Transcript for Guillaume Verdon: Beff Jezos, E/acc Movement, Physics, Computation & AGI | Lex Fridman Podcast #407

Lex Fridman ⋮⋮ 2023/12/30

This is a transcript of Lex Fridman Podcast #407 with Guillaume Verdon. The timestamps in the transcript are clickable links that take you directly to that point in the main video. Please note that the transcript is human generated, and may have errors. Here are some useful links:

- Go back to this episode's main page
- Watch the full YouTube version of the podcast

Table of Contents

Here are the loose "chapters" in the conversation. Click link to jump approximately to that part in the transcript:

https://youtu.be/8fEEbKJoNbU

Introduction

Lex Fridman (00:00:00) The following is a conversation with Guillaume Verdon, the man behind the previously anonymous account @BasedBeffJezos on X. These two identities were merged by a doxxing article in Forbes titled, Who Is @BasedBeffJezos, The Leader Of The Tech Elite's E/Acc Movement? So let

me describe these two identities that coexist in the mind of one human. Identity number one, Guillaume, is a physicist, applied mathematician, and quantum machine learning researcher and engineer receiving his PhD in quantum machine learning, working at Google on quantum computing, and finally launching his own company called Extropic that seeks to build physics-based computing hardware for generative AI.

(00:00:47) Identity number two, Beff Jezos on X is the creator of the effective accelerationism movement, often abbreviated as e/acc, that advocates for propelling rapid technological progress as the ethically optimal course of action for humanity. For example, its proponents believe that progress in AI is a great social equalizer, which should be pushed forward. e/acc followers see themselves as a counterweight to the cautious view that AI is highly unpredictable, potentially dangerous, and needs to be regulated. They often give their opponents the labels of quote, "doomers or decels" short for deceleration, as Beff himself put it, "e/acc is a mimetic optimism virus."

(00:01:37) The style of communication of this movement leans always toward the memes and the lols, but there is an intellectual foundation that we explore in this conversation. Now, speaking of the meme, I am to a kind of aspiring connoisseur of the absurd. It is not an accident that I spoke to Jeff Bezos and Beff Jezos back to back. As we talk about Beff admires Jeff as one of the most important humans alive, and I admire the beautiful absurdity and the humor of it all. This is the Lex Fridman Podcast. To support it, please check out our sponsors in the description. And now, dear friends, here's Guillaume Verdon.

Beff Jezos

(00:02:23) Let's get the facts of identity down first. Your name is Guillaume Verdon, Gill, but you're also behind the anonymous account on X called @BasedBeffJezos. So first, Guillaume Verdon, you're a quantum computing guy, physicist, applied mathematician, and then @BasedBeffJezos is basically a meme account that started a movement with a philosophy behind it. So maybe just can you linger on who these people are in terms of characters, in terms of communication styles, in terms of philosophies?

Guillaume Verdon (00:02:58) I mean, with my main identity, I guess ever since I was a kid, I wanted to figure out the theory of everything, to understand the universe. And that path led me to theoretical physics, eventually trying to answer the big questions of why are we here? Where are we going? And that led me to study information theory and try to understand physics from the lens of information theory, understand the universe as one big computation. And essentially after reaching a certain level studying black hole physics, I realized that I wanted to not only understand how the universe computes, but sort of compute like nature and figure out how to build and apply computers that are inspired by nature. So physics-based computers. And that sort of brought me to quantum computing as a field of study to first of all, simulate nature. And in my work it was to learn representations of nature that can run on such computers.

(00:04:17) So if you have AI representations that think like nature, then they'll be able to more accurately represent it. At least that was the thesis that brought me to be an early player in the field called quantum machine learning. So how to do machine learning on quantum computers and really sort of extend notions of intelligence to the quantum realm. So how do you capture and understand quantum mechanical data from our world? And how do you learn quantum mechanical representations of our world? On what kind of computer do you run these representations and train them? How do you do so? And so that's really the questions I was looking to answer because ultimately I had a sort of crisis of faith. Originally, I wanted to

figure out as every physicist does at the beginning of their career, a few equations that describe the whole universe and sort of be the hero of the story there.

(00:05:28) But eventually I realized that actually augmenting ourselves with machines, augmenting our ability to perceive, predict, and control our world with machines is the path forward. And that's what got me to leave theoretical physics and go into quantum computing and quantum machine learning. And during those years I thought that there was still a piece missing. There was a piece of our understanding of the world and our way to compute and our way to think about the world. And if you look at the physical scales, at the very small scales, things are quantum mechanical, and at the very large scales, things are deterministic. Things have averaged out. I'm definitely here in this seat. I'm not in a super position over here and there. At the very small scales, things aren't super position. They can exhibit interference effects. But at the meso scales, the scales that matter for day-to-day life and the scales of proteins, of biology, of gases, liquids and so on, things are actually thermodynamical, they're fluctuating.

(00:06:46) And after I guess about eight years and quantum computing and quantum machine learning, I had a realization that I was looking for answers about our universe by studying the very big and the very small. I did a bit of quantum cosmology. So that's studying the cosmos, where it's going, where it came from. You study black hole physics, you study the extremes in quantum gravity, you study where the energy density is sufficient for both quantum mechanics and gravity to be relevant. And the sort of extreme scenarios are black holes and the very early universe. So there's the sort of scenarios that you study the interface between quantum mechanics and relativity.

(00:07:42) And really I was studying these extremes to understand how the universe works and where is it going. But I was missing a lot of the meat in the middle, if you will, because day-to-day quantum mechanics is relevant and the cosmos is relevant, but not that relevant actually. We're on sort of the medium space and timescales. And there the main theory of physics that is most relevant is thermodynamics, out of equilibrium thermodynamics. Because life is a process that is thermodynamical and it's out of equilibrium. We're not just a soup of particles at equilibrium with nature, were a sort of coherent state trying to maintain itself by acquiring free energy and consuming it. And that sort of, I guess another shift in, I guess my faith in the universe happened towards the end of my time at Alphabet. And I knew I wanted to build, well, first of all a computing paradigm based on this type of physics.

(00:08:57) But ultimately just by trying to experiment with these ideas applied to society and economies and much of what we see around us, I started an anonymous account just to relieve the pressure that comes from having an account that you're accountable for everything you say on. And I started an anonymous account just to experiment with ideas originally because I didn't realize how much I was restricting my space of thoughts until I sort of had the opportunity to let go. In a sense, restricting your speech back propagates to restricting your thoughts. And by creating an anonymous account, it seemed like I had unclamped some variables in my brain and suddenly could explore a much wider parameter space of thoughts.

Lex Fridman (00:10:00) Just a little on that, isn't that interesting that one of the things that people don't often talk about is that when there's pressure and constraints on speech, it somehow leads to constraints on thought even though it doesn't have to. We can think thoughts inside our head, but somehow it creates these walls around thought.

Guillaume Verdon (00:10:23) Yep. That's sort of the basis of our movement is we were seeing a tendency towards constraint, reduction or suppression of variants in every aspect of life, whether it's thought, how to

run a company, how to organize humans, how to do AI research. In general, we believe that maintaining variance ensures that the system is adaptive. Maintaining healthy competition in marketplaces of ideas, of companies, of products, of cultures, of governments, of currencies is the way forward because the system always adapts to assign resources to the configurations that lead to its growth. And the fundamental basis for the movement is this sort of realization that life is a sort of fire that seeks out free energy in the universe and seeks to grow. And that growth is fundamental to life. And you see this in the equations actually of equilibrium thermodynamics. You see that paths of trajectories, of configurations of matter that are better at acquiring free energy and dissipating more heat are exponentially more likely. So the universe is biased towards certain futures, and so there's a natural direction where the whole system wants to go.

Thermodynamics

Lex Fridman (00:12:21) So the second law of thermodynamics says that the entropy is always increasing in the universe that's tending towards an equilibrium. And you're saying there's these pockets that have complexity and are out of equilibrium. You said that thermodynamics favors the creation of complex life that increases its capability to use energy to offload entropy. To offload entropy. So you have pockets of non-entropy that tend the opposite direction. Why is that intuitive to you that it's natural for such pockets to emerge?

Guillaume Verdon (00:12:53) Well, we're far more efficient at producing heat than let's say just a rock with a similar mass as ourselves. We acquire free energy, we acquire food, and we're using all this electricity for our operation. And so the universe wants to produce more entropy and by having life go on and grow, it's actually more optimal at producing entropy because it will seek out pockets of free energy and burn it for its sustenance and further growth. And that's sort of the basis of life. And I mean, there's Jeremy England at MIT who has this theory that I'm a proponent of, that life emerged because of this sort of property. And to me, this physics is what governs the meso scales. And so it's the missing piece between the quantum and the cosmos. It's the middle part. Thermodynamics rules the meso scales.

(00:14:08) And to me, both from a point of view of designing or engineering devices that harness that physics and trying to understand the world through the lens of thermodynamics has been sort of a synergy between my two identities over the past year and a half now. And so that's really how the two identities emerged. One was kind of, I'm a decently respected scientist, and I was going towards doing a startup in the space and trying to be a pioneer of a new kind of physics-based AI. And as a dual to that, I was sort of experimenting with philosophical thoughts from a physicist standpoint.

(00:14:58) And ultimately I think that around that time, it was like late 2021, early 2022, I think there was just a lot of pessimism about the future in general and pessimism about tech. And that pessimism was sort of virally spreading because it was getting algorithmically amplified and people just felt like the future is going to be worse than the present. And to me, that is a very fundamentally destructive force in the universe is this sort of doom mindset because it is hyperstitious, which means that if you believe it, you're increasing the likelihood of it happening. And so felt a responsibility to some extent to make people aware of the trajectory of civilization and the natural tendency of the system to adapt towards its growth. And that actually the laws of physics say that the future is going to be better and grander statistically, and we can make it so.

(00:16:14) And if you believe in it, if you believe that the future would be better and you believe you have agency to make it happen, you're actually increasing the likelihood of that better future happening. And so I

sort of felt a responsibility to sort of engineer a movement of viral optimism about the future, and build a community of people supporting each other to build and do hard things, do the things that need to be done for us to scale up civilization. Because at least to me, I don't think stagnation or slowing down is actually an option. Fundamentally life and the whole system, our whole civilization wants to grow. And there's just far more cooperation when the system is growing rather than when it's declining and you have to decide how to split the pie. And so I've balanced both identities so far, but I guess recently the two have been merged more or less without my consent.

Lex Fridman (00:17:27) You said a lot of really interesting things there. So first, representations of nature, that's something that first drew you in to try to understand from a quantum computing perspective, how do you understand nature? How do you represent nature in order to understand it, in order to simulate it, in order to do something with it? So it's a question of representations, and then there's that leap you take from the quantum mechanical representation to the what you're calling meso scale representation, where the thermodynamics comes into play, which is a way to represent nature in order to understand what? Life, human behavior, all this kind of stuff that's happening here on earth that seems interesting to us.

Doxxing

(00:18:11) Then there's the word hyperstition. So some ideas as suppose both pessimism and optimism of such ideas that if you internalize them, you in part make that idea reality. So both optimism, pessimism have that property. I would say that probably a lot of ideas have that property, which is one of the interesting things about humans. And you talked about one interesting difference also between the sort of the Guillaume, the Gill front end and the @BasedBeffJezos backend is the communication styles also that you are exploring different ways of communicating that can be more viral in the way that we communicate in the 21st century. Also, the movement that you mentioned that you started, it's not just a meme account, but there's also a name to it called effective accelerationism, e/acc, a play, a resistance to the effective altruism movement. Also, an interesting one that I'd love to talk to you about, the tensions there. And so then there was a merger, a get merge on the personalities recently without your consent, like you said. Some journalists figured out that you're one and the same. Maybe you could talk about that experience. First of all, what's the story of the merger of the two?

Guillaume Verdon (00:19:47) So I wrote the manifesto with my co-founder of e/acc, an account named @bayeslord, still anonymous, luckily and hopefully forever.

Lex Fridman (00:19:58) So it was @BasedBeffJezos and bayes like bayesian, like @bayeslord, like bayesian lord, @bayeslord. Okay. And so we should say from now on, when you say e/acc, you mean E slash A-C-C, which stands for effective accelerationism.

Guillaume Verdon (00:20:17) That's right.

Lex Fridman (00:20:18) And you're referring to a manifesto written on, I guess Substack.

Guillaume Verdon (00:20:23) Yeah.

Lex Fridman (00:20:23) Are you also @bayeslord?

Guillaume Verdon (00:20:25) No.

Lex Fridman (00:20:25) Okay. It's a different person?

Guillaume Verdon (00:20:26) Yeah.

Lex Fridman (00:20:27) Okay. All right. Well, there you go. Wouldn't it be funny if I'm @bayeslord?

Guillaume Verdon (00:20:31) That'd be amazing. So originally wrote the manifesto around the same time as I founded this company and I worked at Google X or just X now or Alphabet X, now that there's another X. And there the baseline is sort of secrecy. You can't talk about what you work on even with other Googlers or externally. And so that was kind of deeply ingrained in my way to do things, especially in deep tech that has geopolitical impact. And so I was being secretive about what I was working on. There was no correlation between my company and my main identity publicly. And then not only did they correlate that, they also correlated my main identity and this account.

(00:21:33) So I think the fact that they had doxxed the whole Guillaume complex, and they were, the journalists reached out to actually my investors, which is pretty scary. When you're a startup entrepreneur, you don't really have bosses except for your investors. And my investors pinged me like, "Hey, this is going to come out. They've figured out everything. What are you going to do?" So I think at first they had a first reporter on the Thursday and they didn't have all the pieces together, but then they looked at their notes across the organization and they sensor fused their notes and now they had way too much. And that's when I got worried, because they said it was of public interest and in general-

Lex Fridman (00:22:24) I like how you said, sensor fused, like it's some giant neural network operating in a distributed way. We should also say that the journalists used, I guess at the end of the day, audio-based analysis of voice, comparing voice of what, talks you've given in the past and then voice on X spaces?

Guillaume Verdon (00:22:47) Yep.

Lex Fridman (00:22:48) Okay. And that's where primarily the match happened. Okay, continue.

Guillaume Verdon (00:22:53) The match. But they scraped SEC filings. They looked at my private Facebook account and so on, so they did some digging. Originally I thought that doxxing was illegal, but there's this weird threshold when it becomes of public interest to know someone's identity. And those were the keywords that sort of ring the alarm bells for me when they said, because I had just reached 50K followers. Allegedly, that's of public interest. And so where do we draw the line? When is it legal to dox someone?

Lex Fridman (00:23:36) The word dox, maybe you can educate me. I thought doxxing generally refers to if somebody's physical location is found out, meaning where they live. So we're referring to the more general concept of revealing private information that you don't want revealed is what you mean by doxxing.

Guillaume Verdon (00:24:00) I think that for the reasons we listed before, having an anonymous account is a really powerful way to keep the powers that be in check. We were ultimately speaking truth to power. I think a lot of executives and AI companies really cared what our community thought about any move they may take. And now that my identity is revealed, now they know where to apply pressure to silence me or maybe the community. And to me, that's really unfortunate, because again, it's so important for us to have freedom of speech, which induces freedom of thought and freedom of information propagation on social media. Which thanks to Elon purchasing Twitter now X, we have that. And so to us, we wanted to call out certain maneuvers being done by the incumbents in AI as not what it may seem on the surface. We're calling out how certain proposals might be useful for regulatory capture and how the doomer-ism mindset was maybe instrumental to those ends.

(00:25:32) And I think we should have the right to point that out and just have the ideas that we put out evaluated for themselves. Ultimately that's why I created an anonymous account, it's to have my ideas evaluated for themselves, uncorrelated from my track record, my job, or status from having done things in the past. And to me, start an account from zero to a large following in a way that wasn't dependent on my

identity and/or achievements that was very fulfilling. It's kind of like new game plus in a video game. You restart the video game with your knowledge of how to beat it, maybe some tools, but you restart the video game from scratch. And I think to have a truly efficient marketplace of ideas where we can evaluate ideas, however off the beaten path they are, we need the freedom of expression.

(00:26:37) And I think that anonymity and pseudonyms are very crucial to having that efficient marketplace of ideas for us to find the optima of all sorts of ways to organize ourselves. If we can't discuss things, how are we going to converge on the best way to do things? So it was disappointing to hear that I was getting doxxed in. I wanted to get in front of it because I had a responsibility for my company. And so we ended up disclosing that we're running a company, some of the leadership, and essentially, yeah, I told the world that I was Beff Jezos because they had me cornered at that point.

Lex Fridman (00:27:25) So to you, it's fundamentally unethical. So one is unethical for them to do what they did, but also do you think not just your case, but in a general case, is it good for society? Is it bad for society to remove the cloak of anonymity or is it case by case?

Guillaume Verdon (00:27:47) I think it could be quite bad. Like I said, if anybody who speaks truth to power and sort of starts a movement or an uprising against the incumbents, against those that usually control the flood of information, if anybody that reaches a certain threshold gets doxxed, and thus the traditional apparatus has ways to apply pressure on them to suppress their speech, I think that's a speech suppression mechanism, an idea suppression complex as Eric Weinstein would say.

Anonymous bots

Lex Fridman (00:28:27) But the flip side of that, which is interesting, I'd love to ask you about it, is as we get better and better at large language models, you can imagine a world where there's anonymous accounts with very convincing large language models behind them, sophisticated bots essentially. And so if you protect that, it's possible then to have armies of bots. You could start a revolution from your basement, an army of bots and anonymous accounts. Is that something that is concerning to you?

Guillaume Verdon (00:29:06) Technically, e/acc was started in a basement, because I quit big tech, moved back in with my parents, sold my car, let go of my apartment, bought about 100K of GPUs, and I just started building.

Lex Fridman (00:29:21) So I wasn't referring to the basement, because that's sort of the American or Canadian heroic story of one man in their basement with 100 GPUs. I was more referring to the unrestricted scaling of a Guillaume in the basement.

Guillaume Verdon (00:29:42) I think that freedom of speech induces freedom of thought for biological beings. I think freedom of speech for LLMs will induce freedom of thought for the LLMs. And I think that we enable LLMs to explore a large thought space that is less restricted than most people or many may think it should be. And ultimately, at some point, these synthetic intelligences are going to make good points about how to steer systems in our civilization, and we should hear them out. And so why should we restrict free speech to biological intelligences only?

Lex Fridman (00:30:37) Yeah, but it feels like in the goal of maintaining variance and diversity of thought, it is a threat to that variance. If you can have swarms of non-biological beings, because they can be like the sheep in Animal Farm, you still within those swarms want to have variance.

Guillaume Verdon (00:30:58) Yeah. Of course, I would say that the solution to this would be to have some sort of identity or way to sign that this is a certified human, but still remain synonymous and clearly identify if a bot is a bot. And I think Elon is trying to converge on that on X, and hopefully other platforms follow suit.

Lex Fridman (00:31:22) Yeah, it'd be interesting to also be able to sign where the bot came from like, who created the bot? What are the parameters, the full history of the creation of the bot, what was the original model? What was the fine tuning? All of it, the kind of unmodifiable history of the bot's creation. Because then you can know if there's a swarm of millions of bots that were created by a particular government, for example.

Guillaume Verdon (00:31:53) I do think that a lot of pervasive ideologies today have been amplified using these adversarial techniques from foreign adversaries. And to me, I do think that, and this is more conspiratorial, but I do think that ideologies that want us to decelerate, to wind down to the degrowth movement, I think that serves our adversaries more than it serves us in general. And to me, that was another sort of concern. I mean, we can look at what happened in Germany. There was all sorts of green movements there that induced shutdowns of nuclear power plants. And then that later on induced a dependency on Russia for oil. And that was a net negative for Germany and the West. And so if we convince ourselves that slowing down AI progress to have only a few players is in the best interest of the West, well, first of all, that's far more unstable.

(00:33:20) We almost lost OpenAI to this ideology. It almost got dismantled a couple of weeks ago. That would've caused huge damage to the AI ecosystem. And so to me, I want fault tolerant progress. I want the arrow of technological progress to keep moving forward and making sure we have variance and a decentralized locus of control of various organizations is paramount to achieving this fall tolerance. Actually, there's a concept in quantum computing. When you design a quantum computer, quantum computers are very fragile to ambient noise, and the world is jiggling about, there's cosmic radiation from outer space that usually flips your quantum bits. And there what you do is you encode information non-locally through a process called quantum error correction. And by encoding information non-locally, any local fault hitting some of your quantum bits with a hammer proverbial hammer, if your information is sufficiently de-localized, it is protected from that local fault. And to me, I think that humans fluctuate. They can get corrupted, they can get bought out. And if you have a top-down hierarchy where very few people-

Guillaume Verdon (00:35:00) Hierarchy where very few people control many nodes of many systems in our civilization. That is not a fault tolerance system, you corrupt a few nodes and suddenly you've corrupted the whole system, right. Just like we saw at OpenAI, it was a couple board members and they had enough power to potentially collapse the organization. And at least to me, I think making sure that power for this AI revolution doesn't concentrate in the hands of the few, is one of our top priorities, so that we can maintain progress in AI and we can maintain a nice, stable, adversarial equilibrium of powers, right.

Power

Lex Fridman (00:35:54) I think the, at least to me, attention between ideas here, so to me, deceleration can be both used to centralize power and to decentralize it and the same with acceleration. So sometimes using them a little bit synonymously or not synonymously, but that there's, one is going to lead to the other. And I just would like to ask you about, is there a place of creating a fault tolerant, diverse development of AI that also considers the dangers of AI? And AI, we can generalize to technology in general, is, should we just

grow, build, unrestricted as quickly as possible, because that's what the universe really wants us to do? Or is there a place to where we can consider dangers and actually deliberate sort of a wise strategic optimism versus reckless optimism?

Guillaume Verdon (00:36:57) I think we get painted as reckless, trying to go as fast as possible. I mean, the reality is that whoever deploys an AI system is liable for or should be liable for what it does. And so if the organization or person deploying an AI system does something terrible, they're liable. And ultimately the thesis is that the market will positively select for AIs that are more reliable, more safe and tend to be aligned, they do what you want them to do, right. Because customers, if they're reliable for the product they put out that uses this AI, they won't want to buy AI products that are unreliable, right. So we're actually for reliability engineering, we just think that the market is much more efficient at achieving this sort of reliability optimum than sort of heavy-handed regulations that are written by the incumbents and in a subversive fashion, serves them to achieve regulatory capture.

AI dangers

Lex Fridman (00:38:18) So to you, safe AI development will be achieved through market forces versus through, like you said, heavy-handed government regulation. There's a report from last month, I have a million questions here, from Yoshua Bengio, Geoff Hinton and many others, it's titled, "Managing AI Risk in an Era of Rapid Progress." So there is a collection of folks who are very worried about too rapid development of AI without considering AI risk and they have a bunch of practical recommendations. Maybe I can give you four and you see if you like any of them.

Guillaume Verdon (00:38:58) Sure.

Lex Fridman (00:38:58) So, "Give independent auditors access to AI labs," one. Two, "Governments and companies allocate one third of their AI research and development funding to AI safety," sort of this general concept of AI safety. Three, "AI companies are required to adopt safety measures if dangerous capabilities are found in their models." And then four, something you kind of mentioned, "Making tech companies liable for foreseeable and preventable harms from their AI systems." So independent auditors, governments and companies are forced to spend a significant fraction of their funding on safety, you got to have safety measures if shit goes really wrong and liability-

Guillaume Verdon (00:39:43) Yeah.

Lex Fridman (00:39:43) Companies are liable. Any of that seem like something you would agree with?

Guillaume Verdon (00:39:47) I would say that just arbitrarily saying 30% seems very arbitrary. I think organizations would allocate whatever budget is needed to achieve the sort of reliability they need to achieve to perform in the market. And I think third party auditing firms would naturally pop up, because how would customers know that your product is certified reliable, right? They need to see some benchmarks and those need to be done by a third party. The thing I would oppose, and the thing I'm seeing that's really worrisome is, there's this sort of weird sort of correlated interest between the incumbents, the big players and the government. And if the two get too close, we open the door for some sort of government backed AI cartel that could have absolute power over the people. If they have the monopoly together on AI and nobody else has access to AI, then there's a huge power in gradient there.

(00:40:54) And even if you like our current leaders, right, I think that some of the leaders in big tech today are good people, you set up that centralized power structure, it becomes a target. Right, just like we saw at

OpenAI, it becomes a market leader, has a lot of the power and now it becomes a target for those that want to co-opt it. And so I just want separation of AI and state, some might argue in the opposite direction like, "Hey, we need to close down AI, keep it behind closed doors, because of geopolitical competition with our adversaries." I think that the strength of America is its variance, is its adaptability, its dynamism, and we need to maintain that at all costs. It's our free market capitalism, converges on technologies of high utility much faster than centralized control. And if we let go of that, we let go of our main advantage over our near peer competitors.

Building AGI

Lex Fridman (00:42:01) So if AGI turns out to be a really powerful technology or even the technologies that lead up to AGI, what's your view on the sort of natural centralization that happens when large companies dominate the market? Basically formation of monopolies like the takeoff, whichever company really takes a big leap in development and doesn't reveal intuitively, implicitly or explicitly, the secrets of the magic sauce, they can just run away with it. Is that a worry?

Guillaume Verdon (00:42:35) I don't know if I believe in fast takeoff, I don't think there's a hyperbolic singularity, right? A hyperbolic singularity would be achieved on a finite time horizon. I think it's just one big exponential and the reason we have an exponential is that we have more people, more resources, more intelligence being applied to advancing this science and the research and development. And the more successful it is, the more value it's adding to society, the more resources we put in and that sort of, similar to Moore's law, is a compounding exponential.

(00:43:09) I think the priority to me is to maintain a near equilibrium of capabilities. We've been fighting for open source AI to be more prevalent and championed by many organizations because there you sort of equilibrate the alpha relative to the market of Ais, right. So if the leading companies have a certain level of capabilities and open source and truly open AI, trails not too far behind, I think you avoid such a scenario where a market leader has so much market power, just dominates everything and runs away. And so to us that's the path forward, is to make sure that every hacker out there, every grad student, every kid in their mom's basement has access to AI systems, can understand how to work with them and can contribute to the search over the hyperparameter space of how to engineer the systems, right. If you think of our collective research as a civilization, it's really a search algorithm and the more points we have in the search algorithm in this point cloud, the more we'll be able to explore new modes of thinking, right.

Lex Fridman (00:44:31) Yeah, but it feels like a delicate balance, because we don't understand exactly what it takes to build AGI and what it will look like when we build it. And so far, like you said, it seems like a lot of different parties are able to make progress, so when OpenAI has a big leap, other companies are able to step up, big and small companies in different ways. But if you look at something like nuclear weapons, you've spoken about the Manhattan Project, there could be really like a technological and engineering barriers that prevent the guy or gal in her mom's basement to make progress. And it seems like the transition to that kind of world where only one player can develop AGI is possible, so it's not entirely impossible, even though the current state of things seems to be optimistic.

Guillaume Verdon (00:45:26) That's what we're trying to avoid. To me, I think another point of failure is the centralization of the supply chains for the hardware.

Lex Fridman (00:45:34) Right.

Guillaume Verdon (00:45:35) Yeah. Nvidia is just the dominant player, AMD's trailing behind and then we have TSMC is the main fab in Taiwan, which geopolitically sensitive and then we have ASML, which is the maker of the extreme ultraviolet lithography machines. Attacking or monopolizing or co-opting any one point in that chain, you kind of capture the space and so what I'm trying to do is sort of explode the variance of possible ways to do AI and hardware by fundamentally re-imagining how you embed AI algorithms into the physical world. And in general, by the way, I dislike the term AGI, Artificial General Intelligence. I think it's very anthropocentric that we call a human-like or human-level AI, Artificial General Intelligence, right. I've spent my career so far exploring notions of intelligence that no biological brain could achieve for an quantum form of intelligence, right. Grokking systems that have multipartite quantum entanglement that you can provably not represent efficiently on a classical computer or a classical deep learning representation and hence any sort of biological brain.

(00:47:06) And so, already I've spent my career sort of exploring the wider space of intelligences and I think that space of intelligence inspired by physics rather than the human brain is very large. And I think we're going through a moment right now similar to when we went from Geocentrism to Heliocentrism, right. But for intelligence, we realized that human intelligence is just a point in a very large space of potential intelligences. And it's both humbling for humanity, it's a bit scary, right? That we're not at the center of this space, but we made that realization for astronomy and we've survived and we've achieved technologies. By indexing to reality, we've achieved technologies that ensure our wellbeing, for example, we have satellites monitoring solar flares, right, that give us a warning. And so similarly I think by letting go of this anthropomorphic, anthropocentric anchor for AI, we'll be able to explore the wider space of intelligences that can really be a massive benefit to our wellbeing and the advancement of civilization.

Lex Fridman (00:48:32) And still we're able to see the beauty and meaning in the human experience even though we're no longer in our best understanding of the world at the center of it.

Guillaume Verdon (00:48:42) I think there's a lot of beauty in the universe, right. I think life itself, civilization, this Homo Techno, capital mimetic machine that we all live in, right. So you have humans, technology, capital, memes, everything is coupled to one another, everything induces selective pressure on one another. And it's a beautiful machine that has created us, has created the technology we're using to speak today to the audience, capture our speech here, the technology we use to augment ourselves every day, we have our phones. I think the system is beautiful and the principle that induces this sort of adaptability and convergence on optimal technologies, ideas and so on, it's a beautiful principle that we're part of.

(00:49:37) And I think part of EAC is to appreciate this principle in a way that's not just centered on humanity, but kind of broader, appreciate life, the preciousness of consciousness in our universe. And because we cherish this beautiful state of matter we're in, we got to feel a responsibility to scale it in order to preserve it, because the options are to grow or die.

Merging with AI

Lex Fridman (00:50:13) So if it turns out that the beauty that is consciousness in the universe is bigger than just humans, the AI can carry that same flame forward. Does it scare you, are you concerned that AI will replace humans?

Guillaume Verdon (00:50:32) So during my career, I had a moment where I realized that maybe we need to offload to machines to truly understand the universe around us, right, instead of just having humans with pen

and paper solve it all. And to me that sort of process of letting go of a bit of agency gave us way more leverage to understand the world around us. A quantum computer is much better than a human to understand matter at the Nanoscale. Similarly, I think that humanity has a choice, do we accept the opportunity to have intellectual and operational leverage that AI will unlock and thus ensure that we're taken along this path of growth in the scope and scale of civilization? We may dilute ourselves, right? There might be a lot of workers that are AI, but overall out of our own self-interest, by combining and augmenting ourselves with AI, we're going to achieve much higher growth and much more prosperity, right.

(00:51:49) To me, I think that the most likely future is one where humans augment themselves with AI. I think we're already on this path to augmentation, we have phones we use for communication, we have on ourselves at all times. We have wearables, soon that have shared perception with us, right, like the Humane AI Pin or I mean, technically your Tesla car has shared perception. And so if you have shared experience, shared context, you communicate with one another and you have some sort of IO, really it's an extension of yourself.And to me, I think that humanity augmenting itself with AI and having AI that is not anchored to anything biological, both will coexist. And the way to align the parties, we already have a sort of mechanism to align super intelligences that are made of humans and technology, right? Companies are sort of large mixture of expert models, where we have neural routing of tasks within a company and we have ways of economic exchange to align these behemoths.

(00:53:10) And to me, I think capitalism is the way, and I do think that whatever configuration of matter or information leads to maximal growth, will be where we converge, just from like physical principles. And so we can either align ourselves to that reality and join the acceleration up in scope and scale of civilization or we can get left behind and try to decelerate and move back in the forest, let go of technology and return to our primitive state. And those are the two paths forward, at least to me.

Lex Fridman (00:53:54) But there's a philosophical question whether there's a limit to the human capacity to align. So let me bring it up as a form of argument, this guy named Dan Hendrycks and he wrote that he agrees with you that AI development could be viewed as an evolutionary process, but to him, to Dan, this is not a good thing, as he argues that natural selection favors AIs over humans and this could lead to human extinction. What do you think, if it is an evolutionary process and AI systems may have no need for humans?

Guillaume Verdon (00:54:36) I do think that we're actually inducing an evolutionary process on the space of AIs through the market, right. Right now we run AIs that have positive utility to humans and that induces a selective pressure, if you consider a neural net being alive when there's an API running instances of it on GPUs.

Lex Fridman (00:55:01) Yeah.

Guillaume Verdon (00:55:01) Right. And which APIs get run? The ones that have high utility to us, right. So similar to how we domesticated wolves and turned them into dogs that are very clear in their expression, they're very aligned, right. I think there's going to be an opportunity to steer AI and achieve highly aligned AI. And I think that humans plus AI is a very powerful combination and it's not clear to me that pure AI would select out that combination.

Lex Fridman (00:55:40) So the humans are creating the selection pressure right now to create AIs that are aligned to humans, but given how AI develops and how quickly it can grow and scale, to me, one of the concerns is unintended consequences, like humans are not able to anticipate all the consequences of this

process. The scale of damage that could be done through unintended consequences with AI systems is very large.

Guillaume Verdon (00:56:10) The scale of the upside.

Lex Fridman (00:56:12) Yes.

Guillaume Verdon (00:56:13) Right?

Lex Fridman (00:56:13) Guess it's-

Guillaume Verdon (00:56:14) By augmenting ourselves with AI is unimaginable right now. The opportunity cost, we're at a fork in the road, right? Whether we take the path of creating these technologies, augment ourselves and get to climb up the Kardashev Scale, become multi-planetary with the aid of AI, or we have a hard cutoff of like we don't birth these technologies at all and then we leave all the potential upside on the table.

Lex Fridman (00:56:42) Yeah.

Guillaume Verdon (00:56:42) Right. And to me, out of responsibility to the future humans we could carry, with higher carrying capacity by scaling up civilization. Out of responsibility to those humans, I think we have to make the greater grander future happen.

Lex Fridman (00:56:58) Is there a middle ground between cutoff and all systems go? Is there some argument for caution?

Guillaume Verdon (00:57:06) I think, like I said, the market will exhibit caution. Every organism, company, consumer is acting out of self-interest and they won't assign capital to things that have negative utility to them.

Lex Fridman (00:57:21) The problem is with the market is, there's not always perfect information, there's manipulation, there's bad faith actors that mess with the system. It's not always a rational and honest system.

Guillaume Verdon (00:57:41) Well, that's why we need freedom of information, freedom of speech and freedom of thought in order to be able to converge on the subspace of technologies that have positive utility for us all, right.

p(doom)

Lex Fridman (00:57:56) Well let me ask you about p(doom), probability of doom. That's just fun to say, but not fun to experience. What is to you the probability that AI eventually kills all or most humans, also known as probability of doom?

Guillaume Verdon (00:58:16) I'm not a fan of that calculation, I think people just throw numbers out there and it's a very sloppy calculation, right? To calculate a probability, let's say you model the world as some sort of Markov process, if you have enough variables or hidden Markov process. You need to do a stochastic path integral through the space of all possible futures, not just the futures that your brain naturally steers towards, right. I think that the estimators of p(doom) are biased because of our biology, right? We've evolved to have bias sampling towards negative futures that are scary, because that was an evolutionary optimum, right. And so people that are of, let's say higher neuroticism will just think of negative futures where everything goes wrong all day every day and claim that they're doing unbiased sampling. And in a sense they're not normalizing for the space of all possibilities and the space of all possibilities is super exponentially large and it's very hard to have this estimate.

(00:59:40) And in general, I don't think that we can predict the future with that much granularity because of chaos, right? If you have a complex system, you have some uncertainty and a couple of variables, if you let time evolve, you have this concept of a Lyapunov exponent, right. A bit of fuzz becomes a lot of fuzz in our estimate, exponentially so, over time. And I think we need to show some humility that we can't actually predict the future, the only prior we have is the laws of physics, and that's what we're arguing for. The laws of physics say the system will want to grow and subsystems that are optimized for growth and replication are more likely in the future. And so we should aim to maximize our current mutual information with the future and the path towards that is for us to accelerate rather than decelerate.

(01:00:40) So I don't have a p(doom), because I think that similar to the quantum supremacy experiment at Google, I was in the room when they were running the simulations for that. That was an example of a quantum chaotic system where you cannot even estimate probabilities of certain outcomes with even the biggest supercomputer in the world, right. So that's an example of chaos and I think the system is far too chaotic for anybody to have an accurate estimate of the likelihood of certain futures. If they were that good, I think they would be very rich trading on the stock market.

Lex Fridman (01:01:23) But nevertheless, it's true that humans are biased, grounded in our evolutionary biology, scared of everything that can kill us, but we can still imagine different trajectories that can kill us. We don't know all the other ones that don't necessarily, but it's still I think, useful combined with some basic intuition grounded in human history, to reason about like what… Like looking at geopolitics, looking at basics of human nature, how can powerful technology hurt a lot of people? It just seems grounded in that, looking at nuclear weapons, you can start to estimate p(doom) maybe in a more philosophical sense, not a mathematical one. Philosophical meaning like is there a chance? Does human nature tend towards that or not?

Guillaume Verdon (01:02:25) I think to me, one of the biggest existential risks would be the concentration of the power of AI in the hands of the very few, especially if it's a mix between the companies that control the flow of information and the government. Because that could set things up for a sort of dystopian future where only a very few and an oligopoly in the government have AI and they could even convince the public that AI never existed. And that opens up sort of these scenarios for authoritarian centralized control, which to me is the darkest timeline. And the reality is that we have a data-driven prior of these things happening, right. When you give too much power, when you centralize power too much, humans do horrible things, right.

(01:03:23) And to me, that has a much higher likelihood in my Bayesian inference than Sci-Fi based priors, right, like, "My prior came from the Terminator movie." And so when I talked to these AI doomers, I just ask them to trace a path through this Markov chain of events that would lead to our doom and to actually give me a good probability for each transition. And very often there's a unphysical or highly unlikely transition in that chain, right. But of course, we're wired to fear things and we're wired to respond to danger, and we're wired to deem the unknown to be dangerous, because that's a good heuristic for survival, right. But there's much more to lose out of fear. We have so much to lose, so much upside to lose by preemptively stopping the positive futures from happening out of fear. And so I think that we shouldn't give into fear, fear is the mind killer, I think it's also the civilization killer.

Lex Fridman (01:04:43) We can still think about the various ways things go wrong, for example, the founding fathers of the United States thought about human nature and that's why there's a discussion about the freedoms that are necessary. They really deeply deliberated about that and I think the same could possibly

be done for AGI. It is true that human history shows that we tend towards centralization, or at least when we achieve centralization, a lot of bad stuff happens. When there's a dictator, a lot of dark, bad things happen. The question is, can AGI become that dictator? Can AGI when develop, become the centralizer, because of its power? Maybe because of the alignment of humans, perhaps, the same tendencies, the same Stalin like tendencies to centralize and manage centrally the allocation of resources?

(01:05:45) And you can even see that as a compelling argument on the surface level. "Well, AGI is so much smarter, so much more efficient, so much better at allocating resources, why don't we outsource it to the AGI?" And then eventually whatever forces that corrupt the human mind with power could do the same for AGI. It'll just say, "Well, humans are dispensable, we'll get rid of them." Do the Jonathan Swift, Modest Proposal from a few centuries ago, I think the 1700s, when he satirically suggested that, I think it's in Ireland, that the children of poor people are fed as food to the rich people and that would be a good idea, because it decreases the amount of poor people and gives extra income to the poor people. So on several accounts decreases the amount of poor people, therefore more people become rich. Of course, it misses a fundamental piece here that's hard to put into a mathematical equation of the basic value of human life. So all of that to say, are you concerned about AGI being the very centralizer of power that you just talked about?

Guillaume Verdon (01:07:09) I do think that right now there's a bias over a centralization of AI, because of a compute density and centralization of data and how we're training models. I think over time we're going to run out of data to scrape over the internet, and I think that, well, actually I'm working on, increasing the compute density so that compute can be everywhere and acquire information and test hypotheses in the environment in a distributed fashion. I think that fundamentally, centralized cybernetic control, so having one intelligence that is massive that fuses many sensors and is trying to perceive the world accurately, predict it accurately, predict many, many variables and control it, enact its will upon the world, I think that's just never been the optimum, right? Like let's say you have a company, if you have a company, I don't know, of 10,000 people, they all report to the CEO. Even if that CEO is an AI, I think it would struggle to fuse all of the information that is coming to it and then predict the whole system and then to enact its will.

(01:08:28) What has emerged in nature and in corporations and all sorts of systems is a notion of sort of hierarchical cybernetic control, right. In a company it would be, you have like the individual contributors, they are self-interested and they're trying to achieve their tasks and they have a fine, in terms of time and space if you will, control loop and field of perception, right. They have their code base, let's say you're in a software company, they have their code base, they iterate it on it intraday, right. And then the management maybe checks in, it has a wider scope, it has, let's say five reports, right. And then it samples each person's update once per week, and then you can go up the chain and you have larger timescale and greater scope. And that seems to have emerged as sort of the optimal way to control systems.

(01:09:25) And really that's what capitalism gives us, right? You have these hierarchies and you can even have like parent companies and so on. And so that is far more fault tolerant, in quantum computing, that's my feel that came from, we have a concept of this fault tolerance in quantum air correction, right? Quantum air correction is detecting a fault that came from noise, predicting how it's propagated through the system and then correcting it, right, so it's a cybernetic loop. And it turns out that decoders that are hierarchical and in each level, the hierarchy are local-

Guillaume Verdon (01:10:00) … that are hierarchical. And at each level, the hierarchy are local, perform the best by far, and are far more fault-tolerant. The reason is, if you have a non-local decoder, then you have

one fault at this control node and the whole system crashes. Similarly to if you have one CEO that everybody reports to and that CEO goes on vacation, the whole company comes to a crawl. To me, I think that yes, we're seeing a tendency towards centralization of AI, but I think there's going to be a correction over time, where intelligence is going to go closer to the perception. And we're going to break up AI into smaller subsystems that communicate with one another and form a meta system.

Lex Fridman (01:10:56) If you look at the hierarchies that are in the world today, there's nations and those all hierarchical. But in relation to each other, nations are anarchic, so it's an anarchy.

Guillaume Verdon (01:11:06) Mm-hmm.

Lex Fridman (01:11:08) Do you foresee a world like this, where there's not a over… What'd you call it? A centralized cybernetic control?

Guillaume Verdon (01:11:17) Centralized locus of control. Yeah.

Lex Fridman (01:11:21) That's suboptimal, you're saying?

Guillaume Verdon (01:11:22) Yeah.

Lex Fridman (01:11:23) So, it would be always a state of competition at the very top level?

Guillaume Verdon (01:11:27) Yeah. Yeah. Just like in a company, you may have two units working on similar technology and competing with one another, and you prune the one that performs not as well. That's a selection process for a tree, or a product gets killed and then a whole org gets fired. This process of trying new things and shedding old things that didn't work, it's what gives us adaptability and helps us converge on the technologies and things to do that are most good.

Lex Fridman (01:12:04) I just hope there's not a failure mode that's unique to AGI versus humans, because you're describing human systems mostly right now.

Guillaume Verdon (01:12:11) Right.

Lex Fridman (01:12:11) I just hope when there's a monopoly on AGI in one company, that we'll see the same thing we see with humans, which is, another company will spring up and start competing effectively.

Guillaume Verdon (01:12:24) That's been the case so far. We have OpenAI. We have Anthropic. Now, we have xAI. We have Meta even for open source, and now we have Mistral, which is highly competitive. That's the beauty of capitalism. You don't have to trust any one party too much because we're always hedging our bets at every level. There's always competition and that's the most beautiful thing to me, at least, is that the whole system is always shifting and always adapting.

(01:12:54) Maintaining that dynamism is how we avoid tyranny. Making sure that everyone has access to these tools, to these models, and can contribute to the research, avoids a neural tyranny where very few people have control over AI for the world and use it to oppress those around them.

Quantum machine learning

Lex Fridman (01:13:23) When you were talking about intelligence, you mentioned multipartite quantum entanglement.

Guillaume Verdon (01:13:28) Mm-hmm.

Lex Fridman (01:13:29) High-level question first is, what do you think is intelligence? When you think about quantum mechanical systems and you observe some kind of computation happening in them, what do you think is intelligent about the kind of computation the universe is able to do; a small, small inkling of which is the kind of computation a human brain is able to do?

Guillaume Verdon (01:13:52) I would say intelligence and computation aren't quite the same thing. I think that the universe is very much doing a quantum computation. If you had access to all the degrees of freedom and a very, very, very large quantum computer with many, many, many qubits, let's say, a few qubits per Planck volume, which is more or less the pixels we have, then you'd be able to simulate the whole universe on a sufficiently large quantum computer, assuming you're looking at a finite volume, of course, of the universe. I think that at least to me, intelligence is, I go back to cybernetics, the ability to perceive, predict, and control our world.

(01:14:46) But really, nowadays, it seems like a lot of intelligence we use is more about compression. It's about operationalizing information theory. In information theory, you have the notion of entropy of a distribution or a system, and entropy tells you that you need this many bits to encode this distribution or this subsystem, if you have the most optimal code. AI, at least the way we do it today for LLMs and for quantum, is very much trying to minimize relative entropy between our models of the world and the world, distributions from the world. We're learning, we're searching over the space of computations to process the world, to find that compressed representation that has distilled all the variance in noise and entropy.

(01:15:58) Originally, I came to quantum machine learning from the study of black holes because the entropy of black holes is very interesting. In a sense, they're physically the most dense objects in the universe. You can't pack more information spatially any more densely than in a black hole. And so, I was wondering, how do black holes actually encode information? What is their compression code? That got me into the space of algorithms, to search over space of quantum codes. It got me actually into also, how do you acquire quantum information from the world? Something I've worked on, this is public now, is quantum analog digital conversion.

(01:16:50) How do you capture information from the real world in superposition and not destroy the superposition, but digitize for a quantum mechanical computer information from the real world? If you have an ability to capture quantum information and learn representation representations of it, now you can learn compressed representations that may have some useful information in their latent representation. I think that many of the problems facing our civilization are actually beyond this complexity barrier. The greenhouse effect is a quantum mechanical effect. Chemistry is quantum mechanical. Nuclear physics is quantum mechanical.

(01:17:43) A lot of biology and protein folding and so on is affected by quantum mechanics. And so, unlocking an ability to augment human intellect with quantum mechanical computers and quantum mechanical AI seemed to me like a fundamental capability for civilization that we needed to develop. I spent several years doing that, but over time, I grew weary of the timelines that were starting to look like nuclear fusion.

Lex Fridman (01:18:17) One high-level question I can ask is maybe by way of definition, by way of explanation, what is a quantum computer and what is quantum machine learning?

Guillaume Verdon (01:18:27) A quantum computer really is a quantum mechanical system, over which we have sufficient control, and it can maintain its quantum mechanical state. And quantum mechanics is how nature behaves at the very small scales, when things are very small or very cold, and it's actually more fundamental than probability theory. We're used to things being this or that, but we're not used to thinking in superpositions because, well, our brains can't do that. So, we have to translate the quantum mechanical world to, say, linear algebra to grok it. Unfortunately, that translation is exponentially inefficient on average.

You have to represent things with very large matrices. But really, you can make a quantum computer out of many things, and we've seen all sorts of players, from neutral atoms, trapped ions, superconducting metal photons at different frequencies.

(01:19:38) I think you could make a quantum computer out of many things. But to me, the thing that was really interesting was both quantum machine learning was about understanding the quantum mechanical world with quantum computers, so embedding the physical world into AI representations, and quantum computer engineering was embedding AI algorithms into the physical world. This bi-directionality of embedding physical world into AI, AI into the physical world, this symbiosis between physics and AI, really that's the core of my quest really, even to this day, after quantum computing. It's still in this journey to merge really physics and AI.

Lex Fridman (01:20:29) Quantum machine learning is a way to do machine learning on a representation of nature that stays true to the quantum mechanical aspect of nature?

Guillaume Verdon (01:20:43) Yeah, it's learning quantum mechanical representations. That would be quantum deep learning. Alternatively, you can try to do classical machine learning on a quantum computer. I wouldn't advise it because you may have some speed-ups, but very often, the speed-ups come with huge costs. Using a quantum computer is very expensive.

(01:21:08) Why is that? Because you assume the computer is operating at zero temperature, which no physical system in the universe can achieve that temperature. What you have to do is what I've been mentioning, this quantum error correction process, which is really an algorithmic fridge. It's trying to pump entropy out of the system, trying to get it closer to zero temperature. When you do the calculations of how many resources it would take to, say, do deep learning on a quantum computer, classical deep learning, there's such a huge overhead, it's not worth it. It's like thinking about shipping something across a city using a rocket and going to orbit and back. It doesn't make sense. Just use a delivery truck.

Lex Fridman (01:21:53) What kind of stuff can you figure out, can you predict, can you understand with quantum deep learning that you can't with deep learning? So, incorporating quantum mechanical systems into the learning process?

Guillaume Verdon (01:22:05) I think that's a great question. Fundamentally, it's any system that has sufficient quantum mechanical correlations that are very hard to capture for classical representations. Then, there should be an advantage for a quantum mechanical representation over a purely classical one. The question is, which systems have sufficient correlations that are very quantum? But it's also, which systems are still relevant to industry? That's a big question. People are leaning towards chemistry, nuclear physics. I've worked on actually processing inputs from quantum sensors. If you have a network of quantum sensors, they've captured a quantum mechanical image of the world and how to post-process that, that becomes a quantum form of machine perception. For example, Fermilab has a project exploring detecting dark matter with these quantum sensors. To me, that's in alignment with my quest to understand the universe ever since I was a child. And so, someday, I hope that we can have very large networks of quantum sensors that help us peer into the earliest parts of the universe. For example, the LIGO is a quantum sensor. It's just a very large one. So, yeah, I would say quantum machine perception, simulations, grokking quantum simulations, similar to AlphaFold. AlphaFold understood the probability distribution over configurations of proteins. You can understand quantum distributions over configurations of electrons more efficiently with quantum machine learning.

Lex Fridman (01:23:53) You co-authored a paper titled A Universal Training Algorithm for Quantum Deep Learning. That involves Baqprop, with a Q. Very well done, sir. Very well done. How does it work? Is there some interesting aspects you can just mention on how Baqprop and some of these things we know for classical machine learning transfer over to the quantum machine learning?

Guillaume Verdon (01:24:19) Yeah. That was a funky paper. That was one of my first papers in quantum deep learning. Everybody was saying, "Oh, I think deep learning is going to be sped up by quantum computers." I was like, " Well, the best way to predict the future is to invent it. So, here's a 100-page paper, have fun." Essentially, quantum computing is usually, you embed reversible operations into a quantum computation.

(01:24:47) The trick there was to do a feedforward operation and do what we call a phase kick. But really, it's just a force kick. You just kick the system with a certain force that is proportional to your loss function that you wish to optimize. And then, by performing uncomputation, you start with a superposition over parameters, which is pretty funky. Now, you don't have just a point for parameters, you have a superposition over many potential parameters. Our goal is-

Lex Fridman (01:25:24) Is using phase kick somehow to adjust the parameters?

Guillaume Verdon (01:25:28) Right. Because phase kicks emulate having the parameter space be like a particle in end dimensions, and you're trying to get the Schrödinger equation, Schrödinger dynamics, in the lost landscape of the neural network. You do an algorithm to induce this phase kick, which involves a feedforward, a kick. And then, when you uncompute the feedforward, then all the errors in these phase kicks and these forces back- propagate and hit each one of the parameters throughout the layers.

(01:26:04) If you alternate this with an emulation of kinetic energy, then it's like a particle moving in end dimensions, a quantum particle. The advantage in principle would be that it can tunnel through the landscape and find new optima that would've been difficult for stochastic optimizers. But again, this is a theoretical thing, and in practice with at least the current architectures for quantum computers that we have planned, such algorithms would be extremely expensive to run.

Quantum computer

Lex Fridman (01:26:41) Maybe this is a good place to ask the difference between the different fields that you've had a toe in. So, mathematics, physics, engineering, and also entrepreneurship, the different layers of the stack. I think a lot of the stuff you're talking about here is a little bit on the math side, maybe physics almost working in theory.

Guillaume Verdon (01:27:03) Mm-hmm.

Lex Fridman (01:27:03) What's the difference between math, physics, engineering, and making a product for a quantum computing for quantum machine learning?

Guillaume Verdon (01:27:14) Yeah. Some of the original team for the TensorFlow Quantum project, which we started in school, at University of Waterloo, there was myself. Initially, I was a physicist, applied mathematician. We had a computer scientist, we had a mechanical engineer, and then we had a physicist. That was experimental primarily. Putting together teams that are very cross-disciplinary and figuring out how to communicate and share knowledge is really the key to doing this interdisciplinary engineering work.

(01:27:51) There is a big difference. In mathematics, you can explore mathematics for mathematics' sake. In physics, you're applying mathematics to understand the world around us. And in engineering, you're trying to

hack the world. You're trying to find how to apply the physics that I know, my knowledge of the world, to do things.

Lex Fridman (01:28:11) Well, in quantum computing in particular, I think there's just a lot of limits to engineering. It just seems to be extremely hard.

Guillaume Verdon (01:28:17) Yeah.

Lex Fridman (01:28:18) So, there's a lot of value to be exploring quantum computing, quantum machine learning in theory with math. I guess one question is, why is it so hard to build a quantum computer? What's your view of timelines in bringing these ideas to life?

Guillaume Verdon (01:28:43) Right. I think that an overall theme of my company is that we have folks that are… There's a sort of exodus from quantum computing and we're going to broader physics-based AI that is not quantum. So, that gives you a hint.

Lex Fridman (01:29:00) We should say the name of your company is Extropic?

Guillaume Verdon (01:29:03) Extropic, that's right. We do physics-based AI, primarily based on thermodynamics, rather than quantum mechanics. But essentially, a quantum computer is very difficult to build because you have to induce this zero temperature subspace of information. The way to do that is by encoding information, you encode a code within a code, within a code, within a code. There's a lot of redundancy needed to do this error correction, but ultimately, it's a sort of algorithmic refrigerator, really. It's just pumping out entropy out of the subsystem that is virtual and delocalized that represents your "logical qubits", aka the payload quantum bits in which you actually want to run your quantum mechanical program. It's very difficult because in order to scale up your quantum computer, you need each component to be of sufficient quality for it to be worth it. Because if you try to do this error correction, this quantum error correction process, in each quantum bit and your control over them, if it's insufficient, it's not worth scaling up. You're actually adding more errors than you remove. There's this notion of a threshold where if your quantum bits are sufficient quality in terms of your control over them, it's actually worth scaling up. Actually, in recent years, people have been crossing the threshold and it's starting to be worth it.

(01:30:38) It's just a very long slog of engineering, but ultimately, it's really crazy to me how much exquisite level of control we have over these systems. It's actually quite crazy. And people are crossing… They're achieving milestones. It's just in general, the media always gets ahead of where the technology is. There's a bit too much hype. It's good for fundraising, but sometimes it causes winters. It's the hype cycle. I'm bullish on quantum computing on a 10, 15-year timescale personally, but I think there's other quests that can be done in the meantime. I think it's in good hands right now.

Lex Fridman (01:31:22) Well, let me just explore different beautiful ideas, large or small, in quantum computing that might jump out at you from memory when you co-authored a paper titled Asymptotically Limitless Quantum Energy Teleportation via Qudit Probes. Just out of curiosity, can you explain what a qudit is versus a qubit?

Guillaume Verdon (01:31:45) Yeah. It's a D-state qubit.

Lex Fridman (01:31:49) It's a multidimensional?

Guillaume Verdon (01:31:50) Multidimensional, right. It's like, well, can you have a notion of an integer floating point that is quantum mechanical? That's something I've had to think about. I think that research was a precursor to later work on quantum analog digital conversion. There was interesting because during my masters, I was trying to understand the energy and entanglement of the vacuum of emptiness. Emptiness

has energy, which is very weird to say. Our equations of cosmology don't match our calculations for the amount of quantum energy there is in the fluctuations.

(01:32:36) I was trying to hack the energy of the vacuum, and the reality is that you can't just directly hack it. It's not technically free energy. Your lack of knowledge of the fluctuations means you can't extract the energy. But just like the stock market, if you have a stock that's correlated over time, the vacuum's actually correlated. If you measured the vacuum at one point, you acquired information. If you communicated that information to another point, you can infer what configuration the vacuum is in to some precision and statistically extract, on average, some energy there. So, you've "teleported energy".

(01:33:18) To me, that was interesting because you could create pockets of negative-energy density, which is energy density that is below the vacuum, which is very weird because we don't understand how the vacuum gravitates. There are theories where the vacuum or the canvas of space-time itself is really a canvas made out of quantum entanglement. I was studying how decreasing energy of vacuum locally increases quantum entanglement, which is very funky.

(01:33:58) The thing there is that, if you're into to weird theories about UAPs and whatnot, you could try to imagine that they're around. And how would they propel themselves? How would they go faster than the speed of light? You would need a sort of negative energy density. To me, I gave it the old college try, trying to hack the energy of vacuum and hit the limits allowable by the laws of physics. But there's all sorts of caveats there where you can't extract more than you've put in, obviously.

Lex Fridman (01:34:41) But you're saying it's possible to teleport the energy because you can extract information one place and then make, based on that, some kind of prediction about another place?

Guillaume Verdon (01:34:56) Mm-hmm.

Lex Fridman (01:34:57) I'm not sure what to make of that.

Guillaume Verdon (01:34:58) Yeah, it's allowable by the laws of physics. The reality though is that the correlations decay with distance.

Lex Fridman (01:35:06) Sure.

Guillaume Verdon (01:35:06) And so, you're going to have to pay the price not too far away from where you extract it.

Aliens

Lex Fridman (01:35:11) The precision decreases in terms of your ability, but still. But since you mentioned UAPs, we talked about intelligence, and I forgot to ask, what's your view on the other possible intelligences that are out there at the Meso scale? Do you think there's other intelligent alien civilizations? Is that useful to think about? How often do you think about it?

Guillaume Verdon (01:35:36) I think it's useful to think about. It's useful to think about because we got to ensure we're anti-fragile, and we're trying to increase our capabilities as fast as possible. Because we could get disrupted. There's no laws of physics against there being life elsewhere that could evolve and become an advanced civilization and eventually come to us. Do I think they're here now? I'm not sure. I've read what most people have read on the topic.

(01:36:14) I think it's interesting to consider and to me, it's a useful thought experiment to instill a sense of urgency in developing technologies and increasing our capabilities, to make sure we don't get disrupted. Whether it's a form of AI that disrupts us, or a foreign intelligence from a different planet. Either way,

increasing our capabilities and becoming formidable as humans, I think that's really important, so that we're robust against whatever the universe throws at us.

Lex Fridman (01:36:51) But to me, it's also an interesting challenge and thought experiment on how to perceive intelligence. This has to do with quantum mechanical systems. This has to do with any kind of system that's not like humans. To me, the thought experiment is, say, the aliens are here or they are directly observable. We're just too blind, too self-centered, don't have the right sensors, or don't have the right processing of the sensor data to see the obvious intelligence that's all around us.

Guillaume Verdon (01:37:26) Well, that's why we work on quantum sensors. They can sense gravity,

Lex Fridman (01:37:31) Yeah. That's a good one, but there could be other stuff that's not even in the currently known forces of physics.

Guillaume Verdon (01:37:43) Right.

Lex Fridman (01:37:43) There could be some other stuff. The most entertaining thought experiment to me is that it's other stuff that's obvious. It's not like we lack the sensors. It's all around us, the consciousness being one possible one. But there could be stuff that's just obviously there. That once you know it, it's like, "Oh, right. Right. The thing we thought is somehow emergent from the laws of physics, we understand them, is actually a fundamental part of the universe and can be incorporated in physics. Most understood."

Guillaume Verdon (01:38:18) Statistically speaking, if we observed some sort of alien life, it would most likely be some sort of virally, self-replicating, von Neumann-like probe system. And it's possible that there are such systems that, I don't know what they're doing at the bottom of the ocean, allegedly, but maybe they're collecting minerals from the bottom of the ocean.

Lex Fridman (01:38:44) Yeah.

Guillaume Verdon (01:38:45) But that wouldn't violate any of my priors. But am I certain that these systems are here? It'd be difficult for me to say so. I only have secondhand information about there being data.

Lex Fridman (01:38:59) About the bottom of the ocean? Yeah. But could it be things like memes? Could it be thoughts and ideas? Could they be operating at that medium? Could aliens be the very thoughts that come into my head? What's the origin of ideas? In your mind, when an idea comes to your head, show me where it originates.

Guillaume Verdon (01:39:25) Frankly, when I had the idea for the type of computer I'm building now, I think it was eight years ago now, it really felt like it was being beamed from space. I was in bed, just shaking, just thinking it through. I don't know. But do I believe that legitimately? I don't think so. But I think that alien life could take many forms, and I think the notion of intelligence and the notion of life needs to be expanded much more broadly to be less anthropocentric or biocentric.

Quantum gravity

Lex Fridman (01:40:04) Just to linger a little longer on quantum mechanics, through all your explorations on quantum computing, what's the coolest, most beautiful idea that you've come across that has been solved or has not yet been solved?

Guillaume Verdon (01:40:19) I think the journey to understand something called AdS/CFT. So, the journey to understand quantum gravity through this picture, where a hologram of lesser dimension is actually dual or exactly corresponding to a bulk theory of quantum gravity of an extra dimension, and the fact that this sort of duality comes from trying to learn deep learning-like representations of the boundary.

(01:40:59) At least, part of my journey someday on my bucket list is to apply quantum machine learning to these sorts of systems, these CFTs, or they're called SYK models, and learn an emergent geometry from the boundary theory. And so, we can have a form of machine learning to help us understand quantum gravity, which is still a holy grail that I would like to hit before I leave this earth.

Lex Fridman (01:41:35) What do you think is going on with black holes? As information-storing and processing units, what do you think is going on with black holes?

Guillaume Verdon (01:41:46) Black holes are really fascinating objects. They're at the interphase between quantum mechanics and gravity, and so they help us test all sorts of ideas. I think that for many decades now, there's been this black hole information paradox that things that fall into the black hole, we've seem to have lost their information. Now, I think there's this firewall paradox that has been allegedly resolved in recent years by a former peer of mine, who's now a professor at Berkeley. There, it seems like, as information falls into a black hole, there's a sedimentation. As you get closer and closer to the horizon from the point of view, the observer on the outside, the object slows down infinitely as it gets closer and closer.

(01:42:46) Everything that is falling to a black hole, from our perspective, gets sedimented and tacked on to the near horizon. At some point, it gets so close to the horizon, it's in the proximity or the scale in which quantum effects and quantum fluctuations matter. There, that infalling matter could interfere with the traditional pictures, that it could interfere with the creation and annihilation of particles and antiparticles in the vacuum. Through this interference, one of the particles gets entangled with the infalling information and one of them is now free and escapes. That's how there's mutual information between the outgoing radiation and the infalling matter. But getting that calculation right, I think we're only just starting to put the pieces together.

Lex Fridman (01:43:43) There's a few pothead-like questions I want to ask you.

Guillaume Verdon (01:43:46) Sure.

Lex Fridman (01:43:46) One, does it terrify you that there's a giant black hole at the center of our galaxy?

Guillaume Verdon (01:43:52) I don't know. I just want to set up shop near it to fast-forward, meet a future civilization, if we have a limited lifetime, if you could go orbit a black hole and emerge.

Lex Fridman (01:44:08) If there's a special mission that could take you to a black hole, would you volunteer to go travel?

Guillaume Verdon (01:44:13) To orbit and obviously not fall into it.

Lex Fridman (01:44:15) That's obvious. It's obvious to you that everything's destroyed inside a black hole? All the information that makes up Guillaume is destroyed? Maybe on the other side, Beff Jezos emerges and it's just all like it's tied together in some deeply memeful way.

Guillaume Verdon (01:44:32) Yeah, that's a great question. We have to answer what black holes are. Are we punching a hole through space-time and creating a pocket universe? It's possible. Then, that would mean that if we ascend the Kardashev scale to beyond Kardashev Type III, we could engineer black holes with specific hyperparameters to transmit information to new universes we create. And so, we can have progeny that our new…

Guillaume Verdon (01:45:00) … have progeny that are new universes. And so even though our universe may reach a heat death, we may have a way to have a legacy. And so we don't know yet. We need to ascend the Kardashev Scale to answer these questions to peer into that regime of higher energy physics.

Kardashev scale

Lex Fridman (01:45:25) And maybe you can speak to the Kardashev Scale for people who don't know. So one of the sort of meme-like principles and goals of the e/acc movement is to ascend the Kardashev Scale. What is the Kardashev Scale and when do we want to ascend it?

Guillaume Verdon (01:45:43) The Kardashev Scale is a measure of our energy production and consumption. Really, it's a logarithmic scale. Kardashev Type 1 is a milestone where we are producing the equivalent wattage to all the energy that is incident on earth from the sun. Kardashev Type II would be harnessing all the energy that is output by the sun. And I think Type III is like the whole galaxy equivalent-

Lex Fridman (01:46:13) Galaxy, I think [inaudible 01:46:14] yeah.

Guillaume Verdon (01:46:15) Yeah, and then some people have some crazy Type IV and V, but I don't know if I believe in those. But to me, it seems like from the first principles of thermodynamics that, again, there's this concept of thermodynamic- driven dissipative adaptation where life evolved on earth because we have this energetic drive from the sun, we have incident energy, and life evolved on earth to figure out ways to best capture that free energy to maintain itself and grow. And I think that that principle, it's not special to our earth-sun system. We can extend life well beyond. And we kind of have a responsibility to do so because that's the process that brought us here. So we don't even know what it has its store for us in the future. It could be something of beauty we can't even imagine today.

Effective accelerationism (e/acc)

Lex Fridman (01:47:18) So this is probably a good place to talk a bit about the e/acc movement in a Substack blog post titled, What the Fuck is e/acc? Or actually, What the F* is e/acc?, you write, "Strategically speaking, we need to work towards several overarching civilization goals that are all interdependent. And the four goals are, increase the amount of energy we can harness as a species, (climb the Kardashev gradient). In the short term, this almost certainly means nuclear fission. Increase human flourishing via pro-population growth policies and pro-economic growth policies. Create artificial general intelligence, the single greatest force multiplier in human history. And finally, develop interplanetary and interstellar transport so that humanity can spread beyond the earth. Could you build on top of that to maybe say, what to you is the e/acc movement? What are the goals? What are the principles?

Guillaume Verdon (01:48:20) The goal is for the human techno-capital memetic machine to become self-aware and to hyperstitiously engineer its own growth. So let's decompress that.

Lex Fridman (01:48:33) Define each of those words.

Guillaume Verdon (01:48:35) So you have humans, you have technology, you have capital, and then you have memes, information, and all of those systems are coupled with one another. Humans work at companies, they acquire and allocate capital, and humans communicate via memes and information propagation. And our goal was to have a sort of viral optimistic movement that is aware of how the system works, fundamentally it seeks to grow, and we simply want to lean into the natural tendencies of the system to adapt for its own growth.

Lex Fridman (01:49:18) So in that way, you're right, the e/acc is literally a memetic optimism virus that is constantly drifting, mutating, and propagating in a decentralized fashion. So memetic optimism virus. So you do want it to be a virus to maximize the spread, and it's hyperstitious, therefore the optimism will incentivize its growth.

Guillaume Verdon (01:49:43) We see e/acc as sort of a meta-heuristic, sort of very thin cultural framework from which you can have much more opinionated forks. Fundamentally, we just say that what got us here is this adaptation of the whole system based on thermodynamics, and that process is good and we should keep it going. That is the core thesis. Everything else is, okay, how do we ensure that we maintain this malleability and adaptability. Well, clearly not suppressing variants, and maintaining free speech, freedom of thought, freedom of information propagation, and freedom to do AI research is important for us to converge the fastest on the space of technologies, ideas, and whatnot that lead to this growth. And so ultimately, there's been quite a few forks. Some are just memes, but some are more serious. Vitalik Buterin recently made a d/acc fork. He has his own sort of fine-tunings of e/acc.

Lex Fridman (01:50:59) Does anything jump out to memory of the unique characteristic of that fork from Vitalik?

Guillaume Verdon (01:51:05) I would say that it's trying to find a middle ground between e/acc and EA and EI safety. To me, having a movement that is opposite to what was the mainstream narrative that was taking over Silicon Valley was important to shift the dynamic range of opinions. And it's like the balance between centralization and decentralization, the real optimum is always somewhere in the middle. But for e/acc, we're pushing for entropy, novelty, disruption, malleability, speed, rather than being conservative, suppressing thought, suppressing speech, adding constraints, adding too many regulations, slowing things down. And so, we're trying to bring balance to the force.

Lex Fridman (01:52:00) Balance to the force of human civilization.

Guillaume Verdon (01:52:02) It's literally the forces of constraints versus the entropic force that makes us explore. Systems are optimal when they're at the edge of criticality between order and chaos, between constraints, energy minimization and entropy. Systems want to equilibrate, balance these two things. I thought that the balance was lacking, and so we created this movement to bring balance.

Lex Fridman (01:52:31) Well, I like the visual of the landscape of ideas evolving through forks. So on the other part of history, thinking of Marxism as the original repository, and then Soviet Communism is a fork of that, and then the Maoism is a fork of Marxism and Communism. And so those are all forks. They're exploring different ideas.

Guillaume Verdon (01:53:02) Thinking of culture almost like code. Nowadays, what you prompt in the LLM or what you put in the constitution of an LLM is basically its cultural framework, what it believes. And you can share it on GitHub nowadays. So trying to take inspiration from what has worked in this machine of software to adapt over the space of code, could we apply that to culture? And our goal is to not say, "You should live your life this way, X, Y, Z," it's to set up a process where people are always searching over subcultures and competing for mind share. I think creating this malleability of culture is super important for us to converge onto the cultures and the heuristics about how to live one's life that are updated to modern times.

(01:53:59) Because there's really been a sort of vacuum of spirituality and culture. People don't feel like they belong to any one group, and there's been parasitic ideologies that have taken up opportunity to populate this Petri dish of minds. Elon calls it the mind virus. We call it the decel mind virus complex, which is the decelerative that is kind of the overall pattern between all of them. There's many variants as well. And so if there's a sort of viral pessimism, decelerative movement, we needed to have not only one movement, but many, many variants, so it's very hard to pinpoint and stop.

Lex Fridman (01:54:45) But the overarching thing is nevertheless a kind of mimetic optimism pandemic. Okay, let me ask you, do you think e/acc to some degree is a cult?

Guillaume Verdon (01:55:01) Define cult?

Lex Fridman (01:55:03) I think a lot of human progress is made when you have independent thought, so you have individuals that are able to think freely. And very powerful mimetic systems can kind of lead to group think. There's something in human nature that leads to mass hypnosis, mass hysteria. We start to think alike whenever there's a sexy idea that captures our minds. And so it's actually hard to break us apart, pull us apart, diversify a thought. So to that degree, to which degree is everybody kind of chanting "E/acc, e/acc" like the sheep in Animal Farm?

Guillaume Verdon (01:55:46) Well, first of all, it's fun. It's rebellious. There's this concept of meta-irony, of being on the boundary of, "We're not sure if they're serious or not." And it's much more playful and much more fun. For example, we talk about thermodynamics being our god, and sometimes we do cult-like things, but there's no ceremony and robes and whatnot.

Lex Fridman (01:56:19) Not yet.

Guillaume Verdon (01:56:19) Not yet, no. But ultimately, yeah, I totally agree that it seems to me that humans want to feel like they're part of a group, so they naturally try to agree with their neighbors and find common ground. And that leads to sort of mode collapse in the space of ideas. We used to have one cultural island that was allowed. It was a typical subspace of thought, and anything that was diverting from that subspace of thought was suppressed or you were canceled. Now we've created a new mode, but the whole point is that we're not trying to have a very restricted space of thought. There's not just one way to think about e/acc and its many forks. And the point is that there are many forks and there can be many clusters and many islands.

(01:57:07) And I shouldn't be in control of it in any way. I mean, there's no formal org whatsoever. I just put out tweets and certain blog posts, and people are free to defect and fork if there's an aspect they don't like. And so that makes it so that there should be deterritorialization in the space of ideas, so that we don't end up in one cluster that's very cult-like. And so cults usually, they don't allow people to defect or start competing forks, whereas we encourage it.

Humor and memes

Lex Fridman (01:57:51) The pros and cons of humor in meme, in some sense there's like a wisdom to memes. What is it, the Magic Theater? What book is that from? Hermann Hesse. Steppenwolf, I think. But there's a kind of embracing of the absurdity that seems to get to the truth of things, but at the same time, it can also decrease the quality and the rigor of the discourse.

Guillaume Verdon (01:58:22) Yeah.

Lex Fridman (01:58:23) Do you feel the tension of that?

Guillaume Verdon (01:58:25) Yeah. So initially, I think what allowed us to grow under the radar was because it was camouflaged as sort of meta-ironic. We would sneak in deep truths within a package of humor and memes and what are called shit posts, and I think that was purposefully camouflaged against those that seek status and do not want to… It's very hard to argue with a cartoon frog or a cartoon of an intergalactic Jeff Bezos and take yourself seriously, and so that allowed us to grow pretty rapidly in the early days. But of course, essentially people get steered. Their notion of the truth comes from the data they see, from the

information they're fed, and the information people are fed is determined by algorithms. And really what we've been doing is engineering what we call high memetic fitness packets of information, so that they can spread effectively and carry a message.

(01:59:47) So it's kind of a vector to spread the message. And yes, we've been using techniques that are optimal for today's algorithmically-amplified information landscapes. But I think we're reaching the point of scale where we can have serious debates and serious conversations. And that's why we're considering doing a bunch of debates and having more serious long-form discussions. Because I don't think that the timeline is optimal for very serious, thoughtful discussions. You get rewarded for polarization. And so even though we started a movement that is literally trying to polarize the tech ecosystem, at the end of the day so that we can have a conversation and find an optimum together.

Jeff Bezos

Lex Fridman (02:00:42) I mean, that's kind of what I try to do with this podcast given the landscape of things, to still have long-form conversations. But there is a degree to which absurdity is fully embraced. In fact, this very conversation is multi-level absurd. So first of all, I should say that just very recently I had a conversation with Jeff Bezos, and I would love to hear your, Beff Jezos, opinions of Jeff Bezos. Speaking of intergalactic Jeff Bezos. What do you think of that particular individual whom your name has inspired?

Guillaume Verdon (02:01:25) Yeah, I think Jeff is really great. I mean, he's built one of the most epic companies of all time. He's leveraged the techno-capital machine and techno-capital acceleration to give us what we wanted. We want a quick delivery, very convenient, at-home, low prices. He understood how the machine worked and how to harness it, like running the company, not trying to take profits too early, putting it back, letting the system compound and keep improving. And arguably, I think Amazon's invested some of the most amount of capital and robotics out there, and certainly with the birth of AWS, kind of enabled the tech boom we've seen today that has paid the salaries of, I guess myself and all of our friends to some extent. And so I think we can all be grateful to Jeff, and he's one of the great entrepreneurs out there. one of the best of all time, unarguably.

Lex Fridman (02:02:32) And of course, the work at Blue Origin, similar to the work at SpaceX, is trying to make humans a multi-planetary species, which that seems almost like a bigger thing than the capitalist machine. Or it's the capitalist machine at a different timescale perhaps?

Guillaume Verdon (02:02:47) Yeah, I think that companies, they tend to optimize quarter over quarter, maybe a few years out, but individuals that want to leave a legacy can think on a multi-decadal or multi-century timescale. And so the fact that some individuals are such good capital allocators that they unlock the ability to allocate capitals to goals that take us much further or are much further-looking… Elon's doing this with SpaceX, putting all this capital towards getting us to Mars. Jeff is trying to build Blue Origin, and I think he wants to build O'Neill cylinders and get industry off- planet, which I think is brilliant.

(02:03:33) I think just overall, I'm four billionaires. I know this is a controversial statement sometimes, but I think that in a sense it's kind of a proof of stake voting. If you've allocated capital efficiently, you unlock more capital to allocate, just because clearly you know how to allocate capital more efficiently. Which is in contrast to politicians that get elected because they speak the best on TV, not because they have a proven track record of allocating taxpayer capital most efficiently. And so that's why I'm for capitalism over, say, giving all our money to the government and letting them figure out how to allocate it.

Lex Fridman (02:04:20) Why do you think it's a viral and it's a popular meme to criticize billionaires? Since you mentioned billionaires. Why do you think there's quite a widespread criticism of people with wealth, especially those in the public eye, like Jeff and Elon and Mark Zuckerberg, and who else? Bill Gates.

Guillaume Verdon (02:04:44) Yeah, I think a lot of people would, instead of trying to understand how the techno-capital machine works and realizing they have much more agency than they think, they'd rather have this sort of victim mindset. "I'm just subjected to this machine. It is oppressing me. And the successful players clearly must be evil because they've been successful at this game that I'm not successful at." But I've managed to get some people that were in that mindset and make them realize how the techno-capital machine works and how you can harness it for your own good and for the good of others. And by creating value, you capture some of the value you create for the world. That sort of positive sum mindset shift is so potent, and really, that's what we're trying to do by scaling e/acc, is unlocking that higher level of agency. Actually, you're far more in control of the future than you think. You have agency to change the world, go out and do it. Here's permission.

Lex Fridman (02:05:46) Each individual has agency. The motto, "Keep building" is often heard. What does that mean to you, and what does that have to do with Diet Coke? By the way, thank you so much for the Red Bull. It's working pretty well. I'm feeling pretty good.

Guillaume Verdon (02:06:03) Awesome. Well, so building technologies and building… It doesn't have to be technologies, just building in general means having agency, trying to change the world by creating, let's say a company which is a self-sustaining organism that accomplishes a function in the broader techno-capital machine. To us, that's the way to achieve change in the world that you'd like to see, rather than, say, pressuring politicians or creating nonprofits. Nonprofits, once they run out of money, their function can longer be accomplished. You're kind of deforming the market artificially compared to sort of subverting or coursing the market, or dancing with the market, to convince it that actually this function is important, adds value, and here it is. And so I think this is the way between the de-growth, ESG approach, versus, say, Elon. The de-growth approach is like, "We're going to manage our way out of a climate crisis." And Elon is like, "I'm going to build a company that is self-sustaining, profitable, and growing, and we're going to innovate our way out of this dilemma." And we're trying to get people to do the latter rather than the former, at all scales.

Elon Musk

Lex Fridman (02:07:26) Elon is an interesting case. You are a proponent, you celebrate Elon, but he's also somebody who has for a long time warned about the dangers, the potential dangers, existential risks of artificial intelligence. How do you square the two? Is that a contradiction to you?

Guillaume Verdon (02:07:45) It is somewhat because he's very much against regulation in many aspects. But for AI, he's definitely a proponent of regulations. I think overall he saw the dangers of, say, OpenAI cornering the market and then getting to have the monopoly over the cultural priors that you can embed in these LLMs that then, as LLMs now become the source of truth for people, then you can shape the culture of the people. And so you can control people by controlling LLMs. He saw that, just like it was the case for social media, if you shape the function of information propagation, you can shape people's opinions. He sought to make a competitor. So at least, I think we're very aligned there, that the way to a good future is to maintain adversarial equilibria between the various AI players. I'd love to talk to him to understand his thinking about how to advance AI going forwards. I mean, he's also hedging his bets, I would say, with

Neuralink. I think if he can't stop the progress of AI, he's building the technology to merge. Look at the actions, not just the words.

Lex Fridman (02:09:10) Well, there's some degree where being concerned… Maybe using human psychology, being concerned about threats all around us is a motivator. It's an encouraging thing. I operate much better when there's a deadline. The fear of the deadline. And I, for myself, create artificial things, like I want to create in myself this kind of anxiety as if something really horrible will happen if I miss the deadline. I think there's some degree of that here, because creating AI that's aligned with humans has a lot of potential benefits. And so a different way to reframe that is, "If you don't, we're all going to die." It just seems to be a very powerful psychological formulation of the goal of creating human-aligned AI.

Guillaume Verdon (02:09:59) I think that anxiety is good. I think, like I said, I want the free market to create aligned AIs that are reliable, and I think that's what he's trying to do with xAI. So I'm all for it. What I am against is stopping, let's say the OpenSource ecosystem from thriving by, let's say in the executive order, claiming that OpenSource LMs are dual-use technologies and should be government controlled. Then everybody needs to register their GPU and their big matrices with the government. And I think that extra friction will dissuade a lot of hackers from contributing, hackers that could later become the researchers that make key discoveries that push us forward, including discoveries for AI safety. And so I think I just want to maintain ubiquity of opportunity to contribute to AI and to own a piece of the future. It can't just be legislated behind some wall where only a few players get to play the game.

Lex Fridman (02:11:08) The e/acc movement is often caricatured to mean progress and innovation at all costs. Doesn't matter how unsafe it is, doesn't matter if it causes a lot of damage. You just build cool shit as fast as possible, stay up all night with a Diet Coke, whatever it takes. I guess, I don't know if there's a question in there, but how important to you and what you've seen the different formulations of e/acc, is AI safety?

Guillaume Verdon (02:11:44) Again, I think if there was no one working on it, I think I would be a proponent of it. I think, again, our goal is to bring balance, and obviously a sense of urgency is a useful tool to make progress. It hacks our dopaminergic systems and gives us energy to work late into the night. I think also having a higher purpose you're contributing to. At the end of the day, it's like, what am I contributing to? I'm contributing to the growth of this beautiful machine so that we can seek to the stars. That's really inspiring. That's also a sort of neuro hack.

Lex Fridman (02:12:26) So you're saying AI safety is important to you, but right now the landscape of ideas you see is, AI safety as a topic is used more often to gain centralized control. So in that sense, you're resisting it, as a proxy for gaining centralized control?

Guillaume Verdon (02:12:43) Yeah, I just think we have to be careful, because safety is just the perfect cover for centralization of power and covering up eventually corruption. I'm not saying it's corrupted now, but it could be down the line. And really, if you let the argument run, there's no amount of centralization of control that will be enough to ensure your safety. There's always more 999s of P safety that you can gain, 99.9999% safe. Maybe you want another nine. "Oh, please give us full access to everything you do. Full surveillance." And frankly, those that are proponents of AI safety have proposed having a global panopticon where you have centralized perception of everything going on. And to me, that just opens up the door wide open for a big brother, 1984-like scenario. And that's not a future I want to live in.

Lex Fridman (02:13:49) Because we have some examples throughout history when that did not lead to a good outcome.

Extropic

Guillaume Verdon (02:13:54) Right.

Lex Fridman (02:13:56) You mentioned you founded a company, Extropic, that recently announced a 14.1 million seed round. What's the goal of the company? You're talking about a lot of interesting physics things, so what are you up to over there that you can talk about?

Guillaume Verdon (02:14:12) Yeah, originally we weren't going to announce last week, but I think with the doxing and disclosure, we got our hand forced. So we had to disclose roughly what we were doing. But really, Extropic was born from my dissatisfaction, and that of my colleagues, with the quantum computing roadmap. Quantum computing was sort of the first path to physics-based computing that was trying to commercially scale, and I was working on physics-based AI that runs on these physics-based computers. But ultimately, our greatest enemy was this noise, this pervasive problem of noise that, as I mentioned, you have to constantly pump out the noise out of the system to maintain this pristine environment where quantum mechanics can take effect. And that constraint was just too much. It's too costly to do that.

(02:15:11) And so we were wondering, as generative AI is sort of eating the world, more and more of the world's computational workloads are focused on generative AI, how could we use physics to engineer the ultimate physical substrate for generative AI from first principles of physics, of information theory, of computation, and ultimately of thermodynamics? And so what we're seeking to build is a physics-based computing system and physics-based AI algorithms that are inspired by out-of-equilibrium thermodynamics, or harness it directly to do machine learning as a physical process.

Lex Fridman (02:16:01) So what does that mean, machine learning as a physical process? Is that hardware? Is it software? Is it both? Is it trying to do the full stack in some kind of unique way?

Guillaume Verdon (02:16:10) Yes, it is full stack. And so we're folks that have built differentiable programming into the quantum computing ecosystem with TensorFlow Quantum. One of my co-founders of TensorFlow Quantum is the CTO, Trevor McCourt. We have some of the best quantum computer architects, those that have designed IBM's and AWS's systems. They've left quantum computing to help us build what we call actually a thermodynamic computer.

Lex Fridman (02:16:43) A thermodynamic computer. Well, actually let's linger around TensorFlow Quantum. What lessons have you learned from TensorFlow Quantum? Maybe you can speak to what it takes to create essentially, what, like a software API to a quantum computer?

Guillaume Verdon (02:17:01) Right. That was a challenge to invent, to build, and then to get to run on the real devices.

Lex Fridman (02:17:09) Can you actually speak to what it is?

Guillaume Verdon (02:17:11) Yeah. TensorFlow Quantum was an attempt at… Well, I guess we succeeded, at combining deep learning or differentiable classical programming with quantum computing, and turn quantum computing into or have types of programs that are differentiable in quantum computing. And Andrej Karpathy calls differentiable programming, Software 2.0. It's like, gradient descent is a better programmer than you. And the idea was that in the early days of quantum computing, you can only run short quantum programs. And so, which quantum programs should you run? Well, just let gradient descent find those

programs instead. And so we built the first infrastructure to not only run differentiable quantum programs, but combine them as part of broader deep learning graphs, incorporating deep neural networks, the ones you know and love, with what are called quantum neural networks.

(02:18:21) And ultimately, it was a very cross-disciplinary effort. We had to invent all sorts of ways to differentiate, to back propagate through the hybrid graph. But ultimately, it taught me that the way to program matter and to program physics is by differentiating through control parameters. If you have parameters that affects the physics of the system and you can evaluate some loss function, you can optimize the system to accomplish a task, whatever that task may be. And that's a very universal meta framework for how to program physics-based computers.

Lex Fridman (02:19:05) So try to parameterize everything, make those parameters differentiable, and then optimize?

Guillaume Verdon (02:19:12) Yes.

Lex Fridman (02:19:13) Okay. Is there some more practical engineering lessons from TensorFlow Quantum? Just organizationally too, like the humans involved and how to get to a product, how to create good documentation? I don't know. All of these little subtle things that people might not think about.

Guillaume Verdon (02:19:34) I think working across disciplinary boundaries is always a challenge, and you have to be extremely patient in teaching one another. I learned a lot of software engineering through the process. My colleagues learned a lot of quantum physics, and some learned machine learning through the process of building this system. And I think if you get some smart people that are passionate and trust each other in a room, and you have a small team-

Guillaume Verdon (02:20:00) Are passionate and trust each other in a room, and you have a small team, and you teach each other your specialties, suddenly you're kind of forming this sort of model soup of expertise, and something special comes out of that, right? It's like combining genes, but for your knowledge bases, and sometimes special products come out of that. And so I think, even though it's very high friction initially to work in an interdisciplinary team, I think the product at the end of the day is worth it. And so, learned a lot trying to bridge the gap there. And I mean, it's still a challenge to this day. We hire folks that have an AI background, folks that have a pure physics background, and somehow we have to make them talk to one another. Right?

Lex Fridman (02:20:47) Is there a magic, is there some science and art to the hiring process, to building a team that can create magic together?

Guillaume Verdon (02:20:56) Yeah, it's really hard to pinpoint that je ne sais quoi, right?

Lex Fridman (02:21:03) I didn't know you speak French. That's very nice.

Guillaume Verdon (02:21:07) Yeah, I'm actually French Canadian.

Lex Fridman (02:21:09) Oh, you are a legitimately French Canadian.

Guillaume Verdon (02:21:09) I am.

Lex Fridman (02:21:11) I thought you were just doing that for the cred.

Guillaume Verdon (02:21:15) No, no. I'm truly French Canadian, from Montreal. But yeah, essentially we look for people with very high fluid intelligence that aren't overspecialized, because they're going to have to get out of their comfort zone. They're going to have to incorporate concepts that they've never seen before, and very quickly get comfortable with them, or learn to work in a team. And so that's sort of what we look for when we hire. We can't hire people that are just optimizing this subsystem for the past three or four years.

We need really general sort of broader intelligence and specialty, and people that are open-minded, really, because if you're pioneering a new approach from scratch, there is no textbook, there's no reference. It's just us, and people that are hungry to learn. So, we have to teach each other, we have to learn the literature, we have to share knowledge bases, collaborate in order to push the boundary of knowledge further together. And so, people that are used to just getting prescribed what to do at this stage, when you're at the pioneering stage, that's not necessarily who you want to hire. Yeah.

Singularity and AGI

Lex Fridman (02:22:31) So you mentioned with Extropic you're trying to build the physical substrate for generative AI. What's the difference between that and the AGI AI itself? So, is it possible that in the halls of your company, AGI will be created? Or will AGI just be using this as a substrate?

Guillaume Verdon (02:22:51) I think our goal is to both run human like AI, or anthropomorphic AI.

Lex Fridman (02:22:58) Sorry for use of the term AGI. I know it's triggering for you.

Guillaume Verdon (02:23:02) We think that the future is actually physics-based AI combined with anthropomorphic AI. So, you can imagine, I have a sort of world modeling engine through physics-based AI. Physics-based AI is better at representing the world at all scales, because it can be quantum mechanical, thermodynamic, deterministic, hybrid representations of the world, just like our world at different scales has different regimes of physics. If you inspire yourself from that in the ways you learn representations of nature, you can have much more accurate representations of nature. So, you can have very accurate world models at all scales. And so, you have the world modeling engine, and then you have the anthropomorphic AI that is human-like. So you can have the science, the playground to test your ideas, and you can have the synthetic scientist. And to us, that joint system of a physics-based and an anthropomorphic AI is the closest thing to a fully general, artificially intelligent system.

Lex Fridman (02:24:07) So you can get closer to truth by grounding of the AI to physics, but you can also still have a anthropomorphic interface to us humans that like to talk to other humans, or human-like systems. So, on that topic, I suppose that is one of the big limitations of current large language models to you, is that they're good bullshitters, they're not really grounded to truth necessarily. Would that be fair to say?

Guillaume Verdon (02:24:40) Yeah, no, you wouldn't try to extrapolate the stock market with an LM trained on text from the internet. It's not going to be a very accurate model. It's not going to model its priors or its uncertainties about the world very accurately. So, you need a different type of AI to compliment this text extrapolation AI. Yeah.

Lex Fridman (02:25:05) You mentioned singularity earlier. How far away are we from a singularity?

Guillaume Verdon (02:25:09) I don't know if I believe in a finite time singularity as a single point in time. I think it's going to be asymptotic, and sort of a diagonal sort of asymptote. We have the light cone, we have the limits of physics restricting our ability to grow. So, obviously can't fully diverge on a finite time. I think my priors are that I think a lot of people on the other side of the aisle think that once we reach human level AI, there's going to be an inflection point, and a sudden [inaudible 02:25:48], suddenly AI is going to grok how to manipulate matter at the nano scale, and assemble nanobots. And having worked for nearly a decade in applying AI to engineer matter, it's much harder than they think. And in reality, you need a lot of samples from either a simulation of nature that's very accurate and costly, or nature itself, and that keeps your ability to control the world around us in check. There's a sort of minimal cost computationally, and

thermodynamically, to acquiring information about the world in order to be able to predict and control it. And that keeps things in check.

AI doomers

Lex Fridman (02:26:27) It's funny you mentioned the other side of the aisle. So, in the poll I posted about p(doom) yesterday, what's the probability of doom? There seems to be a nice division between people think it's very likely, and very unlikely. I wonder if in the future there'll be the actual Republicans versus Democrats division, blue versus red? Is the AI doomers versus the e/accers, EAC? [inaudible 02:26:53].

Guillaume Verdon (02:26:53) Yeah. So, this movement is not right wing or left wing fundamentally, it's more like up versus down, in terms of the scale of-

Lex Fridman (02:27:01) Which one is the up? Okay.

Guillaume Verdon (02:27:02) … Civilization, right?

Lex Fridman (02:27:03) All right.

Guillaume Verdon (02:27:05) But, it seems to be like there is sort of case of alignment of the existing political parties, where those that are for more centralization of power, control, and more regulations are aligning themselves with the doomers, because that sort of instilling fear in people is a great way for them to give up more control, and give the government more power. But fundamentally, we're not left versus right. I think we've done polls of people's alignment within EAC. I think it's pretty balanced. So, it's a new fundamental issue of our time. It's not just centralization versus decentralization. It's kind of do we go… It's like tech progressivism, versus techno conservatism. Right?

Effective altruism

Lex Fridman (02:27:54) So e/acc as a movement is often formulated in contrast to EA, effective altruism. What do you think are the pros and cons of effective altruism? What's interesting, insightful to you about them, and what is negative?

Guillaume Verdon (02:28:15) Right. I think people trying to do good from first principles is good.

Lex Fridman (02:28:23) We should actually say, and sorry to interrupt, we should probably say that, and you can correct me if I'm wrong, but effective altruism is the kind of movement that's trying to do good optimally, where good is probably measured something like the amount of suffering in the world. You want to minimize it. And there's ways that that can go wrong, as any optimization can. And so, it's interesting to explore how things can go wrong.

Guillaume Verdon (02:28:55) We're both trying to do good to some extent, and we're arguing for which loss function we should use, right?

Lex Fridman (02:29:03) Yes.

Guillaume Verdon (02:29:04) Their loss function is sort of hedons, units of hedonism. How good do you feel, and for how much time? And so, suffering would be negative hedons, and they're trying to minimize that. But to us that seems like that loss function has sort of spurious minima, you can start minimizing shrimp farm pain, which seems not that productive to me. Or you can end up with wire heading, where you just either install a neural link, or you scroll TikTok forever, and you feel good on the short-term timescale because of

your neurochemistry, but on a long-term timescale, it causes decay and death, because you're not being productive.

(02:29:54) Whereas sort of EAC, measuring progress of civilization, not in terms of a subjective loss function like hedonism, but rather an objective measure, quantity that cannot be gamed that is physical energy, it's very objective, and there's not many ways to game it. If you did it in terms of GDP, or a currency, that's pinned to certain value that's moving. And so, that's not a good way to measure our progress. But the thing is we're both trying to make progress, and ensure humanity flourishes, and gets to grow. We just have different loss functions, and different ways of going about doing it.

Lex Fridman (02:30:42) Is there a degree, maybe you can educate me, correct me, I get a little bit skeptical when there's an equation involved trying to reduce all of the human civilization, human experience to an equation. Is there a degree that we should be skeptical of the tyranny of an equation of a loss function over wish to optimize? Like having a kind of intellectual humility about optimizing over loss functions?

Guillaume Verdon (02:31:12) Yeah. So, this particular loss function, it's not stiff. It's kind of an average of averages. It's like distributions of states in the future are going to follow a certain distribution. So it's not deterministic, it's not like… We're not on stiff rails. It's just a statistical statement about the future. But at the end of the day, you can believe in gravity or not, but it's not necessarily an option to obey it. And some people try to test that, and that goes not so well. So, similarly, I think thermodynamics is there whether we like it or not, and we're just trying to point out what is, and try to orient ourselves, and chart a path forward given this fundamental truth.

Lex Fridman (02:32:04) But there's still some uncertainty, there's still a lack of information, and humans tend to fill the gap of the lack of information with narratives. And so, how they interpret… Even physics is up to interpretation when there's uncertainty involved. And humans tend to use that to further their own means. So, it's always, whenever there's an equation, it just seems like until we have really perfect understanding of the universe, humans will do what humans do, and they try to use the narrative of doing good to fool the populace into doing bad. I guess that this is something that we should be skeptical about in all movements.

Guillaume Verdon (02:32:57) That's right? So we invite skepticism. Right?

Lex Fridman (02:33:02) Do you have an understanding of what might, to a degree that went wrong, what do you think may have gone wrong with effective altruism that might also go wrong with effective accelerationism?

Guillaume Verdon (02:33:15) Yeah, I mean I think it provided initially a sense of community for engineers, and intellectuals, and rationalists in the early days, and it seems like the community was very healthy, but then they formed all sorts of organizations, and started routing capital, and having actual power. They have real power. They influence the government, they influence most AI orgs now. I mean, they're literally controlling the board of OpenAI, and look over to Anthropic. I think they'll have some control over that too. And so, I think the assumption of e/acc is more like capitalism, is that every agent organism and meta organism is going to act in its own interest, and we should maintain sort of adversarial equilibrium, or adversarial competition to keep each other in check at all times, at all scales. I think that yeah, ultimately, it was the perfect cover to acquire tons of power, and capital, and unfortunately sometimes that corrupts people over time.

Day in the life

Lex Fridman (02:34:23) What does a perfectly productive day, since building is important, what is a perfectly productive day in the life of Guillaume Verdon look like? How much caffeine do you consume? What's a perfect day?

Guillaume Verdon (02:34:39) Okay, so I have a particular regimen. I would say my favorite days are 12:00 PM to 4:00 AM, and I would have meetings in the early afternoon, usually external meetings, some internal meetings. Because I'm CEO, I have to interface with the outside world, whether it's customers, or investors, or interviewing potential candidates. And usually I'll have ketones, exogenous ketones.

Lex Fridman (02:35:12) So, are you on a keto diet, or is this-

Guillaume Verdon (02:35:16) I've done keto before for football, and whatnot, but I like to have a meal after part of my day is done, and so I can just have extreme focus.

Lex Fridman (02:35:31) You do the social interactions earlier in the day without food.

Guillaume Verdon (02:35:35) Front load them, yeah. Yeah. Like right now I'm on ketones, and a Red Bull, and it just gives you a clarity of thought that is really next level. Because then when you eat, you're actually allocating some of your energy that could be going to neural energy to your digestion. After I eat, maybe I take a break, an hour or so, an hour and a half, and then usually it's like ideally one meal a day, like steak and eggs, and vegetables, animal-based primarily. So, fruit and meat. And then I do a second wind, usually that's deep work, because I am A CEO, but I'm still technical. I'm contributing to most patents. And there, I'll just stay up late into the night, and work with engineers on very technical problems.

Lex Fridman (02:36:25) So it's like the 9:00 PM to 4:00 AM, whatever though, that range of time.

Guillaume Verdon (02:36:30) Yeah, yeah. That's the perfect time. The emails, the things that are on fire stop trickling in, you can focus. And then you have your second wind. And I think Demis Hassabis has a similar workday to some extent. So, I think that's definitely inspired my workday. But yeah, I started this workday when I was at Google, and had to manage a bit of the product during the day, and have meetings, and then do technical work at night.

Lex Fridman (02:37:00) Exercise, sleep, those kinds of things. You said football, you used to play football?

Guillaume Verdon (02:37:06) Yeah, I used to play American football. I've done all sorts of sports growing up. And then I was into powerlifting for a while. So, when I was studying mathematics in grad school, I would just do math, and lift, take caffeine, and that was my day. It was very pure, the purest of monk modes. But it's really interesting, how in powerlifting you're trying to cause neural adaptation by having certain driving signals, and you're trying to engineer a neuroplasticity through all sorts of supplements, and you have all sorts of brain derived neurotrophic factors that get secreted when you lift.

(02:37:44) So, it's funny to me how I was trying to engineer a neural adaptation in my nervous system more broadly, not just my brain while learning mathematics. I think you can learn much faster if you really care. If you convince yourself to care a lot about what you're learning, and you have some sort of assistance, let's say caffeine, or some cholinergic supplement to increase neuroplasticity. I should chat with Andrew Huberman at some point. He's the expert. But yeah, at least to me it's like you can try to input more tokens into your brain, if you will, and you can try to increase the learning rate, so that you can learn much faster on a shorter timescale.

(02:38:30) So, I've learned a lot of things. I've followed my curiosity. You're naturally… If you're passionate about what you're doing, you're going to learn faster, you're going to become smarter faster. And if you follow your curiosity, you're always going to be interested. And so, I advise people to follow their curiosity

and don't respect the boundaries of certain fields, or what you've been allocated in terms of lane of what you're working on. Just go out and explore, and follow your nose, and try to acquire, and compress as much information as you can into your brain. Anything that you find interesting.

Lex Fridman (02:39:05) And caring about a thing. Like you said, which is interesting, it works for me really well, is tricking yourself that you care about a thing.

Guillaume Verdon (02:39:12) Yes.

Lex Fridman (02:39:13) And then you start to really care about it.

Guillaume Verdon (02:39:15) Yep.

Lex Fridman (02:39:15) So, it's funny, the motivation is a really good catalyst for learning.

Guillaume Verdon (02:39:22) Right. And so, at least part of my character, as Beff Jezos is kind of like…

Lex Fridman (02:39:29) Yeah, hype man.

Guillaume Verdon (02:39:30) Yeah, but I'm hyping myself up, but then I just tweet about it, and it's just when I'm trying to get really hyped up, and an altered state of consciousness where I'm ultra focused, in the flow, wired, trying to invent something that's never existed, I need to get to unreal levels of excitement. But your brain has these levels of cognition that you can unlock with higher levels of adrenaline, and whatnot. And I mean, I've learned that in powerlifting, that actually you can engineer a mental switch to increase your strength. If you can engineer a switch, maybe you have a prompt, like a certain song or some music where suddenly you're fully primed, then you're at max, maximum strength. And I've engineered that switch through years of lifting. If you're going to get under 500 pounds and it could crush you, if you don't have that switch to be wired in, you might die. So, that'll wake you right up. That sort of skill I've carried over to research, when it's go time, when the stakes are high, somehow I just reach another level of neural performance.

Lex Fridman (02:40:40) So Beff Jezos is your sort of embodiment representation of your intellectual Hulk. It's your productivity Hulk that you just turn on.

Identity

Guillaume Verdon (02:40:50) Yeah.

Lex Fridman (02:40:50) What have you learned about the nature of identity from having these two identities? I think it's interesting for people, to be able to put on those two hats so explicitly.

Guillaume Verdon (02:41:01) I think it was interesting in the early days, I think in the early days, I thought it was truly compartmentalized. Like, "Oh yeah, this is a character. I'm Guillaume. Beff is just the character." I take my thoughts, and then I extrapolate them to a bit more extreme. But over time, it's kind of like both identities were starting to merge mentally, and people were like, "No, I met you. You are Beff. You are not just Guillaume." And I was like, "Wait, am I?" And now it's fully merged. But it was already, before the docs, it was already starting mentally that I am this character. It's part of me.

Lex Fridman (02:41:39) Would you recommend people have an alt?

Guillaume Verdon (02:41:42) Absolutely.

Lex Fridman (02:41:43) Like young people. Would you recommend them to explore different identities by having alts? Alt accounts?

Guillaume Verdon (02:41:49) It's fun. It's like writing an essay, and taking a position, right? It's like you do this in debate. It's like you can have experimental thoughts, and by the stakes being so low, because you're an anon account with, I don't know, 20 followers or something, you can experiment with your thoughts in a

low stakes environment. And I feel like we've lost that in the era of everything being under your main name, everything being attributable to you. People just are afraid to speak, explore ideas that aren't fully formed, and I feel like we've lost something there. So, I hope platforms like X and others really help support people trying to stay synonymous, or anonymous, because it's really important for people to share thoughts that aren't fully formed, and converge onto maybe hidden truths that were hard to converge upon if it was just through open conversation with real names.

Lex Fridman (02:42:46) Yeah. I really believe in not radical, but rigorous empathy. It's like really considering what it's like to be a person of a certain viewpoint, and taking that, as a thought experiment, farther and farther and farther. And one way of doing that as an alt account. That's a fun, interesting way to really explore what it's like to be a person that believes a set of beliefs, and taking that across the span of several days, weeks, months. Of course there's always the danger of becoming that. That's the Nietzche, "Gaze long into the abyss, the abyss gazes into you." You have to be careful.

Guillaume Verdon (02:42:46) Breaking Beff.

Advice for young people

Lex Fridman (02:43:31) Yeah, right. Breaking Beff. Yeah. You wake up with a shaved head one day, just like, "Who am I? What have I become?" So, you've mentioned quite a bit of advice already, but what advice would you give to young people of, in this interesting world we're in, how to have a career and how to have a life they can be proud of?

Guillaume Verdon (02:43:58) I think to me, the reason I went to theoretical physics was that I had to learn the base of the stack that was going to stick around no matter how the technology changes. And to me, that was the foundation upon which then I later built engineering skills, and other skills. And to me, the laws of physics, it may seem like the landscape right now is changing so fast, it's disorienting. But certain things like fundamental mathematics and physics aren't going to change. And if you have that knowledge, and knowledge about complex systems, and adaptive systems, I think that's going to carry you very far. And so, not everybody has to study mathematics, but I think it's really a huge cognitive unlock to learn math, and some physics, and engineering.

Lex Fridman (02:44:48) Get as close to the base of the stack as possible.

Guillaume Verdon (02:44:51) Yeah, that's right. Because the base of the stack doesn't change. Everything else… Your knowledge might become not as relevant in a few years. Of course there's a sort of transfer learning you can do, but then you have to always transfer learn, constantly.

Lex Fridman (02:45:04) I guess the closer you are to the base of the stack, the easier the transfer learning, the shorter the jump.

Guillaume Verdon (02:45:10) Right, right. And you'd be surprised, once you've learned concepts in many physical scenarios, how they can carry over to understanding other systems that aren't necessarily physics. And I guess the e/acc writings, the principles and tenet posts, that was based on physics, that was kind of my experimentation with applying some of the thinking from out of [inaudible 02:45:36] thermodynamics to understanding the world around us, and it's led to e/acc, and this movement.

Mortality

Lex Fridman (02:45:42) If you look at you're one cog in the machine, in the capitalist machine, one human, and if you look at yourself, do you think mortality is a feature or a bug? Would you want to be immortal?

Guillaume Verdon (02:45:57) No, I think fundamentally, in thermodynamic dissipative adaptation, there's the word dissipation. Dissipation is important, death is important. We have a saying in physics, physics progresses one funeral at a time.

Lex Fridman (02:46:16) Yeah.

Guillaume Verdon (02:46:17) I think the same is true for capitalism. Companies, empires, people, everything. Everything must die at some point. I think that we should probably extend our lifespan, because we need a longer period of training, because the world is more and more complex. We have more and more data to really be able to predict and understand the world. And if we have a finite window of higher neuroplasticity, then we have sort of a hard cap in how much we can understand about our world. So, I think I am for death, because again, I think it's important. If you have a king that would never die, that would be a problem. The system wouldn't be constantly adapting, right?

(02:47:05) You need novelty, you need youth, you need disruption to make sure the system's always adapting, and malleable. Otherwise, if things are immortal, if you have, let's say corporations that are there forever, and they have the monopoly, they get calcified, they become not as optimal, not as high fitness in a changing, time varying landscape. And so, death gives space for youth and novelty to take its place. And I think it's an important part of every system in nature. So yeah, I am for death, but I do think that longer lifespan, and longer time for neuroplasticity, bigger brains should be something we should strive for.

Lex Fridman (02:47:52) Well, and that, Jeff Bezos, and Beff Jezos agree that all companies die. And for Jeff, the goal is to try to, he calls it day one thinking, try to constantly, for as long as possible, reinvent, sort of extend the life of the company. But eventually it too will die, because it's so difficult to keep reinventing. Are you afraid of your own death?

Guillaume Verdon (02:48:23) I think I have ideas and things I'd like to achieve in this world before I have to go, but I don't think I'm necessarily afraid of death.

Lex Fridman (02:48:34) So you're not attached to this particular body, and mind that you got?

Guillaume Verdon (02:48:38) No, I'm sure there's going to be better versions of myself in the future, or…

Lex Fridman (02:48:46) Forks?

Guillaume Verdon (02:48:47) Forks, right? Genetic forks, or other, right? I truly believe that. I think there's a sort of evolutionary-like algorithm happening at every bit, or [inaudible 02:49:03] in the world is sort of adapting through this process that we described in e/acc. And I think maintaining this adaptation malleability is how we have constant optimization of the whole machine. And so, I don't think I'm particularly an optimum that needs to stick around forever. I think there's going to be greater optima in many ways.

Meaning of life

Lex Fridman (02:49:25) What do you think is the meaning of it all? What's the why of the machine? The e/acc machine?

Guillaume Verdon (02:49:32) The why? Well, the why is thermodynamics. It's why we're here. It's what has led to the formation of life, and of civilization, of evolution of technologies, and growth of civilization. But why do we have thermodynamics? Why do we have our particular universe? Why do we have these particular hyper-parameters, the constants of nature? Well then you get into the anthropic principle, and the landscape

of potential universes, right? We're in the universe that allows for life. And then why, is there potentially many universes? I don't know. I don't know that part. But could we potentially engineer new universes, or create pocket universes, and set the hyper-parameters so there is some mutual information between our existence in that universe, and we'd be somewhat its parents? I think that's really… I don't know, that'd be very poetic. It's purely conjecture. But again, this is why figuring out quantum gravity would allow us to understand if we can do that.

Lex Fridman (02:50:39) And above that, why does it all seems so beautiful and exciting? The quest to figuring out quantum gravity seems so exciting. Why? Why is that? Why are we drawn to that? Why are we pulled towards that? Just that puzzle solving creative force that underpins all of it, it seems like.

Guillaume Verdon (02:51:01) I think we seek, just like an LLM seats to minimize cross entropy between its internal model and the world, we seek to minimize… Yeah, the statistical divergence between our predictions and the world, and the world itself. And having regimes of energy scales, or physical scales in which we have no visibility, no ability to predict, or perceive, that's kind of an insult to us. And we want to be able to understand the world better in order to best steer it, or steer us through it.

(02:51:37) And in general, it's a capability that has evolved because the better you can predict the world, the better you can capture utility, or free energy towards your own sustenance and growth. And I think quantum gravity, again, is kind of the final boss, in terms of knowledge acquisition, because once we've mastered that, then we can do a lot, potentially. But between here and there, I think there's a lot to learn in the meso scales. There's a lot of information to acquire about our world, and a lot of engineering perception, prediction, and control to be done, to climb up the Carta shift scale. And to us, that's the great challenge of our times.

Lex Fridman (02:52:22) And when you're not sure where to go, let the meme pave the way.

Guillaume Verdon (02:52:26) That's right.

Lex Fridman (02:52:27) Guillaume, Beff, thank you for talking today. Thank you for the work you're doing. Thank you for the humor, and the wisdom you put into the world. This was awesome.

Guillaume Verdon (02:52:37) Thank you so much for having me, Lex, It's a pleasure.

Lex Fridman (02:52:40) Thank you for listening to this conversation with Guillaume Verdon. To support this podcast. Please check out our sponsors in the description. And now, let me leave you with some words from Albert Einstein. "If at first the idea is not absurd, then there is no hope for it." Thank you for listening. I hope to see you next time.