Anders Sandberg – A talk about morphological freedom

https://www.youtube.com/watch?v=5XKTdx5BWic

“If you don’t want to work on technology that could transform the world, well what… what are you developing?” – Anders Sandberg

Join our discussion with Anders Sandberg, one of the world’s leading scholars in the principles of morphological freedom. Anders, a professor at Oxford’s Future of Humanity Institute, has written dozens of groundbreaking studies that investigate risks, solutions, and ethics about developing technologies, existential threats, people’s rights, artificial intelligence, and of course, morphological freedom.

Here, we get the chance to ask him questions going into specifics of how to protect rights, how to develop technologies safely, and looking for analogies that could be educational on how to approach the impact of morphological freedom on society.

Please forgive a bit of rough editing and technical difficulties. We think the content is truly enlightening and a pleasure to watch and listen to. Hope you enjoy!
If you like, you can follow us on Twitter at https://twitter.com/FreedomOfForm. And, if you wish to support our work, we’d be grateful for your support on our Patreon!
https://www.patreon.com/freedomofform

-Dan Davies and the rest of us at the Freedom of Form Foundation

Transcript contents

Introduction

D: ...transformative purposes,

A: Mm.

D: it's also for getting rid of diseases and, you know,

A: Yeah.

D: the whole gamut of um, transhumanist sort of points of view and, generally health, you know…

A: Mmhm… Ja, that makes sense, it's funny because I was just in a meeting with some biosafety researchers around here, we have one team working on biotechnology and biosafety, and uh, we are mostly worried about the darker applications of these technologies

D: Yeah.

A: so, how one can then go about getting the good ones working while minimizing the bad ones, so we're working on a paper about what you call 'differential technology development',

D: Yeah.

A: basically how do you speed up stuff that, because, somebody's always going to be there going about some bad application, but you might actually be able to prevent that by having protective technologies done earlier.

D: Yeah.

A: So this is a fairly high-level theory paper we're working on. But I was uh, last week in the Bay Area for this biosecurity conference, that was pretty fun.

D: Oh, right, well, last week I was in Boston, and…

A: Ah ja. That other biotech place, yeah… [chuckling]

D: hahaha, where we were doing um, a furry convention. [laughing]

A: Ah, ja. Ja.

D: And so we were presenting the Freedom of Form Foundation to, uh, its home crowd, so to speak…

A: Ah, ja, ja.

D: Um, but there's always somebody new, um, there that hasn't heard of us before and in this case there were quite a few, so it was well worthwhile.

A: Mmhm, cool. Ja. Yeah, it's a very interesting small world, when you were saying Boston I was thinking of George Church, is, is around there and is of course involved in all biotechn-

D: Oh yes, George Church, yes, he's another fellow that we're attempting to get in touch with.

A: Ja. Yeah. He is pretty astonishing, it's also probably harder to grab him than Nick Bolstrom, because he's involved in so much uhh, stuff.

D: Yeah.

A: Among oth, the other weird things I have in my books up there, is a one dollar bill from George Church, uh, it's startup money.

D: Haha…

A: So, w-we were in an elevator at the conference, uh, talking about something-

D: I mean, I've heard of stingy, but this is getting silly! [laughing]

A: So, uh, so basically what was happening was, we were chatting about something, and then he remarked, "Look, we are breaking elevator etiquette here by talking in an elevator, we should have an awkward silence!"

D: Haha!

A: And so I responded to say, "but aren't y- all you Americans just doing startup pitches in your elevator?!"

A: "You convinced me son, here is one dollar!"

D: [laughing]

A: So if I ever were to make a biotech startup, I am gonna include that one dollar.

D: You've gotta get me and George Church together in an elevator now,

A: Ahh yeah, ja.

D: [laughing] just to carry that joke on.

A: Ja. Yes, you might have a good uh, pitch here.

D: That'd be brilliant.

A: Yup.

D: Okay.

A: Hehehe. Yeah, it's a funny situation. So, I met several people over at my conference of course, who were

D: Yeah.

A: somewhere orbiting in his network, which is not very hard because it's so vast

D: Oh, it is.

A: but everybody's involved.

D: Uh, I've seen how many people are involved in the Lifeboat Foundation and SENS,
and um, you know, the Future of Humanity Institute itself of course, and it's all quite a widespread variety of people, but they're always so busy it's difficult to get hold of any particular individual!

A: Yes, yes. Ah, it's super-annoying and the worst thing is that many of us are working very hard on understanding priority setting and working on the most important things, which typically means 'oh, no time for anything else.'

D: Yeah.

A: Uh, which is super-annoying, because I want to have time to do everything.

D: Same here! [laughing]

A: Ja.

D: It really, um, is quite irritating just, just how little time there is in this world for how much we want to do.

A: Mmm, mm. But, on the other hand, the alternative might be worse, imagine a world where we actually didn't want very much so there woul- would be plenty of time.

D: Well, for some people that's how it is.

A: Yeah! For some people. Heheheh!

D: That certainly isn't me, though. Um, so I've finally found my notes that I made earlier for the interview, so…

A: Cool. Ja.

D: Things are getting better already.

Why would nature be considered unwise sometimes?

D: Um, so… I would like to ask you, um… based on what you said on, on another of your websites (I forget exactly which), why would nature be considered unwise sometimes?

A: Well, wisdom is about not just intelligence and problem solving, but uh, knowing what problems you should want to solve. And nature doesn't actually want to solve any particular problem, it just does stuff; evolution just generates uh, the patterns, with no heed for how much pain and suffering may be caused, or whether evolution barrels down a dead end that then leads to a mass extinction. So I would argue that nature is deeply unwise and deeply inhumane.

D: Um, from my perspective, I mean, I've had, actually, probably an annoyingly nice childhood compared to what some people have had to put up with. Um, and I should feel very lucky, really.

D: You know, for my human upbringing that I've had. No real issues there. Um, yes, my parents eventually split up, but it wasn't in any way that I couldn't cope with. And, I feel like, you know, this is going to be viewed by some people, you know, this whole concept of morphological freedom, as a bit of a first world problem.

Um, but when you look at the potential that can be unlocked, um, y'know, by enabling technologies for morphological freedom purposes, they affect the whole world. And, so, you know, it's, it's something that, clearly evolution has got us a certain far- a certain distance, um, already, but it can only operate at a certain speed as well.

A: Hmm. And, uh, also saying that 'oh, that's just a luxury problem' quite often misses that solving that luxury problem actually solves other things. It's a bit like spaceflight. Uh, that's probably one of the least cost-effective ways we can right now reduce global existential risk. However, developing automated mining systems that can work without having human labor, automated manufacturing systems that work under very harsh conditions, uh, automated recycling systems and enclosed biospheres, these things are really useful down on Earth and actually if one can motivate that by having space, bingo. Similarly, morphological freedom, yeah, we just might say, 'yeah, that's only relevant once you have political freedom and uh, enough wealth, etcetera.'

A: But it also allows us actually to cure many of the involuntary morphological variants which tend to cause disease. So, obviously, er, achieving that technology would actually allow us to solve a lot of more down-to-Earth problems. It's not like they're separate, I think quite a lot of people assume that when you develop something for a particular goal, it can only be used for that goal.

D: Yeah.

A: But computers are used to transmit cat photos beside they crack German codes. They are a general purpose technology. And developing general purpose technologies, that is so important, because that gives us a way of transforming the world.

D: Indeed. Sorry, I don't mean to be rude or anything, I just have to set something up on here, I forgot to have a backup recording.

A: And I should be turning off the sound of at least my phone.

D: Yeah. I just need to have a backup recording just in case.

A: Ahahah, yup, that's a great point, ja. Ja.

D: Uhh... where did this sound recording app go... Oh, yes.

A: Again, the phone's a beautiful example, actually, they're not just a beautiful, they're a perfect example.

D: Hmm.

A: So, uhh, when the cellphone was introduced, at least in Sweden the nickname was uhh, 'Yuppinala', literally 'Yuppie's teddy-bear',

D: [laughing]

A: because they were kind of fondling it, and showing off this expensive thing - and 'who needs that when there are good phones connected to the wall?'

D: Hmm.

A: And gradually, they became more and more common, and then we got to smartphones. Which, when you think about it, is using the most hi-tech technology possible, and for many, would appear to be a luxury product. Except that the smartphone also brought banking to a lot of parts of Africa.
By actually integrating cameras, so that, well, suddenly police brutality became a lot, lot harder to do, because somebody might be filming you and putting it up on Youtube. It had a lot of complex unforeseen social consequences, many of which I think are profoundly useful for especially the weakest ones in the world.

D: Yeah.

A: So generally, when people talk about technology, they talk about it in terms of images of what they imagine doing, like uh, scenes in a movie. The problem is of course, when something happens in a movie, there is nothing else about it. In reality, most of our technologies are going on in everyday life, they are intermeshed with everything else in society.
A car is not just a conveyance from point A to point B. To Americans, it's freedom, and a way of expressing oneself, and uh, has a lot of other connotations. The phone is not just talking to people.

D: Yeah.

A: And this goes even more of course for biological technologies. That we actually have painkillers that work today, is amazingly powerful, that allows a lot of other medicine to work. It has, uh, drawbacks when they are misused, but they enable many things. As we are getting better able to control our biome, biology or our morphology, we should expect that to have profound effects on quite a lot of domains. The problem is, of course, most people will think about it like the movie-clip, where somebody changes shape.

D: Yeah.

A: That's it. That's the only thing they do in the plot, and that misses that why they are changing shape, how they are changing shape, what this does in their family life, how this fits into society, what functions this enables,

D: Exactly.

A: That's the actual moral, uh, releva-, morally relevant part.

What if somebody decides to misuse the technology?

D: What if somebody decides to misuse the technology, which is of course a major concern we've got to consider ethically speaking.

A: Ja. Uh, um, because any technology that is capable of transforming the world, is also going to be capable of doing things much worse. The, the thing is, however, if you don't want to work on the technology that could transform the world, w- well, what- what are you developing in that case?

D: Exactly.

A: So, the real question is, how can one control that kind of transformation in a useful way, & this is of course where both ethics and thinking about innovation and technology and control comes in, and even what kinds of control actually are morally allowed.

D: And, when you look at licensing as well of technologies that you have developed, things like open source licensing, I find, tend to help, um, usually, in terms of, um, preventing the, um, sort-of restriction of use of a technology to only negative or only um, a specific company or government or individual's purposes. Instead, it has to be available to everybody, so at least there's a fair crack of the whip for those who want to use it for a positive purpose.

A: Yeah, a lot of it of course depends on the scaling of the positive or negative purposes. So if we think about a classic problem about computer security; uh, if I tell the world about a vulnerability in an operating system, I'm making life much easier for some hackers. However, we can probably expect that plenty of hackers already know about that vulnerability so, it's probably better that a large group of people know, we should be careful about using that operating system, or rather, we should be updating or patching it very very quickly.

D: Yeah.

A: The problem comes of course when you have technologies that are harder to patch.

D: Yeah.

A: It's one thing if I discover a vulnerability in the global ecosystems. Now, who do I tell about that? In fact, it might be rather dangerous to talk about it too much. Yet, there is nowhere I can file a bug report. This is of course where things get really interesting if you could /actually/ do a patch.

D: Yes.

A: If I find that I have a weakness in my immune system, in that case, could I actually add an update? That seems to be very useful.

D: Exactly, yes. To be able to treat the body as software, to be able to file bug reports, patch it and update it. Um, I mean, we can see that already kind of happening with what um, Tesla are doing for cars,

A: Hmm.

D: um, where they are allowing over the air updates, and so when you buy a car and it's actually getting better over time, more capable and autonomous driving.

A: As long as you pay for it.

D: and, obviously, if you've paid for the car in the first place, you're getting your over-the-air updates, if you've paid for autonomous driving, it includes that. But um, I see that as being a very good model of um, you know, the, the future sort of approaches that we need to take to some things in life that haven't yet really been fully realized in terms of their digital potential.

A: Mmhm.

D: And healthcare is certainly one of those things.

A: Hmm. And generally of course, we are groping in that direction. After all, um, vaccines are in a sense, a globalized, collective immune system. It's a technological immune system and a really astonishing achievement.

A: Now, what we really would like is of course, a way of updating that. So with the new coronavirus as well, sooner or later, we are going to know how to make a vaccine.

D: Yes.

A: At that point, it would be very nice if everybody could get it instantly.

D: Yes.

A: We want to speed this up of course, tremendously. The problem is, of course, too rapid update cycles are also risky, as many of us know from our computers, when you send out a bad update, that can actually harm a lot of computers simultaneously. S-

D: So you've got to be able to roll it back.

A: Yep.

D: You've got to be able to, undo the update if it hasn't worked out. Exactly.

A: Mmhm. And this is of course a really interesting problem in biology, because even though we might dream about turning to the software so that I can just go into the right app and kind of update my skin color or fur or tail or something like that, the real problem is of course, real biology doesn't work like that. Cells need to divide, or uh, disappear, it takes time, and they are all enmeshed in this really messy complex system. This is of course a real challenge to overcome in the long run. But first things first, we firstly might just want to change some biochemistry of the, uh, genome, that's hard enough as it already is.

D: And indeed before even that, you know, you might just want to improve on what's available in terms of neuroprosthetics and bioprinting.

A: Mmm. Oh yeah. And, it's an interesting problem, what stops things like that. So for example, the world of regenerative medicine, it kind of burst open in the late '90s, and people started saying 'oh, we are going to culture cells and bioprint,' and then things slowed down, and a lot of that was just regulatory uncertainty. Nobody really knew whether this was a drug or a transplant or something else, and since the rules are different for different things, people didn't quite invest in it,

D: Yeah.

A: so the end result was a lot of people were doing fun bioprinting with no aim of actually making a product. So, that's another annoying thing where even if the science is perfect, we might still find the regulatory situation that slows things down.

D: Hmm, that's... very interesting [chuckling].

A: It's also deeply annoying because, for many of us, it would be like, 'oh, I can totally envision how to make a safe product, oh, how to convince the government agents to make the right decision...' and it seems way harder, and... it's only harder of course if you don't have the skills. There are actually people who specialize in doing that, uh, you just need to start talking to them too. [clearing throat] But even understanding where are the skills is quite often hard; I remember vividly having a meeting with some diplomats about global catastrophic risk. And, we were astonished by what they felt were easy stuff. 'Oh, getting a UN convention, yeah, that takes some elbow grease but it's not terribly hard. Oh, getting the Security Council to do something, that can be the crowning pinnacle of your career.' They were on the other hand realized that 'What?! You can do /that/ with /genes/? What?! Blockchains can do /that/?' and they were totally shocked that even this knowledge existed that they had never heard of, and vice versa.

D: I'm very glad you're mentioning blockchains in the same sentence there as well.

A: [chuckling]

D: Because, um, it's something that's been much-maligned by people who've seen the, um, you know, the misuses of um, like, uh, cryptocurrencies for um, you know, scams and so forth, and of course, they forget that that's something you can already do with an ordinary currency and has been done since...

A: Hmm.

D: time immemorial, really, with ordinary currencies. So it doesn't really make much of a difference that it's happening with cryptocurrencies too, the main thing is, /what else can we use blockchain technologies for/, that maybe they are actually more suited to in some regards. Um, and the- I think there's, there's a lot of potential in terms of securing data, and specifically making sure that when you receive data, it's the same as what you sent.

A: Mmhm...

D: Uh, traceability is actually something that you can inherently include in a blockchain if you do it right.

A: Yeah. [clearing throat] The problem is of course, most people don't know what blockchains are good for,

D: Yeah.

A: so they just say you that uh, they shouldn't have it. But for example in biotechnology, there are a lot of very interesting applications. For example, you might want lab equipment that actually documents what it's being used for,

D: Yes.

A: Which is both maybe useful to prevent some misuses, but even more importantly allows you to detect scientific fraud.

D: Mmhm.

A: if I can just send the right part of the blockchain to the scientific journals, should actually show that I actually did these experiments,

D: Yes.

A: I got the results I claimed to, so it's very hard for me to fake unless everybody's in on it.

D: But um, you can swing this, um, back round in favor of the Freedom of Form Foundation conversation-wise, because you can say, um, that the blockchain technologies don't just have the potential to be used with gene editing and with lab equipment, um, but potentially to help improve the, uh, way that we do our research, and the way that we actually communicate it between people around the world. Um, so, I consider that to be a very important thing that, that should be, um, looked into, especially when it comes to um, morphological freedom, because when you're transforming somebody and they're looking different, you're going to need a verifiable record that that took place, and, you know, especially if they edit their genes as well. Um, how do you prove that that person is the same person, um, in their timeline, as they were before? Y'know, so that they've still got the same accounts, and the same place to live, the same rights, the same, y'know, birth certificate, etcetera, etcetera?

A: Yeah. Uh, and in some sense, we already see this with some of the ethical conversation about face transplants.

D: Hmm.

A: Uh, there was a fair bit of hand-wringing about that, 'oh no, maybe this is going to do something bad to people's identity, etcetera.' And then that fell to the wayside, because when people actually tried it, it turned out that nah, that was not much of a problem,

D: Yeah.

A: and people were not turning into exact copies of the face-donor. But, generally, of course, we have this problem, how do you identify people in society? And uh, right now, it's an ecosystem of ways of identification. Quite often, it's based on passwords, something we know, uh, but uh, people are increasingly trying to use biometrics, which I think is a bit of a mistake because biometrics can be updated. And if somebody steals my fingerprint, I can't get new fingerprint as, uh, uh, in a practical way.

D: Indeed. Um, that might be something this system could solve, although that could be used for both good and bad things.

A: Uh, yeah, so basically when we prove that we are who we are,

D: Hmm.

A: what we typically try to do is to show that yeah, there is this causal path from me from that some point in time where it was definitely me by some definition, whether that was when I was born or whether I joined an organization, to the present one. So you could imagine a world where every- there are surveillance cameras everywhere, so then you just track who's uh, that somebody has been moving and that is the same person.

D: Yeah.

A: That's not practical yet, and maybe we don't even want to have that work, so in that case you want other tools for doing identification. But they of course require that now you're tied into a system, and this system might get various powers over you because now it's maintaining your identity. That's also slightly worrisome. So many people would say, 'yeah, I want m- to have that power inside myself, I should be able to prove to the world who I am, without being dependent on the world.' In practice, this is not how things actually work, because much of our identities and especially legal identities, are socially constructed.

D: Yeah.

A: But it's a start.

D: That's um, a very interesting way to put it as well. Haha, um, yeah.

A: And of course in a world where we totally can change up all of this, other things are probably going to be important to signify who we are.

D: Hmm.

A: You're going to want to have certain things that people at a glance, can tell who you are, even though your current uh, shape is very different. That might be things like personal signatures or accessory or just broadcasting the right uh, encrypted identity.

The right to, and origins of the term, morphological freedom

D: And so this kind of, um, drips into people's rights, doesn't it, um, as individuals, um, and how do we have a right of morphological freedom?

A: Ja. And so, I've wrote my essay on morphological freedom, way back, I think I did it in '99. And back then, I think I was very much into this traditional rights framework. Well, it's totally obvious, I have a right to my life, I have a right to my freedom, from that follows that I must have a right to my body, and then I should have a right to update my body.

D: Yes.

A: Relatively straightforward, except then I started studying philosophy and realized that rights are not that easy at all. Indeed, as uh, Bentham famously said, they are "nonsense on stilts." Um, now, the problem is many of the ideas about rights hinge on quite dodgy philosophical foundations,

D: Hmm.

A: in practice rights work really well.

D: Yeah.

A: We declare something a right to society, that means it has to be respected, we set up laws and institutions that protect it, and as long as we more or less agree that this makes sense, this works quite fairly well. And I do think there are pretty profound moral reasons why we should have moral- uh, morphological freedom. It's just that it's going to take a fair bit to get everybody in on that and show that this follows a lot from the other stuff that we're allowing people doing or even obliging people to do in society.

D: Yeah. Um, so, when you um, wrote your book about morphological freedom, um, you say that was back in 1999, was that?

A: Yeah.

D: Um, I think you may have been one of the first to actually develop that as, as a term,

A: Hmm...

D: um, and make it a thing, so to speak.

A: Yeah, Max More beat me to the actual use of the term, and in roughly the same sense, of course, uh, some years earlier, but,

D: Right.

A: Yeah, I was among the first to really start thinking, 'do they actually have some proper formal rights here?' and, 'how does that work in the bioethical debate?' And, it's quite interesting to see what has developed from that paper and the critiques I got and, I've been kind of delighted by some of the disability activists and their critique,

D: Yeah.

A: um, because they think I'm spending too much time on an ableist perspectives, uh, and they're essentially arguing that I'm just arguing- my argument there is negative, nobody should be allowed to stop me from uh, modifying my body. But I'm not arguing that it's a positive right, that I can demand that the healthcare system should do it. And many of my critics from the disability world, are basically arguing that no, it's a positive right, we need to be allowed by this society to do it, and that society should even give us some help. Now this is going way further out than I would, I am taking a fairly liberal libertarian perspective, because I wouldn't dare to actually presume that taxpayers should actually pay

D: Yeah.

A: for my transformation, which I think is an interesting uh, thing...

D: So, you were essentially taking middle ground that you didn't even know was middle ground

A: Yes.

D: [laughing], and you found these 'extreme' people, who uh, wanted to take it further.

Morphological freedom and disability rights

A: Yes. And of course, they don't think they are taking it further, they argue that, well, being disabled in some sense, or differently able, means society needs to accept my different morphology. And I think we have pretty strong arguments for that. But that also means that websites need to be able to function for blind people.

D: Yeah.

A: That means that we need to have wheelchair ramps and elevators in public buildings. This is actually, uh, demanding a fair bit of effort. It's not an enormous one, and as we get wealthier and more technologically advanced, it gets easier to accommodate, but it still means that others actually have an obligation to help people maintain their morphology.

D: And it does mean, it does then bring about a, a very interesting question because, um, if you have come into this life as, you know, obviously, you're in a human body, but say, you feel like you never belonged in that human body and you should have been, for example, a dragon. And you should have these HUUUUUGE wings, and they've gotta be big enough to fly, so they've got to have a massive span. And you've gotta have a long tail as well, and you've got to have horns, and

A: Hmm.

D: not a lot of clothes fit you, not a lot of furniture fits you, and... how much accommodation is supposed to be provided to you by what has hitherto been human society,

A: Ja.

D: when you are going around in a body like that, um, which, in some ways, is providing you with more abilities than you would have had before anyway, because you would be able to fly (uh, maybe for a couple of minutes at most, but still...), and um, you know, it's a question of like, does society have to view this as something that you didn't have a choice in, or something that you did?

A: Mm. And I think the most obvious response is that society will at first of course, definitely say, 'oh, it's only your choice and if nothing fits for you, well, too bad for you.' Um, but this is a little bit like saying, 'I don't want to have a phone'. So, once upon a time, well, everybody would just respond saying 'yeah, that's... normal.' Then, 'okay, that's slightly awkward.' And these days, not having a phone, is actually going to hinder you rather strongly. And indeed, many societal services are kind of dependent on you being possible to access.

D: Hmm.

A: Uh, today it's increasingly hard to get by without actually having a smartphone.

D: Yeah.

A: Um, and-

D: I know there are some people who can and do get away with it, but they find themselves missing out on certain things, as we saw.

A: Hmm. Yeah, and, this missing out, there's always a cost of not fitting into the mainstream.

D: Yeah.

A: And we can say 'yeah, we want to lower some of these costs, but maybe not others.' For example, we have decided that uh, being able to read and write and participate in society is so important that people- that we have mandatory school. So this is a form of enhancement,

D: Yeah.

A: which takes time and resources and money, but it's regarded as so good and so important that we, everybody has to do it. Now, the problem here is when you do something new; at what point does it

become allowed, at what point does it become the favored or obligatory? And I think that depends partially on of course, how much you can easily accommodate things. So there are some people who regard themselves as sensitive to electromagnetic fields. Exactly what's going on there is tricky to tell, but

D: Yeah.

A: many of them want to stay away from electrical and electronical appliances (which is getting rather tough).

D: Yeah.

A: Up until a point, you can oblige them, but beyond that point, yeah, maybe they have to go move out into the forest.

D: Yeah.

A: And, uh, of course, they might then say, 'yeah well, maybe you can just turn off the Wi-Fi and phones in the city.' and mainstream society will respond to say, 'no thanks.' They're a little bit too important for how things work, and you're a little bit too few, and even if we feel really bad about you, you actually can't uh, function.'

D: Hmm.

A: 'Maybe we should be investing more in finding a cure for you, unless you think that your identity's tied to not being, uh, well with electrical fields.' So, the dragon example is interesting, because it seems like, so long as the dragon is a very rare occurrence, in that case it's fairly clear that the onus on society is not to provide for them,

D: Yeah.

A: but there are many other changes that you could do that require much less accommodation. The most important ones of course, are the current struggles about, well, what about gender identity and gender expression. Uh, there is enough interesting struggles over that, even though it's mostly about, how do you maintain identities in an open, liberal society? And, the basic thing is, it's getting complex, and that the cognitive load is tricky. Sometimes, you can solve that in a neat way, like wearing a badge telling what your preferred pronouns is.

D: Yeah.

A: A few years back, that would have been totally crazy. Today, hmmm, starting to get rather uh, reasonable,

D: Yeah.

A: uh, because it saves one problem. Um, you can imagine having apps on the phone uh, discreetly informing whether hugging is okay or not. In this case

D: Mmhm.

A: you have a technological solution. Some people don't like it, because they don't even think this should be done technologically, you should /just know/. And they of course, get very upset when they encounter people who are very hard to know, because they actually got preferences that are very different.

D: Yeah. And sometimes, it's just not easy to remember all of these little details about everybody.

A: Yeah, uh, so I have a bit of face blindness.

I don't know whether it's enough to be clinically diagnosed, but I have a very hard time remembering faces and people etcetera. So, getting pronouns right or genders right, I find tr- troublesome, and maybe I should just, uh, get a diagnosis, so that when somebody gets cross, I can wave my diagnosis at them.

D: Yeah.

A: That's another way we're solving many things in society, we can medicalize them and uh, if I have a diagnosis for something, I get a license in some sense.

A: This is r- rather bizzare.

D: Limited, but yes.

A: It's a limited l- license, but it's also a weird effect. A friend of mine, he got a diagnosis of a bit of mild Asperger's, OCD and... uhh, something else. ADHD.

D: Hmm.

A: And his response was, 'Yeah! I get enhancer drugs!' Uhh, and, my colleague here said, "Yeah, he's got a diagnosis by being an academic anyway..." To some extent of course, by having that diagnosis, now he could get access to certain drugs, and uh, he also found that he could elbow a form of accommodation with his personality q- uh, traits, which he now could ascribe to being on the autism spectrum and having OCD, rather than just being a bossy person who likes to micromanage other people.

D: Yeah.

A: Now, the tricky thing here for morphological freedom is of course, how do you balance these things? And unfortunately, I think the answer, in the end term, seems to be solving kind of cultural politics, which is a horrible mess.

I wish there was some sort of beautiful philosophical argument which would allow me to draw a bright line uh, with laser light, and say this is allowed, and this is not allowed. But that is very rarely what actual ethics, gives you.

The collaborative imperative for the FFF when confronting a risk-averse establishment

D: Hmm. Fair enough. So, when you um, wrote your book, um, you obviously had the idea of morphological freedom in mind from, um, what was it, Max? Um...

A: Max More?

D: Max More, yeah. So, he'd already sort of brought it up once before. Um, Now, that was 1999, wasn't it, when he wrote that, and then, what have we got, about... n- about... 9 years' gap, um, before I even found out that the furry fandom existed,

A: Mmhm.

D: let alone started to look into the ways in which we could potentially make it happen in real life.

A: Mmm.

D: Um, and so, I was getting started in 2008, um, I wanted to write my book, so I started on that, but... um, it was only really from writing my book that I realized, 'hang on a minute', these McGuffins, these little crazy ideas of technology that I'm writing about just to make it seem plausible in the book, they /actually make sense/ and I could actually do something with that; which is kind of what then spurred me on to um, you know, get together a bit of a community and to try to uh, develop these things in real life and to uh, start my company and go to university so I knew what the heck I was talking about when it came to biochemistry...! Um, heh, it's useful! And you know, it's, it took a while to really also get my head around business

A: Hmmm

D: and how that functions, and I'm still learning a few things about that too. I think we're, you know, if we're not learning, then what are we doing? but um, it still strikes me as interesting that uh, over all that time, um, I knew there had been an article about morphological freedom on Wikipedia for a little while.

Um, it had been written by a friend um, he goes by the name /Chrome Dragon/ online, or /Chrontius/. And uh, he said that he basically had written it up from a few scraps that he found here and there. I don't recall it actually cited you, so I might have to mention you to him.

A: Mmhm.

D: Um... so, that's gonna be interesting but um, obviously, eventually we... um, coalesced our sort-of various groups of people online into these um, different organizations that are now working together, including the Freedom of Form Foundation. Um, and it'd be great if we can continue to kind of snowball our um, groups and um, knowledge together to um, greater effect. So, I noticed when I was looking at your um, sort of, biography and your descriptions of what you do on a couple of different websites, that you're part of the ERC Unpredict Project.

Um, that one sort of stuck out to me as interesting, can you tell me, how do you think that might interface, um, with the progress of organizations like the FFF?

A: So, that is a fascinating thing. So, the ERC, the European Research Council,

D: Mmhm.

A: they have given us the research money for this very worthy project about thinking beyond the precautionary principle.

D: Yeah.

A: Basically, when you have a new emerging technology, how do you judge whether that's a good or bad thing, how do you actually think about its risks? And, the standard precautionary approach would be, 'well, you should not assume it's safe, you should actually do a lot of research before starting out, to prove that it's safe or figure it out.' the problem is of course, first of all, that research might be hard to very hard to do without living the technology. Second, the research itself might actually cause the dangers, and the many technologies, you might say, 'okay, I need to do something to safeguard against this bad use', but that, safeguard, might in itself have problems, so at this point, the precautionary principle starts biting its own tail.

For example, climate change is bad, so maybe we should do geoengineering to stop it. But geoengineering could also be bad, so... do we now need to do some kind of anti-geoengineering research, before starting doing geoengineering research, in order to deal with the climate?

D: Yeah. And we finish up going down a bit of a rabbit hole there, don't we, of... theoretical research that never becomes actual.

A: Exactly. And this gets even more problematic when dealing with enormous risks. So for example, a few years back, people were worried that the Large Hadron Collider might generate small black holes, strangelets or vacuum decay that would destroy the world. This seems to be a pretty good reason not to turn on an accelerator. Uh, the physicists were really annoyed by this, and eventually wrote a bunch of papers that are good enough to show that the risk is small enough that we

D: Hmm.

A: probably shouldn't lose sleep over it. We can turn on the accelerator. Of course, that was kind-of written after the accelerator was actually up and running to a large degree; we are not great at this kind of risk assessment, the problem is, just because outcomes can be super-bad, that doesn't mean that we are normally totally precluded from doing anything.

D: Yeah.

A: Maybe that would be the safest possible world, but it would also be a world where nothing new can ever be done.

D: It would be a very boring world.

A: It would be a very boring world, but the real problem is of course that, even something as trivial as dropping a pen, might hypothetically cause the world to get destroyed. It's just that I don't have any information about it.

D: Yeah.

A: So maybe I'm precluded from dropping the pen, but then again, holding the pen might also do it. So at this point, I might make a happy argument that well, we have this symmetry, we have eq-equal amount of information leading in both directions, so it shouldn't uh, ma-matter.

D: Yeah.

A: But what do we do more practically? because, in practice, when we are thinking about biotechnology, we might want to actually develop it and minimize the bad sides. And uh, what about the social effects? Europe loves talking about the social impact of technology. But some impacts would be really extremely upsetting for some people while really approved by other people. Should they say, 'the precautionary principle should tell us to no- never do anything that would deeply upset religious people.' That seems to be a very odd conclusion...

D: Indeed.

A: That we should maybe be polite and try to understand people's point of view, yes, but saying that you're not allowed to do stuff that could upset certain religious groups, is wrong. The problem is, however, these religious groups actually think that you're deal- might be dealing with something that's profoundly mega-important. Many critics of transhumanism have argued that this might threaten human dignity, the most important principle ever. Although, how it threatens the dignity requires cons- quite contested discussions. And again, you can make a transhumanist case that 'no, no, no, we are getting human dignity /by/ modifying ourselves.

A: So, what the Unpredict project is about, is trying to find tricks and heuristics to reason about these uncertain things.

D: To reason with the unreasonable?

A: Yes. So, you can't reason with the unreasonable people, but the unreasonable and unknown things out in the world, we can actually make a stab of trying to understand them.

D: Hmm.

A: Sometimes you can get indirect evidence, in the case of the particle accelerator, the existence of the Moon actually tells us that something has been hit with this high energy particle billions of years, without turning into a black hole. That gives us some useful information. One needs to update or patch this argument a bit, but it actually gives us a pretty good reason to think that accelerators are no problem. We can also do historical things like, 'well, when people predict that certain technology will wreck society, how often are they right?' And the general answer seems to be, there is always somebody claiming that it will wreck society and sometimes they are right, but most of these predictions are not linked to that. That's mostly just the normal moral outrage.

D: Yeah, and a lot of these people who um, peddle moral outrage, and who uh, like to make these doomsday predictions, have um, recently in- in, y'know, social terms, gained quite a hefty mouthpiece called the internet.

A: Oh yeah. But so does the critics.

D: Mmhm.

A: So, the natural dynamics of the internet has basically been, sub-groups of individuals that never knew the others existed, could suddenly talk to each other. D: Yeah.

A: This is how the transhumanist movement in the early '90s kind of got its act together, 'ooh, there are other people that want to update themselves!"

D: Mmhm.

A: This is of course, how the furry subculture and practically every subculture on the Earth right now, are getting together.

D: Yes. Including us. [chuckling]

A: The nex- Yeah. The next step is of course, now they are all encountering each other and that causes a lot of interesting friction.

D: Hmm.

A: The best thing that could ever have happened to transhumanists was actually Francis Fukiyama's book, Our Posthuman Future, where he's really, really trying to show that Transhumanism is the worst thing ever, we really need to not go in that direction. Now, when the bioethicists aligned with President George W. Bush and his conservative bioethicists um, are saying something, immediately the liberal establishment in bioethics, put themselves on the other side, just as... a reflex. And Fukiyama's arguments were not great. But, it was an interesting debate that also forced transhumanists to actually

D: Hmm.

A: step up their game. You couldn't just say stuff on their mailing list, you had to actually write a paper to show why this argument doesn't work, when you're claiming that something is against dignity. Well, what are the weak points, then? So the real problem now is of course now, we have a lot of these discourses, but also groups that are not even used to each other's kind of discourse.

D: Yeah.

A: The big problem is that the traditional academic debate as done by traditional academics, it might not be the most important debate happening.

D: Indeed. And, um, we've found in various groups online, you know, you will get people who sometimes, come in, trying to start a debate just for debate's sake, because they're bored.

A: Oh yeah.

D: And... you know, sometimes, they will waste our time and just make a point of it, just for the sake of, you know, keeping themselves entertained. And yet, we will also get some people who are genuinely interested in finding out why we're doing things the way they are.

A: Hmm.

D: And we try to give them, you know, a fair amount of time and effort to uh, understand things. So-

A: Yeah, and often, the fascinating thing is

D: Yeah.

A: w- I'm sorry?

D: Sorry, I was just going to say so, when- when we're using our online chat rooms and systems, um, you know, it gives a very instant kind of discourse, um, because you don't have to kind of pre-arrange a debate time per se, usually, um, as long as everybody involved has a moment to spare. So, the modern internet is, is giving people a lot of that kind of instant gratification, even with debate. Um, I think that's quite fascinating to see, because that's when it's very fresh in people's minds.

A: Yeah. The, the, the downside is, of course, uh, ongoing debate is usually not the most clever thing. Many of the best debates ever have been in a series of letters or papers, where people take some time formulating something,

D: Yes.

A: the other one responds... that gets too deep into the questions. But that's not going to sway public opinion. Public opinion gets swayed in the long run by good ideas, but also powerful images. And they take time to diffuse. And some groups are very good at generating these ones, and uh, kind of win the War of the Memes.

D: Yeah.

A: But the, the real challenge is, of course, figuring out better ways of actually helping good ideas to get ahead. In the natural sciences it's easy, in some sense, because you can do experiments. If it doesn't work in the lab, or if the observations f- don't fit, that theory is going to be in trouble. In some parts of uh, uh, academia, reality is kind of irrelevant

D: Yeah.

A: and it might be more a matter of being more compelling or more loud than others. And of course, in actual political debates, many other considerations also come in. This is of course, deeply annoying if you're an academic, and think everything should be truth-tracking, but in many ways, that's not how things work out in the world, at least short-term.

Moderation of social anxieties and caution

D: Hmm, yes, I- I can totally agree there. I'm very much of the, uh, y'know, /the truth is what matters/, approach, but also of the approach that, if you've got a good idea that you want to see happen, you shouldn't be dissuaded from attempting it, by the inconvenience of the current um, reality that we're in, if you believe there is a pathway forward to change it.

A: Yeah. Uh, and this is of course also a culturally relevant thing. Europe uh, might be a bit bad at this, visions are something that people get rather worried about actually,

D: Mhmm.

A: because they upset things, they're making things different. America, not all parts of America like visions, but I was recently in the Bay Area, and there of course, their stock and trade is basically, coming out with new crazy visions of how the world could be. Everybody laughs at them, and then they, they start wicked startups and they sell uh, this product to everybody. And then consequences ensue.

D: Yeah.

A: And this happens generation after generation. They have gone through at least four generations, Silicon Valley and the Bay Area, changing the world in various ways. And, I think the world is much better for it. A world that is trying to run on past wishes is actually going to be in deep trouble, because reality out there is constantly changing anyway, and we certainly don't know all of the adaptations we might want to have to it.

D: Yes, an-and that's why we all need to evolve.

A: Yeah.

D: Um, because, if we look at it in the longer term, the world isn't going to sit around and stay the same for us forever.

A: Uh, although there are some people who would probably like to try to come up with a way of achieving that - which probably is going to force a lot of drastic change anyway.

D: Yeah.

A: I think the nice thing about biotechnology, and other forms of technology, is that it has to be truth-tracking enough to actually work. Uh, many other claimed technologies or possibilities can be always vaporware. You can always make up a story about why this is great. But, when you actually are getting into the actual technology, it's testable, it can be improved. And you get feedback from reality which is much more powerful than any amount of political ideology or urgent humanitarian need can give you.

D: Yeah.

Technologies that are relevant to freedom of form

D: Absolutely, so... Um, if you look at um, the current sort of industry progress, um, what are your thoughts on um, things like neuroprosthetics and bioprinting?

A: So, right now, neuroprosthetics are awesome, and still not ready for primetime. Because the basic problem is, that you need to actually do brain surgery, and put electrodes into the brain. This is... kind of a bad thing. It's not something you would want to do unless you really had to. And they're getting better. We're certainly having

D: Yeah.

A: nicer nanosurfaces that don't react so badly to the organism.

A: We're getting electrode arrays with more and more electrodes,

D: mmhm.

A: the recent work by Neuralink of making a sewing machine D: Yes.

A: to put in a lot of electrodes is, really promising. Because it also shows that, you can win by doing engineering on other things than just the electrodes themselves. And there are wonderful things like optogenetics, where you transfect uh, neurons to become light-sensitive and can send signals to them using laser light. For example, Professor Gero Miesenböck here in Oxford, is using this to control fruit fly reproduction.
Basically, he can play the different neural circuits in the mating behavior, by sending laser light at these flies. Which sounds like something out of a bad pulp science-fiction story.

D: (laughing quietly)

A: But, that is a really powerful technique, uh,

D: Hmm.

A: Except of course, that we've got big opaque skulls so that we have some problems getting signals in.

D: Yeah.

A: And there are other interesting ideas such as like using neural dust, little ultrasound transceivers, that you put in the cortex. All of these things are really promising, and it's going to take us years to actually get them to be super-useful. Because the problem is, even if we got it to work perfectly in the lab today, getting it out to-to the market, getting it even safety tested and understood enough, is going to take at least a generation, which is deeply annoying.

D: Yeah.

A: Similarly-

D: So, in some way, I suppose, the regulatory system may have to change too.

A: Yeah. But even if the regulators always said yes to everything, I don't think that means that now, you could immediately get a neural lace implant and in a useful way. Because you need the procedures, you need to figure out how do you do rehabilitation, how do you train people this? This takes time, that's the deeply annoying part about many of these new technologies, because...

D: Yeah.

A: Ideally, we want plug and play. We want to be able to get the technology out of the box and immediately get its full potential. This is why many people are so disappointed in the smart drugs. They're not at all like the movie Limitless, where you become super-genius straight away. No, they help certain cognitive, uh, tricks. They help you do certain things, but you need to understand though, what tasks you're supposed to be working on. This is deeply annoying. Ideally, of course, we want something that works as well as fiction.

D: Yeah.

A: But that is of course, why it's actually fiction. Because reality is full of these tedious things that 'Oh, I need to learn how to use this tool, uh, if I get the brain-computer interface, I'm probably going to have to spend years training myself to be really good at it.' Which might very well be worth it, might very well be what uh, we should be learning at school, but it's still years of annoyance. Um, similarly, if we look at bioprinting, I think that we're getting th-there. We're getting th- rather advanced tissues that we can d-do, but again, you need surgery to put it into the body.

And surgery is messy and bloody and ideally,

D: Yes.

A: we want to get around it. In the best of all possible worlds, we just waved our hands and some nanomachines or stem cells swam into the right place, and started dividing, and did, uh, did what they should. We're rather far away from that. But, we're not that far away as we could have been.

D: Hmmm...

A: We're actually starting to know t- how many cells respond. And then people are working on really clever uses of let's say, of immunotherapies against cancer, where you can actually make use of the body's own resources to destroy cancer cells.

D: Yes.

A transhumanist's midlife crisis

A: So, I think we're getting there, it's just that it's going way more slowly than I would, uh, would uh, wish. And as a transhumanist who's started to get grey hair, I think it's a race against time here but um...

D: Here's your midlife crisis. [laughing]

A: Yeah, well, the-

D: Yeah.

A: The transhumanist midlife crisis is, 'oh no, uh, the-the- it's not happening fast enough, I'm so happy I've signed up for cryonics too!' [laughing]

D: [laughing] Fair play to you. Um, I mean, there's those who would um, even try to upload their minds, as well. I've um, been in touch with um, Martine Rothblatt and um, Lori Rhodes, over at uh, the Terasem Movement Foundation, uh, I don't know if you've heard of them before? Yeah?

A: Oh yeah, oh yeah. They're fun.

D: Indeed. And um, of course, Martine Rothblatt also runs um, United Therapeutics, and its um, little spin-off, which is Lung Bioengineering.

Um, and, what they're doing with bioprinting there is quite fascinating, they're um, literally printing scaffolds for lungs, and perfusing them with lung stem cells, um, to create new lungs from a patient's own cells. Uh...

A: Yeah.

D: Yeah.

Differential technology development

A: Yeah. So, I think we'e- we're getting there, it's just taking a lot of time and work. But that is still progress. The, the real tricky thing is, of course, what ordering of technologies do they want to have? So, that is one of the interesting

D: Mm.

A: outcomes of the UnPredict Project, what we call the principle of differential technology development.

D: Mm.

A: So, there are many technologies that you can foresee that have very bad effects. But there are other technologies that if they existed, would ameliorate these bad effects.

A: So, one simple example would be quantum computers versus current encryption.

D: Yeah.

A: That's not very good for our privacy. If we can get quantum-safe encryption, though... and done that- do that before the quantum computer, everything is more or less fine, except that maybe the future generations would be able to read our past encrypted stuff. Um, but that might be less of a problem.

D: Might- Might you need a quantum computer to develop good quantum encryption, though?

A: No. Um, so you can do quantum encryption where you use the quantum properties themselves to make really, uh, tricky messages that cannot be decrypted even with a quantum computer,

D: Okay.

A: if you do it right - and this is not easy. But you can also do classical encryption, something called lattice-based encryption, that seems to be resistant to quantum computers. Now it might turn out that further in the future that some genius comes up with a way of cracking that, and in that case we are in trouble. But the basic thing here is, we can foresee that someone is going to make a quantum computer because they are darned useful for a lot of things.

D: Yeah.

A: And we want another technology to happen before that, because then we're fairly okay. Similarly, we might imagine that somebody's going to make systems to do geoengineering and reduce solar input, which might be maybe a necessary response to climate change, but that's also potentially dangerous, and might be really bad if you happen to live in a Northerly- Northern country. Technologies are undoing that kind of thing. We probably want that earlier. [Clearing throat]. If you develop a drug, you'd better have an antidote to that drug. So, this goes also for biotechnology. So, one scary story about biotechnology would be, 'oh, if everybody has a gene lab , then of course the crazy people, the school shooters are going to have gene labs and they're going to try to make pandemics.' So, you want technologies that prevent that to happen for.

And part of that might be global disease surveillance. Part of that might be labs that actually don't lend themselves to easily making nasty and dangerous stuff. There might be totally different biocontrol mechanisms that you can use to make sure that you don't get outbreaks. Suddenly a lot of those other biotechnologies become much more benign. But you need to push for these safe-making technologies. This of course, also includes, when you do your biotechnology, you might want to make that one reversible, you might be- want to be responsible to it well documented, it can be fine-tuned so you don't get immune responses etcetera. There is a lot of things one can do as a designer to make things safer.

D: Hmm. So, you, um... with, with the um... what's it called... the UnPredict project, you've um, essentially got um... uhh, a- a way of trying to decide what order technologies should be developed in as well as um, obviously bringing those forth to the various people in charge and saying, "Here's what we, um, think actually matters, and here's what you need to, um, consider in order to not be scared to death about some of these other technologies. Um, which you know, I- I think that's brilliant, kudos to you, you know? [laughing]. Um, So, I mean it's fascinating to me because I'm coming at this from the

perspective where I'm very much personally a strong advocate of the gene editing approaches, as you know with my company that I've started, so, um, to me, um, you know, I'm talking to my colleagues in the Freedom of Form Foundation, and they say to me, 'you know, you shouldn't rule out bioprinting, you shouldn't rule out neuroprosthetics or surgery, um, or a few other possibilities, because you don't know what's gonna happen in the future. And, you don't know if um, gene editing is really going to be developed that fast, compared to, what to them, with current technologies and capabilities, appears to be an easier path, um, of, you know, doing um... neuroprosthetics or bioprinting, etcetera. Now, I- I respect their points of view. And I'll continue to advocate them, even. Because I want to see the Freedom of Form Foundation succeed in its mission just as much as I want to see my own company succeed in its, they are, y'know, parallel to one another. And... I mean... If... if UnPredict was to, um, sort of look at both groupings there of um, y'know, ideologies and perspectives, um, would it say, 'oh, but it's obvious that', say, y'know, 'gene editing has to be done first', or that 'neuroprosthetics has to be done first', and therefore, should we be changing our focus and not funding the other one as much?

A: Well, uh, the idea of differential technology development is usually within one technology. Uh, so, if it as that maybe gene editing was super-dangerous while neural implants worked, then we might say, 'let's go for neural implants, focus more on them, because maybe they are available without the need to do the gene editing. But that's not the case in this case. Uh, it rather... the risk profiles are totally, utterly different. They are about different things, even.

D: Yeah.

A: So, you can't really compare them, so... it's not so much that it's a focus between these ones, rather for Freedom of Form, uh... just go for all the options and try to see which ones become available, including the surprises. Um, I remember vividly being on a panel at the, the futurist debate, where people were asked "So, over the next ten years, what would you expect to be surprising in your own field if it happened?" And, we're all going down the panel and rattling off the different things, and then there was the biotechnologist that said, "I can't actually give any answer, because I wouldn't be surprised by anything, because so much shockingly surprising stuff is happening at such a high rate, that we actually honestly um... can't make the good prediction.

D: Yeah.

A: There is some ultimate limit to what normal biology can do. But they are pretty wide. But for practical purposes, oh dear. This means we are going to get grandiose surprises.

D: Hmm...

A: This was both a good new - a piece of good news and bad news. Because, if you want to try to be uh, cautious about these technologies. It's very hard to do it in a useful way, we can figure out that pathogens are bad, that uh... smart toxins are bad, that people probably shouldn't be allowed to bro- brew up, too-dangerous toxins in their home. Yeah, we should probably come up with a way of handling that. But...

D: Yeah. Perhaps, the best answer then to those who are concerned about the potential um... issues that each of these technologies might present, might be each other. Because, if we have a problem with gene editing um, a solution to it might be to use, y'know, neuroprosthetics or bioprinting for certain

aspects. And, you know, if there's a problem with neuroprosthetics or bioprinting, maybe we can use gene editing to smooth out how that is incorporated into the body.

A: Totally. And, I think this shows an important problem. People tend to get totally hung up about the means to achieve an end, so we might be talking about uh, bioprinting versus uh, gene editing. And certainly, a company might invest in one technology or the other. But remember, I would say as someone in the uh... Humanities faculties, 'so what does it do?' So, if uh... both tools turn you into fox, um the relevant question is do you get turned into a fox.

That's an interesting case. I don't care if this was done by technology A or B, for most cases.

D: Hmm.

A: This goes back to the debate about human enhancement. A lot of people have a lot of opinions about people taking smart drugs as pills.

D: Yeah.

A: But, they're way more willing to take a herbal supplement that supposedly does exactly the same thing.

The moral arguments

A: So um, uh... People's perspective is quite often linked to tools. But that's not where the moral thing happens. The real question is things like, is this risky... is this reversible? Can we find out useful and morally relevant information without doing something too bad? There are some issues for example in drug development where we don't allow the drug testing on children...

D: Yeah.

A: But then the drug gets used on children in the healthcare setting and, this point we might have a moral problem uh, because

D: Yeah.

A: Now we're in a sense giving exactly the, the drug testing... except that uh... yeah. Now the problem is of course, having a proper ethical discussion is tricky. You can do the high-brow version of it in the ethics journals and the seminars around Oxford, uh, which can- typically gets done into nice little details, but that's a personal, substantive ethical discussion

D: Hmm.

A: you actually want in society. Ethical discussion that led uh, to the changes for example in female suffrage, was happening all over the place in the tabloids, in very rakish meetings, in the form of nicely worded letters to the Times, uh, and in all sorts of various graffitis scrawled on the walls.

D: Yeah, up and down all the various different levels of society, exactly.

A: Yeah. Uh, but one can still try to do better ethics. There are some arguments are better than others. And I think for example, when it comes to morphological freedom, it's in line with much else what we do in society. We want people to be able to shape their own lives. We actually think that forcing people to conform to others' ideas about the body, is a really serious matter. Uh, this is of course why trans rights right now are, uh, um, a point a contest, um, because it is causing friction.

D: Yes.

A: But, we have kind of accepted that people are free to actually select how they dress. This was not always true through history. Sartorial laws have been around for a long time, and the idea that you could get thrown into prison um, because you didn't wear a hat in the appropriate setting. That disappeared only about a century ago or so. So, the thing is, we have developed this liberal view that, yeah, people are allowed to do things that basically change their life. Of course, if they look too crazy or stupid, we're going to hold it against them. But, the idea is very clear that, yeah, in a pluralistic society, we must accept different appearances.

D: Mmhm.

A: Then, we might want to figure out how much affordance we can give to them, if they are impractical if someone turns into a dragon, or they require complicated social norms, like uh, labels about pronouns. The reason that annoys people a lot, is that normally we're supposed to know these seamlessly.

D: Yeah.

A: Many people put great stock in their social ability, and they get very annoyed when you make something that's normally fr- free-floating, formalized. And indeed, formalizing some aspects of our society might not be the path we would want to go. Although, that might uh, turn out to be necessary in some situations, like the personal identity; maybe in the end we need to define personal identifying encryption keys on an implanted microchip, rather than appearance or name or anything like that.

D: Indeed.

A: But getting back to that final thing, I think... uh, emphasizing morphological freedom I think is going to be essential because, otherwise we're going to have a very hard problem dealing with people from other culture who might modified themselves. We are going to have a hard time dealing that we will know more about ourselves. Because we're already showing way more diversity in our morphologies, even internally, than most people are aware.

D: Absolutely, it um, amuses me when archaeologists say that they've discovered a new subspecies of human from way back. And, you look at these skeletons that they've uncovered and you sort of compare

them to the morphological differences we already have amongst humans now and you think, you know, is this really a different species?

A: Yeah. Uh, and you could make the argument, maybe it would have been good for us to grow up together with a bunch of similar hominid species. Uhh...

D: But maybe we kind of did, but we just don't realize.

A: Yeah. Th- and, uh... to some extent of course, the myths and stories might hint at that, and certainly the Neanderthal DNA you find mixed into the populations show that yeah, there was some uh, links between... Um, but do you think having the other as part of your society is an important thing? And that of course requires a tolerant, open society. Which is so important, even if you just talk about vanilla tolerance, about people having awkward political views

D: Yes.

A: Or really weird culinary t- uh, tastes. Yeah, that's the start of it, but we really need to have it about the most important things in our society otherwise it's not going to be a workable society. There are some people trying to make a homogeneous society, or insist that there are certain sets of rules and norms define what it means to be a proper citizen, and then nobody else can play. Okay. That's going to be a very small society, not going to be able to link to the rest of the world. I think actually that the twenty-first century is going to be a rather bad era for many of the homogeneous societies, if they insist on being homogeneous. Because, either they have to close themselves off from the world, which means they get very poor, or have t-, have to accept that 'I need to deal with people who are fundamentally different from me.'
D: Yeah.

A: And might be different in s- so important ways, that they can't actually function well.

D: And you know, that fear of the different is something that I seriously hope people finally /get over/.

A: Yeah. But it's an interesting question, how to do that.

D: Hmm.

A: So one interesting possibility might be, we need an enhancement to do that. So, Ingmar Persson and Julian Savulescu wrote a really interesting book called Unfit For The Future, where they argued, 'maybe we need moral enhancement, before we get all these powerful cognitive enhancements.' They got a little worried by my uh, work and Nick Bostrom's work. Uh, they were seeing global catastrophic risk in enhanced humans, "Wait a minute, what about morality, what about being able to have compassion?" And they realized that, as philosophers, philosophers have been going on for thousands of years about this, with limited effect.

D: Yes.

A: Some effect, but limited. Maybe we need biomedical enhancement of compassion and other things...

D: Indeed.

A: This was of course also an outrageous philosophical claim which led to a lot of interesting discussion.

D: (laughing) Because you can philosophize, you know, forever and a day, but if you're not prepared to actually put it into some kind of action, what happens?

A: Yeah. But it's useful to think a little bit before you do it, about 'is there something obviously broken with it?'. And it turns out that the most obvious ways of doing it, actually have problems. And some of them are very interesting and non-trivial, which keeps the philosophers happy. But to the biotechnologists,

D: Yeah.

A: this might be a way of getting into a rut, there might be ways of getting around it. Oxytocin, for example, seems to help us be nice and interesting to our in-group, but might make us slightly less kindly to the out-group. If only we could make oxytocin didn't have that property. It's not given that uhh, just oxytocin has to function like that.

D: Sorry, um, I usually pronounce it 'oxytocin', it took me a minute there to-

A: Ah yes, Ah, yeah, yeah, sorry.

D: Yeah.

A: Oxytocin, I just blame my Swedishness.
D: It's ok.

A: Now, the interesting thing here is that, you could argue that maybe making yourself nicer biomedically might miss authenticity. Indeed, when you ask people whether they would want to enhance things, only 9% of the people you ask, uh, say that they would take a kindness pill.

D: ...hmmm.

A: In some sense, I think, they are wrong. Uh, if we, if uh, existent kindness pills we probably ought to be taking anyway-

D: I think that there are some people who revel in being a little bit mean sometimes.

A: Yeah. B- but you want to be /able/ to be mean, rather than to /have/ to be mean.

D: Yes.

A: Uh, now, um, the interesting thing is, most of the worry here is that if I'm kind because I took a pill, am I truly kind? And as a consequentialist ethicist, I will say "Ha! If you're helpful, that's great. If you do it because of a pill, or because of your intention, doesn't matter too much." But of course, the ontologist would be frowning at this bollocks, "Oh no no no, you need to actually think the right thoughts, and if your emotions tell you to be kind, but your thoughts are not telling you to be kind, maybe you should

not not be kind." At this point, the philosophers end up in a brawl, unless they've taken those kindness pills.

D: Which is of course very unkind of them! Yeah! (laughing)

A: Maybe, yes! And both of us say, this is about truth, beauty and love and all the important things, we need to protect things. The real problem is of course, defining kindness is something that's very social, it's not just a warm fuzzy feeling. It's something we've learned from others, and how we express it might be different in different cultures. So I think, while biomedical enhancement of compassion and tolerance is something I think we should be keeping our eyes on and maybe promote, there are many other ways we can do, we can create institutions that help minorities and safeguard various rights. This is where rights come in very very handy.

D: Yeah.

A: Even though I don't believe there are eternal truths existing out there, they- the universe, is such I think that we can set things up. There are effective shorthands we have, to making functioning societies.

Bioinformatics systems and simulations of the body

D: Hmmm. Okay, thanks for that, that was a good, a good ramble there. I'm going to bring us round now to another question that I um, quickly thought up on the train. And that is, What do you think to the notion... (and this is going back to when you were doing computational biology particularly), what do you think to the notion that a bioinformatic system can be developed to model the whole body, from genetics to macro-scale?

A: I think it's going to be tricky to get the data to really mod- do the modelling right. So, I have a bit of background in, um, computational neuroscience, I was trying to make neural models. And, to some degree, this has been super-successful, because we actually, uh, know what uh, ion channels on the surface of neurons lead to signals going fast... we even have nice measurement of the parameters. For some neurons. That had taken a fair bit of lab work to find. Later bioinformatics have found more of these parameters, but it's still rather hard to get all of the data. For example, to know whether two neurons, found in my brain, are going to have the same parameters or not. So the real problem is that, in many parts of the body, you need to somehow scan the parameters. The length of the neurons, the electrical conductivity of membranes, the chemical concentration, what genes are being expressed. This forms a very complicated pattern. I think one can make a very crude model, relatively easily.

D: Yeah.

A: And that can get refined, the question is whether we can then run the model to actually get the right result.

A: So, I have been working on brain emulation, I'm doing a one to one model of a brain that hopefully produces exactly the same kind of thinking activity. That's a much taller order, of course. I think many would be very happy just to have a model that got the biochemistry roughly right in the body,

D: Indeed.

A: and could tell if I eat a sugar cube, how does my insulin level go up and down? Now, that might be easier in some sense, but it still requires a lot of data. It's not necessarily impossible to do. I don't think there is any law of nature to say 'haha, that can't be done', like some laws in quantum mechanics actually prevent us from knowing certain things. But I do think it's going to be a cumbersome process to make these kinds of models. It's also really valuable. I think it's so valuable that we should probably be throwing enormous research resources into this direction, because having models like that-

D: Glad someone agrees with me! [laughing]

A: [laughing] Um... uh, and you can argue for this from a lot of angles. You could take the animal rights activist angle, and say 'oh, if we could just use those models instead of those lab animals to test things, we would save a lot of lab animals.'

D: Yeah.

A: That's a nice idea and all but it's going to take an awful lot of lab animals to get the model, so I am not entirely certain you end up on the plus-side here, except real long-term. Another aspect is of course, for research, to actually be able to run a lot of experiments and simulations, and try to understand the system truly.

D: Mmhm.

A: In many cases, people say, "yeah, you need to understand it first, and then you can make a model." but sometimes, you have data and try to understand it by having an accurate model of the data, and try to do your experiments on that. This gets in...

D: Consider just how much data there are. In... in those massive databases all the universities are holding, they can't process even one percent of it!

A: Yeah. A-And, one of the big problems is, a lot of the data is also not exactly the data you want.

D: Yeah.

A: One of the big revelations, I think, of the past few years, has been that many of the internet giants can make fairly good use out of crappy data.

D: Hmm.

A: The way I'm clicking on the webpages is certainly telling important things about who I am, and w- m- what my interests are, but there's still a fair bit of noise. From time to time, I accidentally bump my mouse. Accidentally, I misread a word and do something wrong. That doesn't matter if I have enough data. If I am an internet giant, we can collect all the data, and get averages for all this f- d- fairly good, at

least in many domains. The question is, does this apply to the relevant domains in biology? And, it's a deep scientific question, I think we understand enough biology to be able to say that, some parts of the body, we may be able to do this fairly straightforwardly. In others, who knows? Uh, so I think the real challenge here, is making something that also works and then scales. Because, you need to integrate both models of genetics with models of cell biology, with models of tissue dynamics, with hydrodynamics and chemistry, that's a lot of disciplines.

Prof Sandberg's position on bioinformatics

A: Yeah. So, my, uh, to sum up, I think it's doable, it's really worthwhile to do, and it's not going to be easy. But, but, but, that a scientific project is not easy, doesn't mean it can't be done. If you think about the Human Genome Project, that was ridiculous when it was proposed,

D: Yeah.

A: and it remained fairly ridiculous for a long time until it finally just got done. Then of course, the real revelation later turned out to be that, 'Okay, we have the genome, and actually it's not telling us as much as we'd like to know. We didn't understand that before, that was an important revelation to tell. But I think, this kind of big science project's very interesting because they require a lot of very different disciplines working together. And that is hard, because the cell biologists don't normally talk to the tissue modelers, and uh, heaven knows what the behavioral modelers, uh, have in terms of data. To make this stuff compatible with each other is also quite tricky.

D: Oh, indeed. Um, I would like to see a system that really does holistically bring together modules that deal with every aspect of the body, and how it works. Um, and so we can have modules for things like transcription and translation, uh, protein folding, and... lipids, and y'know, all the different components of cells. And then you would have them dealing with the structural integrity of the body, um, and the cytoskeleton, things like that. You'd have them dealing with um... the nutrition, microbiome, and... the electrical properties of cellular um, y'know, membranes, and of... how that interacts on a more macro scale in organs of the body.

A: Mmm, mm, yeah. You, you might want to use general adversarial networks in different organs.

D: Exactly, that's one possibility. Um, I would also, though, caution against the overuse of... um, like, machine learning and neural networks, because I have a feeling from... from what I've seen so far of their use, that they are a bit of a cop-out in some regards...

A: [chuckling]

D: It's like saying, 'Oh, I can't be bothered to think about this the hard way, I'm just gonna let the computer do it', which is exactly what we're doing with a system like this, but if you... um, let the computer train itself, and you can't see what it's trained itself properly, because it's all written in some

hyper-optimized computer language, you might not understand why the computer is then doing something very strange later down the line.

A: Yeah. And biology has also turned out to be tough for machine learning to deal with. So, there is a lot of interest of course still in doing drug design using neural networks, and it's only very recently there have been one, uh, somewhat useful result. On the other hand, protein design using machine learning seems to be going really well.

D: Hmm.

A: So, it might turn out that we are going to find that there are some domains which are going to be helped much more. But generally, what you can simulate, you can turn loose one of these reinforcement learning agents on optimizing.

D: Yeah.

A: So, once we get a really good body simulation or good tissue simulation...

D: That's when you start adding more machine learning, yes.

A: But first we need to get that good simulation. That's going to be annoying and awkward, because we need a lot of data, we need disciplines talking to each other... but I think this is kind of doable.

D: Hmm.

Simulating brains and Schrödinger's Pain

A: Uh, we have seen that to some extent with the Human Brain Project in Europe. So although the political drama has been kind of uh, distracting, it actually has created ways of doing data warehousing with very different forms of data and disciplines together. Because it had all the money to go to all of the cognitive neuroscientists of Europe to talk to each other because they had all of the money. And that was actually helpful at least in that respect. Now we've got so much more interoperable systems. Still far from what we would like to have. The same thing for the body would of course be a rather interesting challenge.

D: Indeed. So, I know you said you had a model of the brain coming together. How complete would you say that is, in terms of being able to say 'Alright, this is ready to, y'know, pit against the actual human brain, and spot the difference?'

A: I think we're a few decades away from that. So, basically what we can do is, making neural simulations that are indistinguishable from the real thing. Now, I have seen actual experimental neuroscientists fail the Turing Test with even fairly simple neural models. So, that works out well. And you can certainly put up bigger network simulation that produce activity which looks suspiciously brain-like.

D: Yeah.

A: Except that they don't have any content. They're just random synaptic connections instead of the actual pattern in a brain, which contains the actual knowledge and the actual peculiarities which make my visual cortex operate different from my auditory cortex.

A: Right now, we have basically, we have a completed connectome of the Drosophila brain. Or at least, half of a Drosophila brain. So, half of a fruit fly. Uh, we have known since the Eighties of the three hundred and three nerve cells in the C. elegans worm, and how they're connected, but we still don't know the synaptic weights there, because it's very hard to do the experiments. So, we're kind of waiting to better experiment in order to get that data out of biology, so we can then start running simulations to see, can we actually get this to work or are we missing something else that's really fundamental?

D: Hmmm... And obviously, when you look at Drosophila and C. Elegans, they're relatively simple brains, obviously, there's not much room in there for anything more complex, um, and it'll be very intriguing to see as we get better and better models and we go up the uh, family tree of life, um, at what point do we start to see any divergence between what we thought was a good model that models everything, and what we actually get with a conscious being?

A: Mmhmm. Yeah. Uh, I uh, personally think that once you get a working mouse model, that really works well, then when you scale it up to higher mammals, you'll probably get to see if it keeps on scaling, I think that

D: Mmhm

A: If you can get your mouse to work well enough, then I think you can do it for a human brain. The problem I think that might be hidden in this, maybe the simulation has a hangover that maybe not be obvious from the outside, until you actually run your first brain emulation and they complain about a perpetual hangover. At that point, you have a deep ethics problem that's going to be really awkward. But of course, even before there, they're not allowed to pinch a mouse, in the lab, the tail of a lab mouse for no good reason in the lab. That's just cruel, that's against the Home Office regulations for how to treat the lab animals...

what about my virtual mouse? If I pinch its virtual tail, and then it gives up a little virtual squeal, I think personally that I don't know whether it's conscious, and that we should probably treat it as I would treat a flesh and blood mouse, and say that mm, perhaps I should abstain from messing with it.

D: Hmm, so there you have the interesting problem of, like, how do you measure the response, um, of a simulation? Without turning the simulation into a being.

A: Yeah. And in some cases, you might, uh, cheat. For example, you could imagine removing all of the pain systems of the virtual mouse. You don't simulate them. Um, or you might actually stop the signals when they reach the pain se- systems. In some cases, I think we can actually do amazing things with virtual lab animals, that would both be avoiding pain and suffering, and uh, would be very efficient. The problem is, animals are complex. Suffering is probably not a simple thing at all. You can suffer without pain, for sure, if you are feeling socially excluded. So, what about lonely lab mice, etcetera?

D: And what about lonely simulations?

A: Oh, yes. Yeah. And, uh, there is a whole host of kind of virtual animal ethics that needs to be done, but I don't think it's necessarily n-, prof, more profound than normal animal ethics in the lab. I think we need to care about it and do it right, but that's kind of it. There's some aspects that get weirder, because you can start your simulation, stop your simulation, you can even backtrack it. 'Oh, that looks painful and bad, let's go back to the state before that, and not do it.'

D: Yeah, in which case, are we then removing that memory? And did it never happen?

A: Yeah, exactly, and I can easily grab you some philosophers around here that say 'yeah, no it didn't happen', and others who will say, 'no, that has happened and uh, it has always happened.'

D: Then it's Schrödinger's pain I suppose, isn't it! [laughing]

A: Oh yeah! And so, we're certainly not getting away from weird ethics. But, even having decent simulations around must also allow you to do a lot of other interesting things. Uh, you could use machine learning methods to optimize and figure out how to do many new things. You could figure out how to influence an organism to make it healthier, or unfortunately, do the opposite. So, there's quite a lot of things that you can do, once you have the simulation. Making simulations, though, is tough.

Artificial Intelligence from a biological perspective

A: So, generally I think uh... the future looks pretty bright. For everything you can simulate, you can probably train one of those uhh, Google uhh, reinforcement learning agents,

D: Yeah.

A: to be really powerful. That also of course might be scary, but uhh, that's very promising. The problem is if we're getting to the simulation right now, requires a lot of human skill. Whether we can automate making simulations of complicated biology, that's another matter.

D: Yeah, when you start automating the creation of the automation.

A: Yep. Uh, well, this is what we care rather a lot about here in the office, because we're doing a lot of work on the AI safety and AI capability.

D: Hmm. Once-

A: Mostly...

D: Yeah, once you've got AIs producing AIs, that's when you're really getting into Ray Kurzweil's territory, isn't it. [chuckling].

A: Yeah. Um...

D: [chuckling]

A: But it's not given that the AIs necessarily are more general AI-creating AIs. So one interesting idea Eric Drexler has, is one called comprehensive AI services. For most applications, you don't want the little box that sits and does everything. No, you want a cameraman robot that is photographing you, you want a robot that generates music, you don't really want an assistant that can do everything in a slightly crappy way. Rather, you take your general system, and use it to generate specialized systems that are really great at some stuff.

D: Yeah.

A: Then you string them together, so you might want to have user interface uh... system that is really good at understanding what use it's for, with an engineer system that is good at coming up with engineering solutions, then a marketing system that sells it to people.

D: yeah. A bit like in Star Trek: The Next Generation, there were places and uses both for the ship's computer and for Lieutenant Commander Data.

A: Yes.

D: And they had different purposes and reasons to exist as AIs.

A: And, uh, Commander Data is interesting as a person or autonomous being, but maybe not the best solution to computational problems.

D: Hmm. Exactly. And um... I think that's where a lot of um... the differences between AIs might come down to, are you going to have that specialist or are you going to have a generalist who isn't that great at specific tasks, but... can do plenty of different things.

A: Uh, and you can vary along a lot of weird dimensions. I'm actually working on a paper on the space of possible minds, trying to map things out and there's some very

D: Hmm

A: strange things out there. Normally, we tend to think about them in a fairly anthropomorphic way of, uh, like, beings.

D: Yes.

A: But actually, could have oracle things, like the Google search engine, where you ask a question and it just generates answers. You can have it be services, which are a bit more like the Amazon cloud. It doesn't want anything, it just does whatever you plug it into. But, you can also have weird conglomerations that, uh, have perspectives of the world that are so utterly alien, that there is very little we can say about them.

D: But you can kind of predict them just by, as you say, mapping out possibilities.

A: Yeah. And uh, in some cases you can do fairly rigorous analysis of super-intelligences, that doesn't exist yet. So, Marcus Hutter defined an interesting one called the AIXI, which is basically the smartest entity that can exist. You can write down equations describing it on a post-it note, and you can prove things like, it doesn't believe in its own existence.

D: Right.

A: There's a zero probability to its own existence. It can write, if you give it the right input, it will write the best possible book about Cogito Ergo Sum, but it will never believe it exists.

D: How strange...

A: it's very strange. Um, by practical standards, you could say this is insane. But, it's as smart as anything can be.

D: So maybe, smart and insane are two sides of the same coin, woooo! [laughing]

A: [laughing] I think people like saying that, because it's very reassuring, uh, it assumes that there's some kind of game-balance to the world. Okay, if you're really smart, you need to have dumped some other stats, but-

D: [chuckling]

A: In practice, the, the real lesson is, being really smart is possible to combine with a very bizarre worldview.

D: Indeed.

A: And very weird values. And getting a handle on why some sensible intelligence is very different from very powerful intelligence, which is why we do so much work on AI safety.

D: Yeah.

A: We want to get the safe ones before we get to the powerful ones.

D: Um, so you're saying that um, you've uh... been sort-of doing this ahead of time, to see what AIs might exist in the future that don't currently... um... I suppose if we sort of look at this from a biological perspective, what the furry fandom has been doing, um, and... y'know, from a more sort of spiritualist perspective, you can say the otherkin and the therianthropes have as well, uh, is mapping out what species might come to exist,

A: Mmm...

D: before they ever do.

A: Yeah, yeah. Uh, although, if we get to a future where we have a kind of post-scarcity economics and a full morphological freedom, what kind of beings would we like to one day want to turn into? What kind of societies would make sense for them? Uh, this needs to be experimented with, the problems need to be noticed better before time, and then of course in any case, one can actually work it out.

D: Yes.

A: The critics and Luddites are kind of an unpaid focus group in many ways.

D: Hmm, and it's I suppose quite productive to look at them that way, because you can basically get ideas for free.

A: Oh yeah. Oh yeah. They're very kindly kind of giving them very helpfully to you. It's just that they're doing it kind of loudly and in a rather rude way, but... it's helpful.

D: Mmhm. That's certainly a positive viewpoint to have on it, so... Uh, I think that's the main bunch of questions asked... umm, at this point, we can ramble on a bit if you like.

A: Yup.

Intercepting exponential development

D: Umm, so I'm just going to sort of prod at your books.

A: Yeah, and I'm very happy to do that. Um, so I'm working on my own book, and um, this of course gives me an excuse to read everybody else's books. This is also a problem because there's so much interesting stuff here so I'm not getting done.

D: Mmhm.

A: yeah. So on one hand, we have Olle Häggström's Here Be Dragons, which is all about, how do they reason well about emerging technologies that don't exist yet? And he's pointing out that uh, ah, he was a member of the, the... Swedish Academy, um, for Sciences and involved in a grant- giving grants to research, and they're told to only focus on academic excellence, not whether they make the world better. Ah-

D: That sounds very narrow minded to me.

A: Totally. And, he felt it was stupid, and everybody he talked to said that's stupid, but

D: Hmm.

A: Updating the rules turned out to be very very hard. So that got him to write this book.

D: That sounds very interesting; so this appears to be dealing with um, the... acceleration of technological development and um, its exponential impacts on people's lives and that.

A: Yeah.

D: Yeah.

A: And um, then of course, the real challenge is, can you reason well about this and uh, since it's very uncertain, and for that, you need to develop various tools and ways of reasoning about it, there are better ways than just guessing.

D: Yeah.

A: And that is partially what is described in there. I find it a very good look at what we do at our institute. Because it's written by somebody outside the institute looking in, so he can actually point out what he thinks are stupid ideas, and which ones he actually think has some merit.

D: And uh, I suppose it's your job to think about these things and to worry about them, and to write papers about them and so on. And here I am, going through life as these technological developments accelerate, I'm just kind of picking and choosing things around me that are developing faster than I can keep track of, because they're useful to me, and putting them into action, and sometimes I'm looking at those trajectories and I'm thinking about, well, 'logically, um, y'know, computing power will be, you know, x number of exponents better by a given year. So, I should try to develop systems that will intersect with that capability of computing, by that year that we expect it would take that amount of time to develop the systems.

A: Yep, yep.

D: So, yeah, I tried to do that rather than aiming at the present day capabilities because, they'll be in the past by the time it's actually developed.

A: Exactly. What one needs to kind of always aim a little bit ahead of the target because it's moving.

D: Yes.

A: And understanding that is of course tricky, so another book related to that, is this one: Vaclav Smil's Growth. Honestly I haven't read this one yet, it just arrived a few days ago and I let one of my students do it. But he's kind of awesome generally, he's doing these enormously data-driven looks at things like in the future, energy and resources, and this is about growth. He's way more pessimistic than I am, but- because he thinks growths always have to end, and usually they end badly, etcetera.

D: Yeah.

A: But what I love about it, is it's trying to look at all sorts of growth processes. Uh, because quite often, you get people specializing on one area now, or question. And um, whether that applies to other domains or not, they don't care. But in practice, we want to find is the robust features of reality, like if you're running a company, okay, you don't know everything going on in the world, but you want to see those patterns that you need to pay heed to.

D: Yes.

A: What are those generally ethical patterns? Obviously, growth processes are one of the most obvious ones.

D: And that's exactly what I was talking about, yes. [short laugh] Spot on!

Interdisciplinary research

A: Uh, so I have uh, this kind of very wide uh, range of topics, uh, everything from how do squids and octopi think…

D: Mmhm...

A: Which I honestly don't understand anything, over to of course, lovely questions like astrophysics. Um, uh, astrophysics is a wonderful subject, uh, because stars and planets are me- much more well behaved than biology. They have much less individuality, actually.

D: Hmmm. Um, on the bright side... um [laughing] The very bright side, if you look at them directly, um, they... there's sort of ways that you can actually combine what you learn from astrophysics and, y'know, cephalopods and... you know, growth in the future, it... it produces lots of interesting new avenues of thought, when you just start to bring together all these different random areas of research. Um, I think holism in that way... uh, is very beneficial to society...? Um, and that we should do more interdisciplinary thinking.

A: Yeah.

D: Because it pays off in so vast, um, a way, every time, you know, every time anyone does this sort of research.

A: When it works. So that is the tricky part. Everybody says they like interdisciplinary work, but quite a lot of interdisciplinary work is actually relatively crappy,

D: Okay.

A: because it's very hard to find those general patterns. When you find them, they are revolution.

Think about Charles Darwin and the theory of Evolution.

D: Alright. I'll rephrase. I- What I meant was, it works really well, when you get onto something that actually fits together.

A: Yes.

D: Yeah.

A: And the problem is, whether you will find something that fits together or not, is going to be a... very much a matter of serendipity. Sometimes it happens, sometimes it doesn't. There are some things that seem to just mysteriously fit well together and also, there are cases when different fields are moving together. Right now, for example, information theory and physics are coming together in an amazing way.

D: Yes.

A: Basically, it turns out that information might be a real physical thing, that you can explain both black holes and thermodynamics in terms of information; but we don't quite know how it is supposed to work yet, it's just that these things are moving together more and more, and it might very well be that some brighter student is coming out today with an idea that gives them a Nobel prize a few years down the line. People have noticed that thermodynamics and entropy, are in some sense substance-independent.

D: Yes.

A: So, there have been people saying, 'ooh, we should apply this to history for example.' And usually, that produces a very very crappy history, because naïvely putting thermodynamics into history doesn't work, it's an open system

D: Yeah.

A: etcetera. So, figuring out what theories apply in a substance independent matter, is a really interesting thing, and we see the same thing in computation, of course. The theory of uh... kind of complexity classes in uh, computer science... um, it's an amazing example because it's not so much about computation as, what questions are hard?

D: Yeah.

A: And that's dependent on whether there's brains or computers trying to solve it.

D: And how do these questions and how does all the information in the universe, link? How does it all fit together and like, how do you get cause and effect?

A: Yeah. Now,

D: Yeah.

A: I think that in this particular part of the discussion, we have to take in several deeply interdisciplinary things and also put them a little bit too close together. I have a feeling that it might turn out that some models are actually not related to each other in the first place. But that's the interesting thing about trying to be properly interdisciplinary. You never know really what things do tend to abut each other in good ways.

D: Yeah.

A: I found for example that there is some aspect of argumentation theory, and reliability theory in engineering, that one is philosophy and about how to have a good discussion, the other one is about building machines that don't break down. You can actually use some theorists from engineering, and slot them straight into the philosophical discussion.

D: Find out how the cogs of your machine are arguing with each other. [laughing]

A: Exactly. Now, the interesting challenge is that building up this interdisciplinary discourse, it takes people being curious about each other's topics. It's not just that uh... you check your boxes to get uh, your research grant, but it's actually that the lawyer needs to be curious, 'how do you astrophysicists actually calculate that stuff?' and the others realize, 'wait a minute, you lawyers don't even have this concept,

D: Hmm...

A: how do you even get through the day in the office?' Now, when you start noticing these things, that's where sparks really start flying.

D: yeah. And... it's, it's interesting if you then get back to say, y'know, law schools, and um... you show them these mathematical concepts at their most basic, and it feeds back into how that is then taught through the system, and eventually, you see the laws changing.

Law, politics, and science meet morphological freedom

A: yeah. Uh, and... law is interesting because you get both several experiments in how to do it. You have the simple law systems where parliaments just define it all, you have American and British common Law systems, where actually precedents does a lot of things, but that both of them actually seem to work is kind of really weird,

D: Hmm.

A: because um, when you have an argument with- between alternates, it seems obvious that one of them should be totally broken.

D: Yeah.

A: Indeed, there was a recent paper looking at people reconstructing a legal system they had never heard of. Because they were ancient ones from Sumeria or China. And it turns out that most people were fairly good at coming up with what laws should be like in this system, because many laws are actually patterned more by what people think makes sense in a social environment, than the formal details of law.

D: Exactly. And, um, there was a recent example down in Australia, um, where... the... uh... um, I - I think it was that they were trying to... uh.... ban um, encryption.

Just, completely ban encryption. And... ther-there was a, an argument that essentially boiled down to, y'know, some-someone with common sense, um, I believe they were a scientist of some degree or other, said to um, the prime minister at the time, well, 'you are running up against the laws of mathematics.' Um, you know, if you're claiming you can ban encryption and at the same time, have data be safe, y'know, and um... the prime minister's response was ridiculously enough, um, 'the laws of mathematics don't apply here, the laws of Australia do.'

A: Yup, yup.

D: And, you know.

A: And it's a bit like Feynman may have uh, challenged this Australian just there when he pointed out that physics always overrules policy.

D: Yeah.

A: Yeah.

D: So.

A: But, the other problem is of course, quite often politicians need to make decisions that make sense within the sphere of law and politics, and they're not always the ones that are best based by evidence for anything. A few years back, here in the UK, for example, Dave Nutt got fired from the government's drug committee, because he kind of pointed out the actual risks of Ecstasy compared to, pastimes like riding. That was really annoying and uh, the fact that he had evidence to back him up, that was irrelevant. Now there was of course a big outcry and now of course he's more powerful than ever because he's not on the committee and can say whatever he wants. But actually I was told that the politician did the sensible thing, because, yeah, the fact that the /facts/ tell a certain story, doesn't necessarily mean policy has to be like that.

D: Yeah.

A: Yeah, now, it's kind of stupid if policy is, uh, diverging from reality. But in some cases, people really want certain things, maybe it's better to give them a crappy law.

D: Yeah.

A: Now, in the case of the drug laws that might harm a lot of people and society at large is a kind of serious thing to get right, but there are a lot of these annoying things for especially people in kind of more truth-oriented areas like science, when they suddenly realize that there are other considerations, the best product doesn't always win on the market. Sometimes, the, b-b- box it's coming in and marketing, is more important than the technical quality.

D: See also, VHS versus V2000 and Betamax.

A: Exactly.

D: Yeah.

A: And the real problem we have of course as a civilization is, can we get our priorities straight about at least the most important things? The less important things, if we flub them, well, that's neither here nor there, but if we flub our own survival, or main moral activities, yeah, then we might fail as a whole.

D: Indeed. Now, we, actually at the... um, sorry, just trying to remember what I was gonna say. [laughing]

A: [laughing]

D: Umm... we were talking about policies, weren't we, and law, that's right - at um, the Freedom of Form Foundation, we recently funded a, um, research from Ben Ramanauskas, who I believe is also linked in some way linked to Oxford University,

A: Oh…!

D: and um, he studied the differences between British and American law, um, with regard to body modification and morphological freedom, and um, thanks to us being able to fund him and um, a few other um, funds coming together as well. Um, this was on top of his previous research that he'd done on the laws in the UK in general with regard to body modification, and what happened with that Dr. Evil case in Birmingham.

A: Mmm, mmm, ja.

D: Yeah.

A: Yeah, that's very very sad, uh, that case, but it also shows that we need to think about, what alternative forms of organization do we have for the healthcare or body care system? It's not just enough to notice that there are some of these laws are bad and that they could do better, but we actually need to have some positive mentions about, 'so how /would/ we go about it?' There is a thought experiment in the debate about human enhancement, against the idea that enhancing people, that's against the goals of medicine, because the goals of medicine is only curing people. And then Eric Parnas pointed out, 'What if there was besides doctors, schmoctors? And schmoctors don't do medicine, they do schmedicine? At schmospitals? And uh, it looks very similar except that they are all doing enhancement instead of ever curing people.

D: Right.

A: Curing people would be wrong for a schmoctor, because that's against the goals of schmedicine. Now, doesn't seem to be any logical reason why you couldn't have this setup, uh, and yet it totally shows that you totally could have enhancement in society. More likely, of course, you could say, 'yeah actually, doctors know a lot about bodies and it's probably a good thing to have an overlap between these disciplines. And the same thing goes of course for a lot of other forms of body modification, but we have these artificial forms of separation partially because of historical reasons. We are kind of lucky that surgery got back in, into medicine, rather than being left to the barbers for a long time. But for a long time, it was left to the barbers because, uh, proper doctors did not cut people. Okay, they could just

have been that we could have had two disciplines now at hospitals, where you have the doctors and the surgeons.

D: Yeah.

A: And in the future, we might want to see, do we need more disciplines, do we want fewer, do we want to organize the certification and quality control in different ways?

D: That's um, a fascinating way to look at it.

A: Yeah...

Wrapping up

D: Thank you very much for taking part in this interview and um...

A: Well, thank /you/! And this is exciting stuff.

D: Okay, I would like to um, obviously chat with you some more, um, later in the day as well, um... I think at the moment, we should take a break.

A: Yeah, yeah.

D: Take stock, you know.

A: You want some tea or coffee?

D: A cup of tea would be lovely, do you have decaf?

A: yeah, yeah. Oh, yeah.

D: Thanks.

A: Let's go and investigate the kitchen because there is too many kinds. But I- I know where there is decaf, because I'm always getting it when I don't want it.

D: [laughing] So, there definitely is some. [laughing]

A: yeah. Yeah, that is the problem. [laughing]

D: Alright, I'm gonna stop the recording here...