The Politics of AI

SPEAKERS

Keith Bloomfield
Bahareh Heravi
Sanjay Sharma
Embodied Audience
Jo Garde Hansen
Tanaya Guha
Speaker 1
Carolina Bandinelli
Sebastián Lehuedé

RECORDED

6 June 2026 at 17:30

KEYWORDS

AI and politics
critical AI
data justice
AI literacy
AI and society
AI and education
AI and mental health
AI and media
AI and environment
AI and inequality
AI and technology
AI and ethics
AI and regulation
AI and industry
AI and innovation.
AI ethics
transparency
accountability
mental health
automation
regulation
nonprofit AI
neural networks
critical minerals
AI development
societal impact
AI training
public education
AI safety
AI culture.

GUESTS

Jo Garde Hansen
Tanaya Guha

Sebastian Lehuede
Bahareh Heravi

Sanjay Sharma

00:00 KEITH BLOOMFIELD
You're live.

00:00 CAROLINA BANDINELLI
We are live! Welcome everybody to this episode of the Centre for Digital Inquiry TV, CDI-TV, at the University of Warwick. I'm Carolina Bandinelli, and I am realising my childhood dream of being a TV presenter, and that is the main reason you are here today.

You might have noticed, the usual aficionados might have noticed, that Michael Dieter, my co-host and CDI co-director, is not here, and that's why I have taken the chance to stand up and do a proper audition. I'll see it as an audition for the next Eurovision, so I am going to try to be very good at presenting our guests today. I've taken some notes, but that doesn't mean I'll pronounce the names correctly and that I'll be able to properly match names, affiliations and titles, okay? But mistakes, glitches, experiments are really part of our process.

Today our topic is AI and politics. These are a little bit of buzzwords today. Nowadays, everybody talks about AI not necessarily knowing what they are talking about. Definitely I don't know what I am talking about when I talk about AI, and that's why I'll be soon joined by people that know, by people that can help us think through what the hell is artificial intelligence. We even have a computer scientist, so: interdisciplinarity. Tick boxes - just talking to the high-level people at Warwick here for my promotion. Not the Eurovision one, the academic one. Interdisciplinarity. And then we want to explore what are the issues that AI may pose on society, possibly going beyond and ahead and on the side of the either super-apocalyptic or techno-enthusiastic accounts. And then we also want to explore whether AI can be used for good things or to tackle some important societal challenges.

When with Sanjay Sharma, who will soon join me on this stage, we thought about organising this, we really talked about can we start thinking of a 'critical AI' field of studies or areas of thought, and what would it mean? So I don't know if this is a pilot, I don't know if it is a laboratory, but definitely it's not something that is a very fixed, determined and all done, all known. It's something in theory, it's something in progress, and we'll try to think together. That's why [you should] feel free to intervene at any moment, and also... No, I'll tell you this later - building some suspense. Can I invite Sanjay Sharma to join me here? Sanjay is actually the co-organiser of this event... Maybe you should...

03:55 EMBODIED AUDIENCE

[Applause, cheers]

03:55 CAROLINA BANDINELLI

We are rehearsing! We are rehearsing for the big guys up there. And Sanjay Sharma is an associate professor at the Centre for Interdisciplinary Methodologies at Warwick. You work a lot on data and justice. Data justice and AI, ethnicity, AI and racism. You'll tell us more about it in a sec. Maybe you can have a seat. And also, because this idea was a little bit more yours than mine, in that you were the one really saying, "Okay, let's think about a 'critical AI'." So maybe to start with, I ask you - before calling the other guests - super quick, if you can tell us what do you have in mind when you think about 'critical AI'?

04:48 SANJAY SHARMA

First of all, it's great to be here and maybe I'll slightly pause, and... I've rethought the phrase of 'critical AI', because if anyone follows the debate, there's a proliferation of prefixes on AI. Ethical AI, critical AI, responsible AI, and I work in those sorts of fields.

There's an author called Yarden Katz who wrote a book about Artificial Intelligence, and he questions this prefixing of AI because it tends to not question what AI is itself. [1] What's the object of study here? We've had a day of events. When we named the day of events, we actually called it 'Critical Dialogues on AI', which I think is an important distinction. Very quickly, the field of AI is exploding. Every day you can wake up and there's a new model that's released, or more funding or some other amazing technology, and for researchers it's hard to keep up.

05:59 CAROLINA BANDINELLI

Oh yes.

06:01 SANJAY SHARMA

And whilst there's critical dialogues circulating for sure, because there's many harms that AI is creating and producing and reproducing, it's still a fairly marginalised discourse. So even whilst Big Tech corporations talk about ethical AI, it's very much on their terms, and more distinctive voices are marginalised. I'm interested in inserting our voices partly into that debate, and also - which I'm sure we'll talk about later - is to actually question the realities of AI and whether AI is actually creating more problems than it's solving. I'll stop there.

00:00 CAROLINA BANDINELLI

No, but it's so true. There are all of these adjectives just put next to the word AI and sometimes it's not even clear what AI means. Also, I think another problem has to do with disciplinary silos and the struggle to communicate across different disciplines and different vocabularies, so that many people like

me - cultural scholars or cultural sociologists - we talk about AI, but we don't really know what we are talking about, so we need to establish more dialogues with people that know what they're talking about. For instance, Tanaya Guha from the University of Glasgow, who is a... Please, please come join me! Join me.

07:46 EMBODIED AUDIENCE
[Applause]

07:47 CAROLINA BANDINELLI
And a round of applause, yes. Tanya, thank you for being here, and you are a computer scientist. So we are sure about that?

07:58 EMBODIED AUDIENCE
[Laughs]

07:58 TANAYA GUHA
Um, no...

07:59 CAROLINA BANDINELLI
No? You see. Damn.

08:01 TANAYA GUHA
Electrical and computer engineer, if I'm being very specific.

08:05 CAROLINA BANDINELLI

Okay, this is even better from the point of view of crossing disciplinary vocabularies and barriers. You work at the University of Glasgow. You are a... What is your title again?

08:19 TANAYA GUHA

Associate professor, senior lecturer.

08:20 CAROLINA BANDINELLI

Associate professor, senior lecturer. And you are part of the Social AI Group. You are working on a Leverhulme project on machine learning and remembering older women on screen. We were talking about it this morning. But to start with, really think about it as an opening statement: What are we talking about when we are talking about AI? For dummies, please? Because it's true that I don't understand.

08:49 TANAYA GUHA

Oh, that's a great question, I think you need to understand that. AI, very simply, they're basically mathematical models, statistical models, that you can code up in computers, and then you can teach machines to do certain tasks. For example, you can teach a machine to identify between a dog and a cat. So, essentially, they are mathematical models programmed and then doing some tasks for us.

09:22 CAROLINA BANDINELLI

And there are plenty of implications of this apparently simple definition. And talking about the implications, I want to invite the third person working on the Leverhulme project on machine learning and older women on screen, who's Jo got Garde Hansen, so super round of applause for Jo.

09:53 EMBODIED AUDIENCE

[Applause]

09:56 CAROLINA BANDINELLI

Jo is a professor of culture, media and communication at the University of Leeds, where she's also head of school. She loves theme parks - that may or may not be relevant, but it is because it has to do with risk. It has to do with being daring and and also being daring to lead and co-lead on a project that has to do with AI coming from more of a media and cultural background. So can you briefly tell me, or tell us, actually, how you got into that? What is your perspective? What kind of entry point did you find for yourself?

10:37 JO GARDE HANSEN

Well, the entry point for the work I was doing, or have been doing with Tanaya, was women and ageing. They were invisible or largely visible in media, film, television, etc. And we'd done a project - years ago, the AHRC, the Women, Ageing and Media Network that was run from the University of Gloucestershire by Professor Ros Jennings. Myself and others like Kristyn Gorton, who's now at Leeds, we were really interested in making visible older women. I think what came from that was this recognition that there was a tremendous amount of life-logging now of older women in social media. I spoke to Tanaya and we had a coffee and we talked about, "Okay, there's all these older women on screen all the time that are ageing on screen." And we got really interested in the messy and disruptive ways in

which older women were in film and TV. And you talked to me then, "Oh, well, you know about the Geena Davis Inclusion Quotient?" And I thought, "No." I didn't know anything about... I didn't even know what an inclusion quotient was. And it was also discovering that, okay, there are computer scientists who are being really thoughtful about their research. And I had met one(!) and I thought, "Oh my God, how lucky am I? I'll hold on to this person!"

12:01 CAROLINA BANDINELLI

[Laughs] Kind of intellectually get married!

12:03 JO GARDE HANSEN

Yeah! I mean, the thing was in our field, you know, you're doing really quantitative stuff and I guess I'm doing really qualitative stuff. And never the twain meets, you know? And I thought, "Okay, this is such a reach and a bridge between methodologies." But at the same time we had a shared value system, which was really interesting. Values around women and diversity and who gets to speak, and how long they get to speak, and how visible they are in media. And we were just kind of, "Okay, there must be ways of bringing this together."

12:39 CAROLINA BANDINELLI

And I was thinking, in a way... You said your entry point was through the topics of your research: the memory aspect, the gender aspect, the media aspect. And in a way, this AI thing -

there is a good reason why everybody, also from the social sciences and humanities and the arts are thinking, "Okay, I need to engage with AI." It is a buzzword, yes, but it's also a buzzword for good reasons, because it puts us in the position that we have to rethink our disciplines, and we have to rethink our topics and subjects in light of... I don't know if it can be called- it's probably a revolution, or, to say the least, a technological (hence also cultural) shift that of course reproduces certain things - so we shouldn't fall in the in the trap of, "Oh, everything is new" - but definitely it also has an agency as a medium, and it is very pervasive, so we need to understand it. Talking about topics, issues, values actually, that are rethought and are put in dialogue with the technological and cultural implications of AI, let me welcome Bahareh Heravi. Bahareh, I asked before how to pronounce it.

14:13 EMBODIED AUDIENCE
[Applause]

14:13 CAROLINA BANDINELLI
Bahareh is from the University of Surrey, where she's a reader on AI and media, and you are also a BBC BRAID fellow, right? Did I say good? These are my notes.

14:19 BAHAREH HERAVI
All perfect.

14:30 CAROLINA BANDINELLI
Fantastic. So I know Barareh a little less than the other guests today, so full disclosure, but I'm very happy that I'm getting to know you now. Also very nice look.

14:45 BAHAREH HERAVI
Thank you. [Laughs] Thank you.

14:47 CAROLINA BANDINELLI
Because also look matters. We are women on screen. Jokes aside, I really like you, and so I'm happy to ask you, is it correct that you work with AI, democracy, journalism... Is it true?

15:06 BAHAREH HERAVI
Is it true?

15:07 CAROLINA BANDINELLI
Is it true or not?

15:09 BAHAREH HERAVI
Is that the rumour?

15:10 CAROLINA BANDINELLI
This is a bit of a true crime narrative. "Is it true that you're working with AI and democracy?"

15:17 BAHAREH HERAVI
I work on AI and journalism, so media and journalism. My work at the moment is mostly focused on how journalists use AI and as part of the BRAID programme that Sanjay is also part of I'm a fellow at the BBC working on responsible AI literacy for

newsrooms and at the BBC. But in terms of the questions that you're asking, "Are you a computer scientist, and where are you coming from? And what's your story?" It surprises me now to say my background is actually computer science.

15:49 CAROLINA BANDINELLI
Ooh, this is a scoop! We already have a scoop talking about journalism.

15:54 BAHAREH HERAVI
Yeah, and somehow I managed to move into the journalism sphere towards the end of my PhD. I can tell you the stories later, but yeah, that's the area I work on.

16:08 CAROLINA BANDINELLI
Okay, great. And then now last but not least is Sebastián Lehuedé from King's College [London]. Please.

16:18 EMBODIED AUDIENCE
[Applause]

16:21 CAROLINA BANDINELLI
Come next to me. Sebastián is a senior lecturer in AI and society.

16:30 SEBASTIÁN LEHUEDÉ
Not yet, only lecturer.

16:31 CAROLINA BANDINELLI
Only lecturer. I promoted you.

16:33 SEBASTIÁN LEHUEDÉ
We're waiting for the promotion, yeah.

16:35 CAROLINA BANDINELLI
We are ready for- When you promote people, it's good. If you drag them down, then they might hold the grudge on you. So that's fun. That's fantastic. But for me, as far as I'm concerned, you can totally be senior lecturer. I have no problem with that.

16:35 SEBASTIÁN LEHUEDÉ
That's right, tell people at my university.

16:50 CAROLINA BANDINELLI
I'll tell them, I'll tell them. Because your research, I find it super interesting. Sebastián came a couple of weeks ago to give a seminar on his research, and he looks at issues around climate change and environmentalism, looking at both the possibilities and the obstacles that AI might pose. So can you tell me - it's a little bit the same question that I asked to Jo - what was your entry point?

17:23 SEBASTIÁN LEHUEDÉ
Yeah, yeah. Well, thank you for the invitation. I love this format. It's very interesting. I think we're not used to

this, so maybe we look a bit nervous, because as academics we don't do this very often. So thank you for the invitation.

17:34 CAROLINA BANDINELLI
Never, c'mon. I'm a mess. How can you feel like...

17:38 SEBASTIÁN LEHUEDÉ
[Laughs]

17:38 CAROLINA BANDINELLI
I'm giving the example.

17:42 SEBASTIÁN LEHUEDÉ
No, basically, during my PhD - some time ago - I did it about astronomy data in Chile. We can talk about that in more detail, but the governance of astronomy data. I met some indigenous people from the Atacama Desert affected by the construction of some of these observatories, and when I talked to them about astronomy and all those kind of things, they would tell me, "Sebastián, that's important, but we have a more important issue here, which is lithium extraction." Lithium is one of the components that are used to build digital devices, even data centres - which are these buildings key for artificial intelligence - use a lot of

lithium. That took me into the material footprint of digital technologies. After that, I turned to data centres, which are these very important buildings we can talk more about. But always with a focus on communities, people, how they are affected, how they are resisting, how they are negotiating these kind of things. So I look at the environmental impact of artificial intelligence from a more grounded or community perspective, if you want.

18:53 CAROLINA BANDINELLI

And what are you finding out?

18:56 SEBASTIÁN LEHUEDÉ

I'm finding out that the discussion that we're having at the moment about AI is very exclusionary and very elitist.

19:03 CAROLINA BANDINELLI

Okay so these are two things that we can think about. Why?

19:07 SEBASTIÁN LEHUEDÉ

Yeah, yeah. Basically because I think the people who are the most affected by the environmental impact are not part of the discussion usually, and I think they should be, because many of them have very valuable ways of

engaging with nature, with the environment, which I think is something we have to highlight in the current context of climate crisis, rather than undermine. And I think it will make us think AI from a different perspective. Most of the time, they are not AI abolitionists. Even people resisting data centres, it's not that they are totally against data centres inherently, but I think we can come up with a vision of AI that's different if we pay more attention to these groups. And that's only one example, but there are also data workers, so many groups that are not part of the discussion.

19:56 CAROLINA BANDINELLI

This is really good. I was hoping that you would give me a nice cross and bridge for me to open up the discussion, and you did. So thank you very much.

20:06 SEBASTIÁN LEHUEDÉ

You're welcome.

20:06 CAROLINA BANDINELLI

Also, do we want to say ciao to your mama and papa?

20:09 SEBASTIÁN LEHUEDÉ

[Laughs] Yeah.

20:10 CAROLINA BANDINELLI
Because they are connected.Ciao!

20:12 SEBASTIÁN LEHUEDÉ
Hola, hola, hola.

20:14 CAROLINA BANDINELLI
Hola. Bienvenidos.

20:17 SEBASTIÁN LEHUEDÉ
I told Carolina that my parents will be looking at me today, so they're on YouTube at the moment. They don't speak English though, but I hope that there's a live translation or something so they can understand.

20:28 CAROLINA BANDINELLI
Hopefully, hopefully.

20:29 SEBASTIÁN LEHUEDÉ
Using AI, yeah. We can talk about the environmental footprint of that tool.

20:33 CAROLINA BANDINELLI
Exclusionary and elitist. Let's stay with these two words, because on the one hand, there is a sense of, "Okay, AI is everywhere. AI is pervasive. AI has been everywhere for a while

on our smartphones, and it's already very much part of the way we live our lives." So we are all kind of implicated in a way and co-living with AI, but at the same time it is exclusionary and elitist, and the way I'm gonna think about it is elitist because a lot of people use AI, but not a lot of people have AI literacy. So there is this divide, and also exclusionary, because of course there are matters of power implicated in who develops AI for what, where the funding comes from, how is that put to work? And when it comes to AI literacy and matters of power, I think... Bahareh, do you want to add something on it? Maybe you can tell us something more about it? Would you say that's true? When you work about AI literacy, is that the direction you're going in?

21:52 BAHAREH HERAVI

I think I see it in two ways. In a sense, I think AI is democratising the access to information and the access to technology. On the other hand, it is creating bias and ethical issues that we can talk about a little bit more.

But in terms of democratising, what do you mean with that? If we look at a lot of... I'm going to look at both sides of it. If we look at certain levels of privileges that people have had, for example, people who are born with language as a native English speaker. They would have forever written much more easily. As you think, you write, and you can very easily get to certain... You can have good proposals, you can have good papers, you can have good dissertations, and so on. And at some level, a kind of divide is being created here. Now using ChatGPT, people who don't have good English, people who are

coming from other socio-economical backgrounds, have access to certain technologies that could help them create output that could be considered of similar quality. Is it good? Is it bad? Different people will have different views on that. I was in a conference or workshop where a journalist basically said, "I think AI should be used for everything, including information gathering, verification and so on, but not for writing." In my opinion, it's like, why? Because all the other phases of information collection, analysis, verification, and so on, to me are the main parts, while the words are how I'm going to explain that. If I'm good at explaining, I'll do a good job. If I'm bad at explaining, which I am, I will not do a good job. So to me writing may not be considered as the most important thing, while to a journalist or to somebody with a very good level of writing in English and so on, that would be the main output. So we might see this differently, as a tool, as an enhancement, or as a threat, or as something that's destroying the likes of journalism and writing and creativity. But at the same time, certain issues that AI is creating because of who is developing it - who has power in integrating certain decisions and so on in AI systems - are creating other challenges in societies and creating possibly bigger divides that I'm sure other speakers could also talk about.

24 : 26 CAROLINA BANDINELLI

Yeah, exactly. It's interesting what you said, because it's a similar... Some parts of what you said, it's similar to the debate that surrounded the internet and basically any new technology. There is, on the one hand, "Oh my God, we are going to be substituted by it. We're going to be made obsolete. We're going to lose our memory. We're going to lose our cognitive capacity." The first- one of the first, at least in the Western tradition, who did it was Socrates. In the fifth century before Christ in his letter against writing, which was then written by Plato because, indeed, he wouldn't write. And he was like, "Oh, writing is going to be the end of our memory, is going to be the end of our tradition, is going to be the end of our cognitive ability and the ability to form a community." And now to us, it seems quite funny, because we see in writing that authenticity of the very analogue media and the materiality of paper, etc, but that was not the case.

In a way, with AI it's the same thing. It's about the widening the access to means of production. A little bit what happened with [the] internet, with user-generated content, but even with smartphones, if you think about audiovisual production. But then, on the other hand, these issues of power, and the issues of power are the issues of who produces it, and the issues of who produces it, it goes both ways. One is, of course, the economic issue. Who's going to profit by it? But also whose world-vision is reproduced and implemented in the gaze of AI, to use a word from your project. Maybe Sanjay, you can tell us something more about it, about what values are reproduced, and what are the exclusionary patterns that you've noticed in your research?

26:33 SANJAY SHARMA

I think the question of having the parallel discussion around the internet, and how the internet develops, is interesting. Not to collapse the two things together, with AI. The internet partly developed from an academic-military complex.

26:50 CAROLINA BANDINELLI
The ARPANET, no?

26:51 SANJAY SHARMA
Yeah. But then it had certain protocols which were open protocols, which allowed effectively anyone who shared the protocol to be part of the internet. And then the Big Tech companies came along and created walled gardens. Facebook became the internet, for instance. They kind of took over the internet(!) So when you when you talk to kids these days, they don't really think about the internet, they think about apps, for instance, instead, which are very much different from the promise of the internet. And the problem with AI I have at the moment is that it simply reproduces this kind of walled garden effect about who develops it, who controls it. Whilst it allows a certain kind of democratisation of, say, writing, we have very little control about the tools we are using.

The tools are effectively very opaque, and we're kind of enslaved by them. And, you know, they have a certain mystique to them. You type in something and it'll summarise a really boring set of texts and produce a report for you really quickly, and you think, "Wow." But I like to step back and think, "Well, what's the context of what's going on here?" For example, I remember reading an article about working-class students being able to improve their university admission statements using ChatGPT, and the argument was it equalises the field, where the middle-class students had all the cultural capital in the world they can apply, and so it kind of democratises it. On one hand, you think, "Yeah, that's a kind of interesting

use of the tool." But then you think, "Does that really question the structure in the first place of the structured inequality, or is it still just effacing that? Obscuring that?" In many ways, I think a key point for me with AI is it's a technology which obscures many, many different kinds of power relations. So when I think about AI, I don't just think about the interface or the code. I think about it being part of a wider socio-technical assemblage, which is in many ways reproducing many inequalities. Some of the side effects might be that it increases certain elements of... where people can participate more or improve their skills. One of the simplest definitions of technologies: it augments human capacities. And AI certainly does. But the conditions through which it augments our human capacities, for me, in many ways, it's quite problematic. And if we are going to have some kind of AI, I think go back to Sebastián's point: how do we imagine it? Who's it for? Who actually benefits? It seems like the benefits are often side effects, where the real benefits are for Big Tech and big corporations to concentrate power and try to recreate even greater wealth.

30:02 CAROLINA BANDINELLI

You seem to be saying the problem is not necessarily AI in itself, but the socio-cultural and also socio-technical context in which it is produced. And another thing that you said that I really liked - well it resonates with me - is that sort of 'mystique'. There is some kind of magic for the ordinary user, like, "Look at what it did." And so I would like to hear from Tanaya also, because Tanaya will need to leave, so when you have to leave, just leave, okay? No problem at all. So how it is from the viewpoint of those that actually work with AI and create AI and code AI? What is the backstage of this 'mystique'? And how would you describe the way in which you, your colleagues, or

the industry - sorry if I'm not using the right word, I acknowledge that I don't master that jargon - but how do you see the relationship between the technical side, the mathematical model, and all the political implications that are there.

31:23 TANAYA GUHA
Okay.

31:25 CAROLINA BANDINELLI
So my first question is: can we unveil the mystic?

31:30 TANAYA GUHA
Alright, so I'll try my best. Can we go back a little bit in time. AI as we know it today wasn't built in a day, right? So it existed before. I think around 1970s there was a big hype that AI is going to take over and everything, and then it went down because we couldn't go beyond - just using some technical terms, so disclaimer -beyond two layers of neural network.

31:59 CAROLINA BANDINELLI
Beyond..?

32:00 TANAYA GUHA
Two layers of neural network. Now we have millions and trillions of parameters.

32:07 CAROLINA BANDINELLI

And in the 70s, two?

32:11 TANAYA GUHA
Let's say 10s.

32:12 CAROLINA BANDINELLI
10s. And now millions?

32:14 TANAYA GUHA
Trillions.

32:14 CAROLINA BANDINELLI
Trillions. Okay, that's a huge step.

32:18 TANAYA GUHA
Back then, AI also existed but nobody was interested in them.

32:23 CAROLINA BANDINELLI
Okay.

32:23 TANAYA GUHA
There were researchers - they were just doing their work, they were solving problems, they were doing science - but there was not much interest from the corporates or from the big tech companies. And then there were big hardware innovations. Nvidia came in, and if you look at- I actually have a nice slide with a timeline of why deep learning and AI

became so big. It's also associated with hardware development. A lot of chips became very easily available. They became powerful. And that gave the power of compute. And now the computer scientists could do things that they could not do in the 70s, not because they did not have the mathematics, [but] partly because they did not have the hardware to do all this kind of computation. That's the first time AI - or AI machine learning - was shown to be doing something that is meaningful, or at least close to what humans can do. And that's when people started taking interest. So if I talk about my career, in 2007 I took my first machine learning class. I used to explain to everybody what it means. Nobody knew what it was, what it does. There was no interest. And now everybody wants to know why you are not doing machine learning if you're in computer science.

32:24 TANAYA GUHA

But what changed? Why do you think at some point it became interesting?

33:57 TANAYA GUHA

The hardware development, which triggered our power to do lot of compute, that was one thing. Another thing is data being available - because of a huge internet, and, you know, you have devices everywhere. Now you can get data very easily. What you could not get earlier... getting an image was not very easy. Now you have millions of images very quickly. It doesn't cost much. So now you can train both using your hardware and your data quite a lot. And that's when it started

showing some signs of result that could be useful for humans. Before that, nobody was interested. And that's when we started seeing companies getting interested in it, and then commercialization. I often say that when we were doing our PhDs - in early 2000s, early 2010s - that's a time when the companies were picking these things up, and there was still more innovation of AI coming from academia. But now, if you see, it's mostly companies, Big Techs, that dominate the big models, or everything that is being used, how they're going to be used. Even in academia, we end up using their models, which is a very different setting than it used to be 10 years ago.

35:19 CAROLINA BANDINELLI

So what happened? Why is...

35:21 TANAYA GUHA

It started giving some results, and companies got interested. More money was bought in.

35:21 CAROLINA BANDINELLI

Yeah, we scholars are good at things that don't produce results. That's our speciality. [Crosstalk] now that it's working!

35:37 TANAYA GUHA

[Crosstalk] labs to the reality. I mean, that's happened with most technology. It was in the lab for many, many years, and then when it started producing some good result, then it moved more towards commercialization. Big Techs got

interested. And then, in the last... after GPT, actually, everything just exploded.

36:03 CAROLINA BANDINELLI

The big change for the general public, so to speak, was with ChatGPT, right?

36:10 TANAYA GUHA

Yes.

36:10 CAROLINA BANDINELLI

So I'm gonna ask another question to you, because I need to take advantage of the time that we have. What kind of technical innovation did ChatGPT bring about that it made a difference? Because at some point it became everybody's concern. Whereas up until that point, for instance, people like me wouldn't know about neural networks. Now I know. From the technical point of view, what's in there?

36:55 TANAYA GUHA

Again, it has been a long process. Researchers will know they have been doing these kinds of development to generate natural language for many, many years. But suddenly what was different in ChatGPT was the scale at which they were producing. The scale at which they were training these models, which are huge models, which takes a lot of money to train. We cannot basically train them in academia. That scale-up in terms of data and in terms of compute, that changed the results. Earlier, you could not generalise to unseen questions. You can ask ChatGPT any question [and] it gives you an

answer. It may not have seen the questions before. That sort of power comes from scaling-up in terms of data, also in terms of the models they could use because they have huge computational power.

37:53 SEBASTIÁN LEHUEDÉ

I just wanted to add something. I completely agree with that description, but I think that at some point this specific type of AI - let's say machine learning, LLMs like ChatGPT, for example - this form of technical progress also found a business model, right? This scale allowed companies such as Google and Microsoft to create a business model out of it by providing the infrastructure or the cloud that would enable the analysis of these huge amounts of data. So you have researchers in academia doing very valuable research for decades. I'm not part of that discipline, but that's how I understand it happens.

38:35 CAROLINA BANDINELLI

Because you are too young.

38:37 SEBASTIÁN LEHUEDÉ

I hope so(!) But then until there was some technical progress that became profitable... It was when these technology companies got it and

scaled it, and made it public for everyone, let's say.

38:55 TANAYA GUHA

If you look at Google or Facebook, they didn't start as AI companies, right? Google was a search engine, then Facebook had social media, but they all saw the power and the profit that the companies can make. And as you said, the clouds were available from Amazon and basically Nvidia making all these hugely efficient GPUs, that was a big turning point. I think sometimes people miss that - how much hardware development actually sped up the whole process.

39:33 CAROLINA BANDINELLI

You don't need me to talk. It's not school. So if you want to add something at any...

39:41 JO GARDE HANSEN

I wanted, while Tanaya's here really, just for you to say a little bit about how film and television were important for your early work. Because I think part of our project is revealing the origin story and the memories [of] how film and television has been forgotten in that early manual machine-learning work that you were doing. I think it would be a really good opportunity to say a little bit about that.

40:06 TANAYA GUHA

Sure, yeah. It came up in the discussion yesterday. Before that, I did not know these things could be

important to anybody. Like, how [in] earlier days we used to train models, small models. One easy way of getting... I said getting data was not easy, right? One easier way of getting human data was using film clips and movies. We used to use those clips and movies, and painstakingly would draw the bounding box around faces that you see almost in every camera or anything, and they would do it frame by frame, and if it was audio, then we'll do the hand-marking, going seconds, microseconds level sometimes, like, who's speaking, what the gender is, things like that. We used to use a lot of film and television data for that to train those models. Of course, things have changed now, but computer scientists still use a lot of film and TV data to test their models. Even actions, like what actions are being performed. The entire video that you are creating now, we could use it for training models and validating models. Postures, actions, spoken words, accents, everything can be used.

41:31 CAROLINA BANDINELLI

Okay, so I'll let you go. Does someone want to ask a question that only Tanaya could answer? Like a technical question? Vincenzo, for instance, do you have a more... I'm asking Vincenzo because I know that he knows this part of the thing.

42:08 EMBODIED AUDIENCE

One thing that now is on my mind is that... One problem that is, I think, an hour... Because three, five hours ago, the team of OpenAI published an article that they exposed fraudulent groups [which] has publicised, as you said, data models to harm and direct public speech, to harm specific groups, or to monopolise the discussion in the USA. And it's

very interesting, because they like to explain all the movement that they use, and it's not so high technology. The system is created with the help of ChatGPT, with the use of the code. So now one question is, okay, OpenAI has the right to effectively block these profiles, but we have some limitations in this model, like censorship and other type of things. But is that the right way? Technically speaking, we have the moment to put some control in this system, or we have to maybe rethink how they are built, these systems.

43:55 CAROLINA BANDINELLI

I love your Neapolitan English, because I have a Florence English. So I say 'because'. And you have the Neapolitan English, we would be a perfect duo. So basically, you're asking about AI and censorship, and whether these models should include some form of censorship or not?

44:16 EMBODIED AUDIENCE

Yes.

44:16 CAROLINA BANDINELLI

Referring to something that happened a few hours ago.

44:17 EMBODIED AUDIENCE

A few hours ago, it's the paper, but this is happening in at least in one year that some groups use ChatGPT for fraudulent action...

44:19 CAROLINA BANDINELLI

One year... For fraudulent action... Okay, so it's about regulation, and about how do we regulate these mathematical models?

44:39 TANAYA GUHA

I think that's a big question. I think everybody is asking this question. I don't have an answer for that. But as long as... I think I'll repeat what I was saying earlier, that it's not the technology itself, but the users - how we are using it - or the producers that need to be more mindful. I am not sure what exactly was the incident that you were referring to? I don't know the incident.

45:12 CAROLINA BANDINELLI

What was the incident that you're referring to? [Italian]

45:14 EMBODIED AUDIENCE

In particular... It's eight different incidents...

45:26 CAROLINA BANDINELLI

Eight different incidents, saying on the microphone.

45:31 EMBODIED AUDIENCE

Eight different incidents. One known as High Five, where group to the immigration in the USA to mobilise the discussion of immigration in the USA. The other ones like Vixen and Keyhole Panda, it's another group on the Chinese sphere of the internet...

46:01 TANAYA GUHA

What are they doing? That's what I did not understand.

46:05 EMBODIED AUDIENCE

It's not that they don't understand. It's that, okay, that OpenAI has this power to control effectively what the user do. It's okay that they can act directly to some malevolent, effective reaction, okay, this is like...

46:25 CAROLINA BANDINELLI

So it's how OpenAI reacts to potentially fraudulent or misuse.

46:33 EMBODIED AUDIENCE

Yeah, because the fact is that okay, this is a misuse, and...

46:37 CAROLINA BANDINELLI

This is a misuse, okay...

46:39 EMBODIED AUDIENCE

...print a correction, but it's not something that has to be directly on OpenAI. But in this case, it's OpenAI that operates directly on it.

46:54 TANAYA GUHA

As far as I understand from the technique, it's not like how technically we can combat that. You're asking whether, the platform that is providing this,

how they should be held responsible or accountable for that model.

47:09 EMBODIED AUDIENCE

I'm asking if it's okay to put [inaudible] in the model, to get the data from the user.

47:17 TANAYA GUHA

Getting the data from the user? The questions that they are getting from the users?

47:22 CAROLINA BANDINELLI

We are getting lost in translation. [Italian]

47:45 EMBODIED AUDIENCE

[Italian]

47:45 CAROLINA BANDINELLI

[Italian]

47:47 EMBODIED AUDIENCE

[Italian]

47:48 CAROLINA BANDINELLI

Okay, so in these cases, what happened is that OpenAI the company detected the misuse by getting data from the users and triangulating with other available data. So has a company, a tech company, the right to investigate on their users? Or

should we introduce a technical limitation that makes this impossible?

48:20 TANAYA GUHA

Well, I'm not exactly sure. But I think when we sign into OpenAI and we agree to terms and conditions, it might be written there. And they are going to use the data that you put in. If you're not using the temporary one, where they explicitly mention that we are not going to use this data, otherwise they use your data. I think same for Facebook. If you upload your images, they will use your images.

48:50 SANJAY SHARMA

I can just chip in. I'm sure there's other people that'll come in. This opens up a bigger question about the proprietary nature of these platforms. Again, I was talking about how much control we have of this tech, and we have very little control. That's the same for social media platforms. If you roll back a little bit, how these models have been trained, as Tanaya was saying, they literally harvested data indiscriminately, and then they try to make it safe. So they use human reinforcement learning, which actually subjects - often - workers in the Global South to horrific content. This is both for social media and for AI, which is often completely hidden from view. We just see the shiny interfaces that we use in the Global North. And the question of censorship is interesting, because on one level, particularly say with image generation software, when people have looked at the data sets they find child pornography in those data sets.

So my question would be, for these models to operate at scale, for large language models to exist and to be

so powerful, they can only do that by indiscriminately harvesting so much data, because they need that much data to train on. So fundamentally we're working with extremely unsafe systems, which these companies then are trying to make safe. I wouldn't use the word censorship. Because that puts up a whole other debate. I think it's more about [how] these practices, how they've developed, are quite flawed. But because these general models give us the kind of answers we want, we kind of live with that. Yet we all know the taglines of each model. "It might produce inaccurate answers, etc." But that's like a sticking plaster on something more deeply problematic about the very nature of the formation of these systems in the first place. Surprise, surprise - people are going to jailbreak these systems, people are going to misuse them. That's the nature of humans. We do stuff like that. We'll experiment. We'll try to exploit the systems. So I'm not surprised at all.

ChatGPT, interestingly, have got so much attention because they were the first to release these models which were relatively safe, where Google was holding back for a long time because they knew how damaging these models can be. But once a genie is out the bottle, everyone then races ahead, and so they're constantly dealing with the problem of AI safety, and I don't think it can be simply solved. There's no simple technical solution, because there's always contingencies and there's always unknown ways people might interact with these systems. So a lot of these companies have what's called Red Teams, where they actually try to jailbreak their own systems to see where the weaknesses are. But it's a kind of cat and mouse game, unfortunately.

52:04 BAHAREH HERAVI

And I think one thing else that I would add is: what is that red line that OpenAI would consider, or anyone, any person, would consider? When now we can see, in politics in the US, what is freedom of speech? What is censorship? What is safe? What is responsible? I think it's very different if you look at it from different points of view. From different people from different locations in the world, you're going to see things differently. Something that would be considered at this moment in time freedom of speech in the US would be something that is considered discrimination in another country and in a third country would be considered as the rule of law or as the expected behaviour from certain groups of society. So I think it's quite hard, unless there are the likes of international standards groups or open groups that can somehow influence these regulations, it is very hard to know which regulation is suiting which part of the society in which part of the world.

53:10 SEBASTIÁN LEHUEDÉ

I wanted to add something. One recommendation, because I'm reading at the moment a book by Karen Hao, *Empire of AI*. [2] It's great because she's a journalist, she's not an academic, even though it's very deep, the book. But it really tells all the nitty gritty behind OpenAI. It's all about OpenAI, hundreds of pages, how it started... And something that I found

very interesting is that AI safety was super key for the project of OpenAI, that's what they wanted to do. That's why they started as a nonprofit, partly because Elon Musk in particular had this concern. He didn't want any company to release a model that would be too dangerous. So he came up with the idea of creating OpenAI, along with Sam Altman, and they constantly had this debate on when to release ChatGPT, right? Because they knew the risks. But then at some point - well, Elon Musk left - but the team of OpenAI decided that for some reason they were the best positioned to release these models in a responsible or safe way. But it's funny to see the evolution, right? How they have completely abandoned the agenda of AI safety, how they became a for-profit company, and that implied having less safeguards and so on and so forth. But it's a very good book, and because this is a show for non-academics also, I think it's very catchy and compelling to read.

54 : 38 TANAYA GUHA

I just want to add something technical about this. The most basic model that OpenAI, or any GPT these days are based on, is transformer.

54 : 50 CAROLINA BANDINELLI

What is transformer?

54:52 TANAYA GUHA

It's called transformer, which came from Google. And Google actually open-sourced it and they completely have it like free in just an academic lab. And that's the driver behind ChatGPT and all other of these technical models.

55:12 SEBASTIÁN LEHUEDÉ

It's in the name, right? I think the name GPT means generative, something, transformers.

55:18 TANAYA GUHA

Pre-trained.

55:19 SEBASTIÁN LEHUEDÉ

Pre-trained, yeah.

55:22 TANAYA GUHA

Transformers are the basic of all these LLMs and all the big models.

55:32 CAROLINA BANDINELLI

Okay, we have put on the table different areas already. We have the technical part. These are models. There's hardware, there's software. There are technical innovations. There is a socio-economic level, meaning these are products, services. There is an industry. There are people that make a profit from it. And then the moment that some technical innovation

become a widespread means become - we could say, in kind of an STS way - an actor on society. Then the problem of regulation comes in. How do we regulate because law catches up on technical innovations. Something is thrown on society, and then we need to understand how to regulate it. Let's think about all that the European Union did when it comes to data and privacy. So there are these levels. And then I would like to think a little bit more with you today about... Okay, if it is an actor on society, if we try to think about AI or even just ChatGPT as a co-agent of our world. Then how do you find their, its, his, her gaze? And I'm here drawing on the notion, the idea, of the AI gaze Jo, Sanjay and Tanaya are developing in their project. What do you think- Maybe Jo, you can answer this question or start throwing... It's not about answering, right? But why do we need to think about this AI gaze, because a lot of the time we think about how we see AI, but how are we seen by AI?

57:52 JO GARDE HANSEN

Yeah, I think our project has been really about accepting that, as Tanaya said, the models are pre-trained, and there's very little we're able to do about that in terms of computing power. But we can do something methodological, in the sense that we can bring some of that model and iterate with it, slow it down, make things a bit messier. In our work, by showing it or asking it to analyse film and television - and in particular representations of older women - we have noticed its sense that it's not always accurate, and it challenges the whole concept of accuracy in relation to trying to identify the different ages and the same person ageing in a film, and the creative aspects of ageing, de-ageing, prosthetic ageing, and then

perceptions of age, and then maybe even a concept of screen age, as it tries to reference other forms of media and scrape through all the life-logging and the other films that it's ever looked at. And so what we've tried to do is observe and think about the machine-learning 'looking at' these representations of ageing at scale, and then at the same time thinking about that as an intervention in itself.

But also I guess as academics who have closely read objects of study. Say your object of study is a film or television text. We've built entire departments of research around reading a text really closely and deeply in terms of its history, and that's a really important cultural and critical way of thinking - slowly, deeply contextually - and then trying to bring that methodological way of thinking alongside something that is doing kind of the opposite of all those things. And so working together with a non-human actor in terms of our team, and then critically reflecting on each iteration as well. We chose those film and television texts not only because, as Tanaya said, they're part of the hidden history of machine learning as well, because you couldn't simply just always go out without consent at the time and collect loads of images of people. So film and TV, rightly or wrongly in relation to IP, that's how they were modelling. We've really thought about film and TV as something that really challenges machine learning itself in its seeing, while at the same time, through Sanjay's work, revealing that the critical data justice aspects of this - in relation to what those models hold

within them - what they can't forget. They can't forget what they were trained on, and we can't extract that out. It has learned a lot of things around the kinds of faces that have been out in the world for the last however-many years that it can scrape from, and the predominance of those faces, or whatever they are. So we're not able to get it to unremember all that work. But what else can we do in revealing that can't-forgetting.

1:01:23 CAROLINA BANDINELLI

This is sort of going back to what Sebastián was saying before, the exclusionary patterns that are somehow embedded in the AI. Because there is a little bit of an idea that ChatGPT or AI in general 'knows the truth' because it is unbiased, because it is about data, and the myth of the 'raw data' is still very much there. Leopoldina Fortunati, who is an Italian scholar, did research lately on young people. I don't know if it is published or not, because she told me over a pizza in Naples, so I don't know. But she did this research with undergraduate students, and for almost the majority of them, they consider ChatGPT cleverer than themselves. Because it is unbiased, because it knows the truth, because intelligence is increasingly associated with computational capability. And so I very much like this concept of accuracy, because we we tend to think about efficiency, and the whole digital technology has very much produced an ideology of efficiency whereby everything that is efficient is normatively better. Because efficiency, it's all we need. But then what about accuracy?

So in what you're saying, and your project is showing, well, a machine learning model can be very efficient at recognising and categorising older women on screen - definitely more efficient than if you put me doing that manually, as Tanaya has said - but that doesn't mean it is accurate, and that has to do with the fact that the AI gaze reproduces the gaze of those

that produce it, and also reproduces the gaze embedded in the data. There's not such a thing as a raw datum, or at least so I remember from my exam on the philosophy of science. And when it comes to relationality.... So it's also about, okay, if AI is a co-agent, who is working with that? Who gets to collaborate? And so I would like to go back to what you were saying before, Sebastián, about the exclusionary pathway, and what what subjects are indeed excluded? And what can we gain from widening, in a way, the collaborators so that this AI thing can learn also from other people? Not the usual suspects.

1:04:22 SEBASTIÁN LEHUEDÉ

Yeah, yeah. That's a very good question. I think you're pointing at my research in Colombia, I guess. But I wanted to say, just to... structure a little bit the conversation. I think your research looks a bit more at the outputs of AI, so to speak, like generative AI, the kind of texts or images that it can give you. In my research, I have looked more at the inputs that AI uses to create those outputs in the first place. So that's why I will talk about infrastructure like data centres or minerals that are used to develop AI, or even data workers in the Global South countries who are paid very, very low wages and also have exploitative conditions, [who] label the data that then AI machine learning systems use to train their models, blah, blah, blah.

However, in my research, I have looked at the environmental dimension of AI, which comprises two

things. On the one hand, the environmental footprint of AI. Basically all the resources or things that it uses in order to work or develop, comprising minerals but also energy, water and so on. And on the other hand, I have also looked at how AI can be used for sustainability. Because if you talk to people in the field about the environmental footprint of AI, some of them will tell you - this is something I have heard - "Oh, it's okay. It uses a lot of minerals, a lot of water, data centres, for example, energy. But then it also has like good uses, right? You can use it for sustainability." And you see that idea replicated in many international organisations and so on, saying that we need AI for climate modelling, for example, to fix the climate, as Sam Altman from OpenAI would say. So basically I've been trying to look at both sides of the debate. Something interesting that I have found - in February, I actually travelled to Colombia because I was interested in projects that might incorporate more people, more voices in these kind of uses of AI for sustainability. I identified like a very interesting project in the High Andes of Colombia, of scientists, mainly biologists working on plant conservation, but that were also increasingly incorporating peasants or farm workers. It's difficult to

translate into English. Campesinos, yeah, who were increasingly participating in this project. I found it so interesting because I don't want to say that AI is always bad, that everything done with AI is problematic for different reasons. I also want to see people trying to shape the development of AI in a different direction that incorporates more voices. So that's one example. And even though these peasants have not been super involved in the technical dimension of AI for different reasons, they are part of the project and they are shaping it in different ways. And you can see that the attitude of people using AI when you incorporate more voices is way more humble, right? Usually programmers - or sometimes computer scientists, sorry that the person who is a computer scientist is not here - but they tend to overstate, sometimes, the capabilities of AI or what it can do. But when you incorporate more voices, you become more humble and more understanding of the complexities involved. Sorry, if it was too long.

1:07:58 CAROLINA BANDINELLI
Improvise!

1:07:58 SANJAY SHARMA

I'd like to just follow up on that. Is it the case then, that... Are they developing more local models which are specific to certain social problems they're facing, or political economic problems which might help them? A lot of the problem I have with AI is these general models which cost huge amount of money to produce, are unsafe, exploit the Global South in terms of the training and labelling of the training data. So when you when you talk about alternative voices, does that also include alternative ways of building the models which are more maybe more specific, more local to certain problems?

1:08:46 SEBASTIÁN LEHUEDÉ

Yeah, yeah. It's a completely different logic and scale. You're right in that sense. I'm actually comparing this small-scale project with a project by Microsoft called Guacamaya that's seeking to tackle deforestation in the Amazon, but that's a way larger-scale project that looks at the entire Amazon, Colombian Amazon at least. Whereas this other project, the smaller one that I was mentioning, mainly it's trying to focus on one more specific problem. I think that also gets us into another discussion, which is very interesting - foregrounded by a colleague, Alison Powell - which is, it's interesting how sometimes we use the 'mainstream' and 'alternative' models of framing different projects. I think it would be interesting for us as researchers to mess the field [up] a little bit and try to come up with

projects that are very attuned to the local context, but that also have the ambition to be more large-scale eventually, and on the other hand larger-scale projects that can also be attuned to the local context. Because right now we have this very comfortable situation. You have large-scale projects by companies, and we criticise them, and they're like the bad actors. And on the other hand, we have small, beautiful projects, and it's very easy to foreground them as very nice and a model to follow. But then what if we mess this up? I do think that we need large-scale AI, maybe not generative AI for all users. But if you talk to climate scientists, for example, we pretty much need them, right? Because the scale of climate change is so big. Even if we degrow our economies tomorrow, we will need to remove carbon from the atmosphere anyway, with technology. So it's highly likely that we're going to need AI. The question is how we deploy those AI, or AI-driven or AI-supported projects, at scale in a way that's just and ethical, but I like your question because I think that's one of the big questions of the moment.

1:10:44 JO GARDE HANSEN

There's something else that we haven't really mentioned, but there's something about the

talent and the kinds of people - like Tanaya, who's not here - staying in the public sector, staying in the university sector and continuing to produce, or at least work with, socially just projects. What we found with our project is that it's really difficult to get really talented researchers to stay in universities. AI researchers and machine learning, they're highly technical work. They get taken up by industry and they want to go into industry. It pays a lot of money. That's the other part of it. To what extent do we, as universities, support financially as governments really invest in this kind of research, and some of it has those origins in really fair, socially-just uses. It wasn't highly corporatized at that time in those early uses. I do think that's one of the significant problems we're facing as researchers on these either big interdisciplinary projects or public third-sector public university projects is that it's really hard to get the talent because they're not attracted to stay in universities or be part of university projects, and they're moving around all the time or they go straight into industry where it pays a lot more money.

1:12:15 BAHAREH HERAVI

I think I would echo what you said. I think even the ones who are interested... At some level it is becoming more and more difficult as life is basically becoming more expensive and the universities are not able, or they are not supported, to support that. It's the same with the journalists. Journalists are also some of the lower paid part of society. These two groups of

people, academics and journalists, are those who are critically looking at these problems, write about them, research about them. And if, in a sense, they're not funded to an extent that is comparable to the funding in the industry, they are going to go to the industry, and we are going to lose more and more this critical aspect of this AI problem that we are looking at now.

1:13:08 CAROLINA BANDINELLI

So again, it's a problem of who is involved and to what extent [they're] involved. Who is an object of AI, and who is a subject of AI? Who is produced and who is more a producer? And so it's about the web of power, to use a Foucauldian notion because I read Foucault in my PhD and never get out of that moment. But I would like now to hear some more voices from our either embodied or disembodied audience here. Please.

1:13:51 EMBODIED AUDIENCE

I was wondering, because I know you mentioned you work with the BBC, but then really this is a question to all of you, which is just the use of AI in kids' media. Because we've seen throughout the years that a lot of the more interesting and innovative kids' media, or especially kids' educational media, has been, on one hand, you could say it has functioned - insofar as media technologies are seen as harmful - as harm reduction. Like, okay, if you're going to be watching TV and you have young kids it's better to watch Sesame Street. But okay, also the whole idea of Sesame Street was like: kids remember ads that come on TV; can't we use this to teach them the alphabet or something like that? And of course, some of these things later had provable educational successes. And of course, questions

about educational technology always relate to like: what is childhood for? What is education for? Is the idea of playing Baby Mozart to your kids that they're gonna wake up and they're gonna know how to use all of these things as marketable skills? But I guess one question is: have there been major attempts by the BBC or by other people to use AI in educational, especially young educational, young kids' media, in any reasonable context? Has this been able to be done in a way that the AI isn't just used for cheating on tests or something like that, which is the bane of all of us here? Have there been any efforts in that regard? Is it okay at all to mix kids and AI? Is this inherently, ethically, a big problem?

1:15:17 BAHAREH HERAVI

Is that question for me? I have one small or short answer that I know that the BBC has been working on an AI tutor project to help with the - I'm not sure what age of children, but children at school - with some of their work. But at the same time, I don't know, I don't have children, so I'm not very familiar with the children's lingo and the programmes that they watch and the ways they learn, so maybe somebody who's more familiar can respond.

1:15:48 CAROLINA BANDINELLI

But can you say something more about this AI children tutor?

1:15:53 BAHAREH HERAVI

It's basically [that] from the content the students learn - maybe at school or from the programmes that are available to

them - they can ask questions and the AI tutor can help them with their work. That's the level I know of the project, but there is some information available online, I think, about the project if you look into it. But there are a lot of universities, which is something that I know more about, or even schools, they're all looking into these AI tutors and how to help the students with tutors. Also that comes up as a discussion in less-privileged parts of society, that they're not able to get private tutors and so on, how they can get help from AI tutors. Is it a good thing? Is it a bad thing? Is the content suited for them or not? That is not something that... Maybe Sanjay has something.

1:16:42 SANJAY SHARMA

Yeah, I've got a daughter of school age, so I see some of the tech that she's bringing home in terms of software her school is using. For example, there's a programme called CENTURY Tech, which helps in English and maths, for instance. I looked it up and it's a black box in terms of its testability, how efficient it is, or how much does it improve a child's learning. Because basically it's a diagnostic tool. A child enters work and then it figures out where the weak points are, then it'll point the child [to], "Oh, you need to improve on this." That's this basic model. Sounds really good in principle. In practice, I don't know whether it actually works or not. It's not my field or area, but there's a whole industry around EduTech that's developed. Massively profitable. Some of the research that's been done on it basically suggests that there's no evidence it improves children's learning. Even bringing a tablet

into a classroom does not necessarily improve children's learning, partly because - against the question of labour - what happens is they can invest less in face-to-face teaching and let tools then replace a teacher in the classroom, so there's less contact for kids. And I think at that age, you need teachers. There's a whole issue about Foucault and education, and education is a prison. We don't need to get into that debate. But technology necessarily isn't a solution, but it's sold as a solution to help kids, and there's very little evidence - empirical evidence - that actually that's working. But it's kind of scary because we're all sold, by EduTech, that it's effective.

1:18:34 BAHAREH HERAVI

[I'll] just maybe add something. Again, I don't know about children. But from my interviews at the BBC - I'm talking to different journalists or creative people at the BBC on how they use AI and how they engage with AI training and AI literacy, and how they upskill themselves, or whether or not they feel like they need to learn AI. One of the things that I've noticed is how the younger journalists learn - and like to learn - is quite different. The older journalists maybe like to have somebody in the room or online, but live. Or even some content that they can go and look into. More structured, more formulated. Some of the younger journalists that I've talked to, they're not really interested in that. So I think the way of learning between children, younger people, and it varies in age - the expectation is a little bit different, and

they don't really have a lot of interest to take one course from beginning to end. They've told me, "We have learned it ourselves and from YouTube and within our own time." So just as a comment to Sanjay, that maybe children have different expectations. Whether or not it makes education better or has a better result, I don't know, but I think they're just different tools.

1:19:49 CAROLINA BANDINELLI

I think it also has to do with the degree of intimacy that we have with the machine, with technology itself. That's the idea I'm flirting with lately, which is this idea of artificial intimacies. Because it looks like the ways in which people use AI, it is quite intimate indeed. So we ask them - I'll give it a "them" - to be a tutor. It's a tutor, it's a master signifier in the life of a kid, but also of an adult. A tutor, a mentor. And then another field in which generative AI, ChatGPT, is really having a huge impact is mental health. Now, for the NHS, I was told by someone I trust - I didn't fact check [laughs] - but the entry level is an AI therapist. And then you can 'work up' to a psychoanalyst-ish, but that really takes months, so if you don't kill yourself before then you'll get there, but the entry level is that. And again, on the one hand, it's kind of scary. I mean, I've been over 10 years in Lacanian psychoanalysis, so for me Lacanian psychoanalysis is really different from an AI, because they don't say anything, and ChatGPT says a lot. So they're very, very different.

1:21:24 EMBODIED AUDIENCE

[Laughs]

1:21:25 CAROLINA BANDINELLI

But also, on the other hand, there's a question of access and there's a question of cost, because who can afford - how it's called? - mental health support ? And can a national health system under increasing financial pressure offer? What kind of salary can it offer? But then I think another level is also what you said - for me, it may be weird and even uncanny to think of having ChatGPT as a shrink, but maybe for another person that is not necessarily true, so it's also how we relate to AI as a subject, and what we ask. We did a project with with the masters students this year on love and AI, sex and AI, and most of them were looking at educational uses, so sex therapists or a sex educator, of a 'love educator'. And that also points at the lack of so-called affective and sexual education. I think that, again, the AI gaze tells something about the state of society, even before. And on the other hand, by the way different people look at AI, we can unveil a lot of dynamics. So this idea of artificial intimacies is the one that goes around my mind. Do we have other voices? Other question? David.

1:23:14 EMBODIED AUDIENCE

I'm not sure I had a question, but I'll formulate one. I thought this was a really interesting discussion. Thank you. I was struck earlier before, when Tanaya was giving her account of the history of AI. It felt almost synonymous with - or was, essentially - a kind of history of computing power. That it was kind of about the development of computing power. And I'm wondering generally - and she touched on this - whether there's a moment... How can we distinguish between... What's the distinctive 'AI' bit of AI, which distinguishes it from just the powerful, logistical, instrumental reason of computing power? I seem to remember in the history of AI, in the imaginaries of AI from science fiction, etc, but also I think immanent in some things like the Turing test, there was a moment when machines would become sentient in some way - touching on some of the things you've just been talking about - that

there'd be a sense of the AI taking a role, a kind of benign role, in looking after people in various ways, in looking after us, in solving problems for us. Where did that go? Where did that promise, that sort of utopian promise, go? And I guess I wonder, is the sort of things that we're seeing now in and around the problems around ChatGPT - the hallucinations of AI, the fact of the mass production of basically incorrect information - what the implications of that are? Given that it's established itself as this particular kind of trustworthy voice within various forms of public discourse? That wasn't bad, was it, for being sprung on me?

1:25:15 SEBASTIÁN LEHUEDÉ

Thank you. That's a very interesting question about what happened with this, or what's particular about AI - what makes it intelligent compared to regular computing. I always think that people, different groups, within the field will give different answers. I would say that the field, if you look at the history, has been split between two big groups called symbolic AI or connectionist AI. If you are a symbolic AI person, you will say that what's intelligent about it is that, as humans, computers can, for example, learn high-level rules and apply to different contexts. So you don't have to explain or describe the entire world, but just give the computer some insights, and then it's going to be able to deploy it in other places. For symbolic AI people, that's what will make it intelligent. On the other hand - which is I think more like the hegemony, if

you want, within AI at the moment - is the connectionist paradigm. For that paradigm, learning is not - or being intelligent is not - about having high-level statements and then applying it in the world, but rather learning from observing, from having data and being able to create categories without any previous input. The model that we have right now, for example, the idea is that you give computers a lot of data - which is why we were talking about data is so important, the amount of data that we have at the moment - and it's able, without any human telling them anything, to create categories. So you will give the computer images of cats and dogs, and it's going to be able to tell you "these seem to be two different kind of animals" and eventually describe them without much human input or with little human input.

The thing about data and computing - why it's so important - is because first, the amount of data that you have to give to the computer, and second, the analysis of each pixel of that image requires a lot of computing power. So that's why computing is so important for the paradigm that we have now of AI. Now something I wanted to say: that's not the only type of AI that we can have. One of the problems with the connectionist paradigm is that it

has some issues with accuracy, as you were saying, right? Hallucinations. Some people have been saying maybe we can experiment and try to come up with a mix between symbolic AI or connectionist AI or try a different kind of path. But the problem - and I'm going to connect this with another thing we've been discussing - is that the connectionist model is very profitable because the same company that provides you with AI - Microsoft, for example, either directly or through its partner OpenAI - is the same one that provides you with cloud computing, and storing and processing data is very expensive, so that's how they make money, right? Basically telling you "use more AI", and the more AI you use, the more dependent you become on their infrastructure. So the problem with not having many AI experts in academia is that, right now, most AI research is shaped by companies which have profit motives, which is fine, but we're missing research that might try to change the paradigm, maybe one that's less profitable, but better for society in different ways. I don't know if that answers your question.

1:28:45 JO GARDE HANSEN

I don't think we have lost it entirely, because I think that we've talked a lot, and we do talk a

lot, about the harms. But the reverse of that is the care and the caring modalities. And I think perhaps we haven't lost it in the cultures of caring that AI replicates that we want access to. We all want access to our PA or concierge or the butler service or the air stewards, kind of rhetoric of "I will come and help you, and it's really turbulent right now, but don't worry, look at my calm exterior and I'll get everything for you." So there's that service model of how it's giving us a culture around caring, and that might be through an entry-level care that is free, and then you get up to the real caring that costs. But I think there's something in there about... we are still getting it discursively when we go into ChatGPT or Claude or any of them, it's, "I'll care in an obsequious way and serve you very quickly." And you can actually see it written, it's always apologising for something. Whenever I ask it a question, it's usually apologising. So I think there's something about that that we want, and we need that intimacy of someone talking to us in that really super caring way that AI is giving us. So some of that promise is there just in the culture of it.

01:30:45 SANJAY SHARMA

But I guess what I would ask is, what's missing in our society that ChatGPT can fulfil in a certain kind of way? So there's that issue. You probably think I'm completely anti-AI - I'm not. But say in the case of mental health, if I was ever to use a language model, I want to know how it's been trained, and I would never want to try to use a general model because we have no idea about what kind of answers it might produce.

There's been some very remarkable cases of kids asking AI about their parents, and the AI eventually saying you need to kill your parents, that will solve your problem. Maybe it's very Freudian, I don't know. But literally that can happen. Okay, they might be edge cases, so often they're dismissed. But it does go back to the questions of how responsible are these models, and how are they being rolled out? Because they're seen as more efficient, it's cheaper than employing real therapists. That's my question. This automation of society and acceleration of automating all kinds of things. There's one hand like, say you're a farmer, a tractor is really useful, it helps you become more efficient. But when you when you start applying it to certain other fields, like education or mental health, they're more brittle. They're more fragile, those spaces. AI is like this train that's left the station and regulation is way behind the curve of the development of AI, and then there's the whole corporate structures of lobbying governments and constantly pushing back the question of safety and regulation for profit. I would like to see non-profit modes of AI development. I think we'd have different kinds of AI and maybe we'd have - like you're saying, Sebastián - models which might not be as efficient, but they're more accurate, potentially they're safer, and they're more transparent. They're less opaque. Because one of the problems with current connectionist models, particularly in neural nets, we don't really know how they work. The software developers don't know how they work. They tweak the weights of them, the inputs/outputs. So we don't know how what the AI gaze is - what is it seeing itself? We just see the results of that. And for me, I find that quite problematic. I think we need to be able to know what these technologies are and how they work and in fact we don't.

01:33:48 CAROLINA BANDINELLI

Can I ask the question of transparency and accountability to Bahareh?

01:33:50 BAHAREH HERAVI

Yes, what about that specifically?

01:33:51 CAROLINA BANDINELLI

What do you think from your viewpoint?

01:33:53 BAHAREH HERAVI

Transparency and accountability are obviously huge problems when it comes to AI and when it comes to journalists' use of AI. One thing that the journalists often need to do whenever they're reporting on something - there's different guidelines and rules, but for example at the BBC it will be - they need to have two original sources if their own reporter is not on the field. And this story with technology, when journalists use it, is treat any information from the internet as any other source that you would. Same with AI. Treat AI as any other source and don't rely on that. But the problem is, when it comes to accountability and transparency, we don't know how it is generated. There is no transparency there, and then when we don't know how it is generated, what is behind it, what data has been involved in training this model, we can also not

necessarily be accountable, because we don't know what's been happening there. So if somebody needs to be accountable, then using AI could be problematic. Or, if they are using it they need to be very specific and very clear rules and regulations and disclosure agreements, and so on, on how AI has been done, to what extent, in what process, and what is communicated to the user and to the audience.

01:35:17 SEBASTIÁN LEHUEDÉ

I just wanted to point out two things. First, that the main AI company, it's called OpenAI, so that's so interesting, and it's not very open, right? We know how opaque it is, but it's interesting because at the moment - I don't know how it has developed, I read this a few months ago - but there seems to be two paradigms within AI companies. On the one hand, you have OpenAI which doesn't really disclose much information about the training of its models, like the parameters or the criteria, so to speak, that it uses or 'weights'. It's a very technical concept, but let's say criteria that it uses to train its models. And on the other hand, Meta, the parent company of Facebook, they say that they want to go farther into openness and transparency, so they want to provide the weights or the criteria that they're using to train their

models. The problem is that people from OpenAI and other circles, they say that that's actually dangerous, because you can have malevolent or maleficent people using those... It's related to the discussion we had before - we have some actors that might misuse that kind of information. So that's why OpenAI, according to them, are not releasing all the information of their models. Now this has been also - again, I want to point out Karen Hao's book because it's so good -it has been a constant question for OpenAI, and I love- because at the beginning they had this promise, the promise of openness and transparency. But then, over time, they came up with different justifications, excuses or arguments to become more and more closed to fulfil what they consider to be their mission, which is to provide safe AI for humanity.

01:37:04 CAROLINA BANDINELLI

This is what Vincenzo was saying before about safety and censorship, accountability, transparency, regulation, freedom and that incidents can unveil these contradictions.

01:37:22 EMBODIED AUDIENCE

The conversation has been very focused on this infrastructure of what we put inside the AI, and how then the machine learning works, etc, etc, and how safe it is, and then how people make decisions or use AI. But I really

wanted to... Because take out AI from the picture, the way that many companies work - and if you talk with people who work in private businesses, they call consultants like Bain and Company and McKinsey, and for them it's also like [how] someone would use ChatGPT. You know, it's a black box. "Should we do this project? McKinsey says so. McKinsey has all this knowledge and know-how, so yeah, let's do this project." With Covid as well - it doesn't matter your position on lockdown, I think lockdown was the right decision - you can't argue that many of the political actors, their judgement was very [much] based on a scientistic... like on scientism, which really handicapped their political judgement on if lockdown or no-lockdown or what should be our strategy to deal with this pandemic, etc. So, the fact of how we're going to use AI, etc, it is, of course, in the infrastructural models and how we make AI, but it's also in our culture which is very rationalist, which is very scientistic, which has all these biases and if academia can't make its own large models - such as OpenAI or Anthropic or other AI companies can because they have all this access to private capital, etc - we can try to make a difference in the culture and show how... In the summer in Lisbon, I went to an AR workshop and someone showed me this visual representation of a neural network that I believe Google did. It's basically two plagues, and there's all these strings of - I don't know the technical term, because I'm not a computer scientist - but the neural network itself. Even the best computer scientists in the world won't be able to explain to you exactly what happens, right? Because you have all this data, which, the last time that OpenAI published what they use for the data, I think 30% was basically Reddit posts, so that's the level we're dealing with. But they themselves don't know. Humans or students, 18 year olds, use ChatGPT and we judge them. I mean, Warner Brothers currently selects their roster of films with this AI software. So how are we expecting that people are going to be responsible when it's a cultural problem? It's an

ideological problem. The technology will continue, but if we make this effort to try to tell people it's a black box and you're not getting too much value out of it, maybe that can make more of a difference than all this infrastructural talk that we don't have resources for, and people are not even going to hear us because they have all these biases.

01:41:38 BAHAREH HERAVI

I just have a question, who's "we" and who's "the people"?

01:41:42 EMBODIED AUDIENCE

"We" is... So you're trying to do like an AI collective, and we're the responsible ones; people in open AI....

01:41:50 BAHAREH HERAVI

They're not, right?

01:41:51 EMBODIED AUDIENCE

They're really not, no.

01:41:53 BAHAREH HERAVI

We are not either. So there's no "we" to go to tell the people what is right and what is not right, or what is good and what is not good.

01:42:01 EMBODIED AUDIENCE

I mean, I think "we" is the people who... I mean, you're in academia or we have less power over this. And the

technologists do.

01:42:19 BAHAREH HERAVI

So you mean people who are not Big Tech...

01:42:22 EMBODIED AUDIENCE

Yeah.

01:42:23 BAHAREH HERAVI

...or the ones outside of... Yeah, well, the idea is it would be very good if "we" could make an impact. But the problem is, the "we" that you're talking about are also the "we" that's not heard as much. It's the "we" that doesn't have as much power. So then the question becomes, how this "we" can be heard at all when it comes in front of the power, when it comes in front of the Big Tech, when it comes in front of the big capitalism.

01:42:55 EMBODIED AUDIENCE

But, you know, we've been seeing political shifts of people who had no power at all, were at the outskirts of public discourse, and suddenly people who had lots of power and had basically a monopoly of discourse and information, etc, have way less power than in the 1990s, and the other ones that were in the outskirts of society now have a lot of power, and are not wielding it responsibly. I would say that everyone in this room is probably more responsible than many people in OpenAI, etc.

01:43:36 CAROLINA BANDINELLI
So in this room, we haven't lost faith in the revolution?

01:43:41 SEBASTIÁN LEHUEDÉ
No, I do have faith in revolution when it comes to AI.

01:43:45 CAROLINA BANDINELLI
I know that you have.

01:43:47 SEBASTIÁN LEHUEDÉ
Something I wanted to say, related to what you were saying, is how to resist AI. Or how to provide more agency to ordinary people when it comes to these kind of things. I think that there are very promising things going on, and we need to highlight them. There are artists who are mobilising against intellectual property theft. There are - in the case I studied, a case in Santiago in Chile - opposing the construction of a Google data centre that got cancelled in the end after local opposition. There are data workers unionising in Africa and so on. So there is a lot of agency. And I think looking at the infrastructure is super important. It also depends on when you're where you're standing. I went to a conference on AI recently to South Africa, Cape Town, and the

main topic was critical minerals. In the Global North, we don't see critical minerals. We are based here right now, but it's taken for granted, right? It's not considered... I went to a conference once, and I told them, "I look at AI, critical minerals." And they were like, "Oh, no, AI itself." And I was like, what does it mean, minerals don't count as AI? But then you go to the Global South and it's the main topic. You mention AI and critical minerals come to the fore, for example. What matters and what mobilises people changes according to the location, the context and so on. What I think might be the next step - and something we can help as academics and acknowledging the power that we have in this kind of Global North university, for example - is to try to facilitate the generation of solidarity among these groups so we can better see the connections between them, and create maybe a bloc of people affected or mobilising in relation to AI. But I think that's already happening. I think that would be my message. We have to foreground those voices instead of thinking what we can do as researchers. It's already happening, and the question will be how we can help, how we can support those groups.

01:45:56 BAHAREH HERAVI

Just one thing I want to say. I think the reason that I highlighted the "we" is that we should not forget that - you mentioned that most of us, or all of us, in this room are more responsible than other people outside of here - I don't think that is the case. That's a very supremist, elitist view to assume that we sitting here know better than people outside of this world, or know better than people outside of academia or outside of this country, or whatever that we are putting our "we" around. So I think we should not forget that we are not better, or we are not in any way... We cannot be the voices of certain people who are living on the ground and be here as a hero or as somebody who wants to save the world, pretending that we can multiply their words. But we can create venues, perhaps, to facilitate or to make our own small effort to have some people heard without forgetting that there's no such thing as we who's saving everyone.

01:47:14 JO GARDE HANSEN

Yeah, I'd echo that. There's something about... in the AI debate of certain parts of universities becoming splendidly isolated in some way, or somehow safe sanctuary spaces. I don't think that's the direction we'd want to think of ourselves or go in. As I've said in other fora, around this in particular, I think that the public - if it is outside of academia, if we want to say that, and of course we are all members of the public, just so we know - the public probably

does have to understand the role of universities. I think I always feel it has to- right now, anyway, particularly with the attacks on universities, or the criticisms of universities across the world, or particularly in the Global North and the West and America and the UK and Europe, is that the general public does have to fall in love with universities again and understand what the purpose of a university is. And we also have to understand, or at least explore or express our purpose more than we do, and the purpose of research as well, and teaching and educating. We are educating generations of people, and that's one of the most impactful ways... I've probably had more impact by teaching one student in my life than I've had by all my research. And face-to-face interactions with one student have probably had more impact for me than all the research I've ever done and all the books I've ever written, so being an educator in that personalised way is really, really important. That's what I'd say. Some cheers in the audience [laughs].

01:48:59 CAROLINA BANDINELLI

Jo, at the beginning you asked me, "Are we really going on for two hours?" Yes. Yes. We made it. I know that there might be other questions in the room. And then perhaps we can have a chat now, but then you can stand up and drink and eat and do whatever you want. Thank you very much. I hope that this experimental format gave you the opportunity to have an interesting chat. Surely, it was very interesting for me, and I think also for the people that are here. And speaking of patterns and formats, somehow we always end up with this question, "Okay, so what do we do?" And who are we? And

what role do we play? And I think this is the right question to end this kind of conversation. So thank you.

01:50:11 SEBASTIÁN LEHUEDÉ
Thank you. Thank you.

FOOTNOTES

1. Yarden Katz, *Artificial Whiteness: Politics and Ideology in Artificial Intelligence*, New York: Columbia University Press, 2020. ↑

2. Karen Hao, *Empire of AI: Inside the Reckless Race for Total Domination*. London: Allen Lane, 2025. ↑