilities to solve problems and take care of and solve tasks, etc. A form of energy, in this case a form of agency, that is of course mindless, not conscious, not intelligent, not nothing. But like my phone playing chess but is extraordinarily powerful and is in the hands of sometimes good people, sometimes not so good people. So the new trends are all about, what are we going to do with this? It's not so much about what digital innovation is going to make here, but the governance of the digital. So what we do with digital innovation, and that's why the center and what Yale is doing, and that's why I'm here, is so important. Understanding that this new form of agency can, for example, help us enormously with climate change, with social problems, with better distribution of resources, with better politics, with better teaching and learning, with improved research. This is like having a fantastic engine in the hands of amazing people. Or, a complete disaster because the same power, the same force, in the wrong hands is going to be more conflicts, more violence, more disinformation. And that is why it's exciting to be at this particular stage in the digital society, because now it's easier to put things on the right foot to send them in the right direction. Fast forward, say, a few decades and things will have started solidifying, we'll have acquired habits. With another analogy, it would be like not having thought about traffic when we invented the car. It's a bit too late now.

Peter: Staying with the Digital Ethics Center, one of the things that you have been promoting is the need for a multidisciplinary approach to these challenges. So science, engineering, public policy, law, economics, business, health and medicine, many others. How does this work in practice? Do people from all these fields come and visit you at the center on Trumbull Street and sit down around a table and philosophize together? How does how does it work?

Luciano: Some do, some do. I have to say that there is a very nice and very welcome flow of people knocking the door. Some are curious, some want to collaborate, some have problems. For example, the brain implant came out of our computer scientist knocking on the door saying, 'We have problems, can you help us?' Most of the times is us looking for the right expertise around a specific problem. It could be, for example, how the insurance sector is using medical devices to screen people, depending on whether or not they qualify or do not qualify for health insurance. Now, this use is a little bit ambiguous. These are tools that have been developed for medical

diagnosis. They haven't been regulated for use by the insurance sector. So what we need is someone who knows a lot about legislation and insurance: the medical people, the people who actually know about policy and industry strategies. And so we go around and say, could you please come and help us? So we have a paper, or there's a problem, and you go and try to collect all the expertise around the table. The thing that I normally recommend is you need to knock the door of your colleagues with a gift. So something that is already kind of clear in terms of problem with a bit of an abstract, maybe a draft, and say we have done all the work we could, we need your help, but the heavy lifting has been done. Could you please add that extra ingredient in this particular dish? Because without that ingredient, this is not going to work. But we've done most of the cooking. And I can see that colleagues say, okay, well fine, so if I have to spend, say, a month and working with you on this, it can be done. But as you said, you need the legal expertise, the technical expertise, the policy and so on. Philosophers used to work by themselves, which is typical of humanities, but I think we need also philosophers who are able to join the table like anyone else now. And that idea of writing, for example, paper or publishing something with another 5 or 6 people...

Peter: With multiple people.

Luciano: Yeah, yeah, which is normal in science, is a bit of a novelty in philosophy. And I think it's a welcome novelty.

Peter: Yeah. I think it has enormous implications for graduate education, how we involve students in our work. I think it makes it more accessible. Philosophers always wrote articles as well as books, but it pushes toward article writing, right?

Luciano: It does.

You know, you have a problem. You try to solve it with collaborators and then move on. I think that style of thinking and publication allows collaboration with students that is much harder in other forms of the humanities.

Luciano: It does. And we're discussing why Yale? Why not elsewhere? It's the right place. And anyone listening, I just remind you, I came here because of this, not the other way around. Yale is the right place for that multidisciplinary approach. Practice not just preach, especially at the undergraduate level or the early-stage career. It really strongly invites that not breaking down the barriers. So we got kids coming from science, from sociology, from international relations, from mass media, from psychology, from philosophy. And they all feel at home. There is no sense, 'this is my subject, you are invited to join me.' But it's really something at the center of the table, and everybody's contributing around this table. That is amazingly rewarding and I cannot recommend that enough. And it's not philosophy pretending to be a science, it's philosophy trying to do the best with as many brains around that problem as you can possibly put.

Peter: Yeah. It's wonderful. I've been talking about a more unified Yale for eleven years, and this is primarily what I meant by it, that we wouldn't be bound by the bricks and mortar walls that separate us into disciplines and schools, but rather would come together across them and find common cause around some interesting problem. And of course, the university is not traditionally set up to support that, right?

Luciano: No, it's not.

Peter: When you're appointing a faculty member, what field are they going to be in? Whose department? Whose so-called slot are you going to use? So you have to build structures that support it. You have to change norms. Can't rely on a traditional organization of the research university, came from the German research universities, but very much influenced what Yale evolved into. But now there's another step we have to take if we're going to do this kind of work.

Luciano: Oh, I agree. I think we can be, if I may put it this way, a little bit less German and a little bit more Greek. You know, the curiosity really knows no boundaries. I remember when I was a kid and when I say, why did you want to do philosophy? I was passionate about mathematics and economics, but then I realized that philosophy was the only discipline that would allow me to go in any department and not feel like I was a stranger. From archaeology to zoology, well everybody puts up with a philosopher, you know, a license to breach boundaries and go anywhere I thought I could explore. And so I thought, you know what, at the end of the day, if I do mathematics or economics, people will not welcome me, say in the modern languages' faculty. But if I'm a philosopher, I can sneak in everywhere and people will say, okay, well, we can put up with this guy. So I think philosophy would be a good place to start. A pinch. Not too much. Like salt. I'm not suggesting that people eat salt, but just a pinch everywhere. I think it helps.

Peter: Yeah, that's so great. Plato's cave didn't have departments.

Luciano: Yes, exactly. So I think that if we enable our students to pursue their interests, and their passions, and their investigative interests, you find them not crossing boundaries. Immediately, they don't have yet that view that this belongs to that department, and this is not within the faculty, or is a different kind of school. They move freely. I think Yale is doing a fantastic job. We can always do better in supporting that. And that's what the center does.

Peter: It's really quite terrific. We've been really focused on Yale, though. We've been focusing internally. Your center also wants to affect the world, and you've done work with the EU and on the UK's Covid app, for example. Maybe you could tell us a little bit about the work that's outwardly facing.

Luciano: Yes, indeed. And that's where philosophy become a little bit translational. So your blue sky, your foundational work, starts having an impact on everyday life. There's a huge need outside. And I keep repeating to everybody, they want to know. They're really curious. I'm not

being overly optimistic. I'm saying your good businessperson, or your good politician. Of course, there's plenty of not so good, but in general, they like to understand more, to design better. When I was at Oxford, I used to spend a lot of time with the government in the UK, of course, the Italian government, or Brussels, the European Union, to shape, for example, the legislation on AI, the so-called AI Act. Now, very happy to tell you that, now I know it's early days, so let us be a little bit careful, that we have started doing the same here. Washington is listening, not just to us, of course, to the general overview of what academia can provide. Practically, that means, for example, advising people on what would have a positive social impact, or what will be ethically aligned in a sort of transatlantic perspective. So imagine you're talking to a senator who, hypothetically, is in charge of AI, and you start thinking, look, keep your mind open to the fact that there's a legislation on the other side that it would be great to have a dialogue. Then there are other forces in the world they are not so positive and not so constructive. So, for example, transatlantic alignment between Washington and Brussels could be something very welcome. And so you try to be the interface between different people, provide the right values, try to highlight how this will bring about better social perspectives. It will be good for the environment. It will be good for society, not just for business, which of course is crucial. And all of a sudden you find yourself contributing to what we said before that big table. There is also a voice for the ethical perspective. And just to be clear, I'm not blind to the pragmatics, to the dirty job that has to be done, the trade offs and so on. But what we want is to have also that extra voice so that there is more of a perspective, and if there's a little space to send things in a better direction, we can take advantage of this. And it happens more regularly than people may think.

Peter: It's interesting. You know, here in New Haven, our mayor speaks often about inclusive growth. How does change and advancement include rather than exclude? How can everyone's lives be improved as the future comes? But just a couple of weeks ago, I saw that the International Monetary Fund cautioned us that AI could actually affect 40% of the jobs around the globe. And what they said, this is their analysis, they thought it would likely worsen inequality. How do we, in looking outward and advising governments and agencies, try to think about technology, in particular AI, as a force for inclusive growth rather than replacement?

Luciano: It's one of the big challenges, inevitably. A fast transformation of society due to this digital innovation, etc. Is going to shake up the job market. So here, for example, is a concrete way in which the work we do at the center here at Yale can help internationally and nationally is what kind of policies can either anticipate and avoid or mitigate also the negative impact of all this. Let me give you an example. Unemployment due to a lack of jobs. Classic. There are just no jobs, so people will not be on the market. Or a misalignment between demand and supply. So there's plenty of jobs, it's just the people don't have the skills to pick up those jobs. Well, clearly the digital revolution is of the second kind. So once you get that, there's not a lot of jobs for everybody doing something kind of solution, but it's education. And again, it's not because we're here and obviously Yale talks about education, but because that is obviously the way in which you mitigate or completely eliminate the problem by saying, what skills do we need to provide to the new generation so that the misalignment doesn't hurt so much? Now, the United States, no matter what other people say, is doing an amazing job here. There are other places, Japan is one

of those, South Korea, not so badly, Europe very badly in this alignment-misalignment. But there is a typical way of doing one, a clear analysis, identifying the problem. Once the problem is clearly identified, find out the solution. Problem is misalignment. We need the new skills. The other thing is, more generally, this revolution is going to provide a lot of benefits in the future, but a lot of growing pain now. And this shows my European origins, I think we should borrow a little bit from the future to mitigate the growing pain now. It's a bit unfair that one generation pays all the costs and the next generation gets all the benefits. And it's no good politics. How do we smooth the curve? How do we make sure that cost and benefits are spread across generations? And that is good legislation, good laws, which ultimately brings us to good politics. There is no replacement for that. Civil society by itself cannot do it. It has to rely on the way in which we take decisions all together. And that is politics. So if politics works, we have a good future. When it doesn't, we are in trouble.

Peter: That gives me food for thought, and a little bit of pause. You know, I like to end when I talk to professors at Yale with a question about students, whether it's interacting with students through teaching in the classroom or through incorporating them into the scholarship going on in a center. You're fairly new to Yale at this point, but how do you anticipate working with undergraduate and graduate students in all of this?

Luciano: Yeah, I've been here for seven-eight months by now or something, and I can tell you that I was astonished. People told me. But I'm still astonished by how research-oriented students are. I know that we describe ourselves as a research university, but it's one thing to read it. Another thing is to live through the experience. It happens almost weekly that one student or a group of students come, they want to see me, and they have an amazing project. I mean, something really extraordinary. For example, yesterday, can we look at the impact that what we need in terms of computational power is having on the energy supply in the United States? And they have already done an analysis. They've spoke to the right agencies. These are undergraduates or master's students. So I know that I'm emphasizing the scholarly research side as opposed to the teaching. But I think the two things here go so much more hand-in-hand than in other places. It's almost like there's a default implicit trust in the ability of early stage or young brain to do more and better than someone my age. So I remind them then, some of them are a bit cautious--can I really put forward a new ideas like, you are never going to be this intelligent ever again in your life, so I can give you some wisdom maybe, and some guidance. But trust me, your brain is way faster than mine at this stage. So let's use the engine and students around us at this stage at Yale, when I say that they can really make a difference with the future, I'm not being optimistic or rosy-kind of glasses. It is their responsibility, too. And I close here with something that my mentor, when I first arrived to Oxford recommended me, your task, Luiciano, is to remove obstacles in front of them, not to tell them where to go, but wherever they're going, just remove the obstacles. And that's what I'm trying to do.

Peter: That's a wonderful place to stop, removing the obstacles for your students and allowing them to think the way you did thirty-five, forty years ago, beginning to imagine this new field. Well, they're imagining new fields as well. Luciano, I want to thank you very much for taking the

time to speak with me today. I believe AI is only beginning to show its value in enhancing and inspiring new kinds of creative work. But as we discussed today, successfully navigating this revolution requires care and forethought. And we have that in abundance at Yale, but especially at the Digital Ethics Center. And I'm excited by the opportunities presented by these emerging tools and forums. And, of course, Luciano, I'm grateful for your work to enhance their potential as a force for good. So delighted that you have chosen Yale, that you have arrived on our shores.

Luciano: Thank you for this invitation, but above all, for the overall invitation to join you here. It wasn't easy, but I'm really very happy to be here.

Peter: We are delighted you are here. To friends and members of the Yale community, thank you for joining me for Yale Talk. Until our next conversation, best wishes and take care.

Peter: The theme music Butterflies and Bees is composed by Yale professor of music and director of university bands, Thomas C. Duffy, and is performed by the Yale Concert Band.

Yale Digital Ethics Center https://dec.yale.edu/