

[Transcript] FYI - For Your Innovation / Why Bitcoin is the Currency for AI

Welcome to FYI, the four-year innovation podcast. This show offers an intellectual discussion on technologically enabled disruption because investing in innovation starts with understanding it. To learn more, visit arc-invest.com.

Arc Invest is a registered investment advisor focused on investing in disruptive innovation. This podcast is for informational purposes only and should not be relied upon as a basis for investment decisions. It does not constitute either explicitly or implicitly any provision of services or products by arc. All statements regarding companies or securities are strictly beliefs and points of view held by arc or podcast guests and are not endorsements or recommendations by arc to buy, sell, or hold any security. Clients of arc investment management may maintain positions in the securities discussed in this podcast.

All right. Welcome to the Bitcoin brainstorm. Alongside my co-host Yassine Almondra, analyst at Arc Invest, I'm your host, Rod Rudy, the co-founder of Bitcoin Park, a community-supported campus here in Nashville, Tennessee, focused on grassroots Bitcoin adoption

and a home for Bitcoiners to work, learn, collaborate, and build. Yassine, my friend, how are you doing, first of all? I'm doing really well. Very, very excited about this conversation. If you don't already know from the title, we're going to be talking AI and Bitcoin. And I think this is particularly compelling at arc because of how often we talk about convergences across different technology platforms. If you're familiar with arc, we're not just focused on Bitcoin. We're focused across multiple technology platforms stemming from genomics and DNA sequencing to energy storage to robotics to, of course, AI and broader crypto. And so maybe we can start with a few intros from our guests around the clock and then get into maybe Kathy, why this conversation is particularly compelling from Arc's point of view. I love it. Let's do it. Let's start with a few intros.

Yeah. And actually, maybe even before we jump into those intros Yassine, just to recap, this is our second episode. And last month, we recorded our first episode building and scaling with some amazing guests, including yourself, Kathy, Jack Mahlers, Obi from FETI, Preston Pish, and Harry Sudak. So if you haven't already listened, be sure to check that out and subscribe to the Arc Invest podcast and YouTube feeds. And also, if this is your first time tuning in, maybe I can do a good explanation of what we're trying to do here as well. And again, thanks again to the support from you, Kathy, and the Arc team. We're taking our monthly Bitcoin topic-based approach here at Bitcoin Park and applying it to this new monthly podcast series we are calling Bitcoin Brainstorm. So each month, we are planning to have a different topic and inviting amazing guests from a variety of areas within the Bitcoin community. And as Yassine eloquently said on episode one, as well as here, our aim is simply to drive conversation around Bitcoin. And the topic this month is AI and Bitcoin. And we're going to try to get through all of this in 60 minutes. And we're going to be covering a lot of ground. So let me do some quick introductions for everyone and then just jump right into it. So joining Yassine and I this month, we have Frank Downing, director of research at Arc Invest, who's focused on the next gen internet. Frank, thanks for joining us. Yeah, super excited for the conversation. Awesome. And Paul Itoy, CEO of StackWork, a platform that makes and combines AI models and humans to solve real enterprise automation problems. Paul, thanks again for joining us. Great to be here. Thanks, Rod. And Lalu, who's best known on the internet as roast beef without

the T, is the CTO and co-founder of Lightning Labs. His focus is on Bitcoin and Lightning protocol engineering. He led the design and implementation of LND and implementation of the Lightning network. And now is working on taproot assets to bring stable coins to Bitcoin and Lightning. Lalu, thanks again for joining us. Great. Yeah, thank you. Yes, super excited. Things give me a lot of fun. Heck yeah. And Cody Lowe, head of developer and product support for FETI. Cody is an open source developer and recently has taken the lead on using Bitcoin to solve and scale monetization problems facing the wave of new generative AI applications through the AI for all hackathon, which I'm sure we're going to touch on today. Cody, thanks for joining us. Thank you, Rod. And Kathy Wood, founder, CEO and CIO at Arc Invest. Kathy, thanks for joining. Really happy to be here, Rod. And so I guess I take it from here. Please do. Okay. All right. Well, I'm very excited about this. As Yacine said, Yacine leads our crypto effort, but because he works with other analysts, he's tuned into this idea of convergence. It's a part of our DNA. And I don't think most of the asset management world understands how powerful the convergences are going to be between and among the technologies that are evolving today. And the way I like to describe why they're going to be so powerful, the convergences, is we're looking at S-curves. So new technologies, they start out slowly and then they gain momentum and you end up into the steep part of this S-curve. Well, I think we are at a very similar point in both Bitcoin and AI. They're both moving into the steep part of the S-curve. So you have S-curves feeding S-curves, creating opportunities for explosive growth potentials. The first convergence that served ARC really well was Tesla. Most people thought Tesla was an auto company. It is not, of course. It is the convergence between among three of our major innovation platforms. Robots, autonomous vehicles are robots and will be robots. Energy storage, they will be electric. It is the most efficient and at least expensive solution. And three, artificial intelligence. They're going to be powered by AI, autonomous taxi platforms. And I think many people have been shocked at the explosive growth of Tesla and we haven't seen anything yet because autonomous hasn't happened yet. So pretty excited about that. And I've always thought there would be something to discuss in terms of Bitcoin converging with some of these other technologies. Our first experience with convergence around Bitcoin was right after Elon had put Bitcoin on Tesla's balance sheet. And then he was faced with an onslaught of controversy around how environmentally destructive Bitcoin is or was. And we ended up working with Block Square, Jack Dorsey and his crypto team. And we put together a seminar, which I believe is still up, right, Yasin? It's still up on. Yes. And I think it was called the B word. So we talked about a convergence that would turn the tide in terms of this thinking that Bitcoin is environmentally damaging. This idea of the convergence between Bitcoin and energy ecosystems. It's been very interesting to watch Exxon put Bitcoin machines in its natural gas fields to capture the gas instead of letting it flare the excess gas into the ecosystem. It powers Bitcoin mining machines. And they did that in six different countries around the world. There are other examples where Bitcoin has become a part of utility ecosystems. So when utilities put in very large energy storage units in a very sunny place, for example, like Florida, it will fill up the storage. Sun will fill up the storage and the excess power goes into minting Bitcoin. So utility ecosystems can even overbuild their solar and wind systems. So again, very exciting convergence. And I also wanted to alert everyone, we are putting out a paper in the fourth quarter about convergence and how powerful the growth trajectories become when these technology convergences take place.

[Transcript] FYI - For Your Innovation / Why Bitcoin is the Currency for AI

Now, I'd like to turn it over to Frank, who's going to give you a sense of how we are thinking about AI and perhaps some of the convergences. Sure. Yeah. Thanks, Kathy. I think what gets us really excited about AI and what we're seeing now with generative AI in particular, this ability for artificial intelligence and in particular neural networks to create new content for the first time. So rather than predicting the next best ad to show somebody when they're shopping on Amazon or the next best movie for you to watch on Netflix, with generative AI along the likes of chat, GPT, we can now create content. You can ask a question and get a really thorough answer as if it was from an expert. You should still fact check it, but it's really remarkable in that capability. You can do text to 3D assets, text to image. All of these things where we're able to create and produce usable assets for the first time. And we think that, especially because it's so generalizable across many different fields and many different job functions, has the potential to bring tremendous productivity gains in particular to knowledge workers. So we estimate there's around a billion knowledge workers in the world today. These are people that are spending their days at their desks and typing on their keyboards. And with generative AI, we can really accelerate and empower them in their daily workflows to be much more productive. And we're already seeing that in many areas today. If you look at software development as a really good example, products like GitHub co-pilot are already showing they can reduce the time of software developers to complete coding tasks by over 50%. So you can think of that as your developers can now produce twice as much code and they're not only more productive, the actual satisfaction when completing coding tasks is up. So we're able to produce more and we're happier as we're doing it, which I think is a really interesting outcome. You can see the same across many different domains, graphic design with all the image models is another example. And we think by 2030, artificial intelligence is going to be embedded into virtually every software product that we use, much like the, and it's a bit cliché to say this, but much like the internet is embedded into all of our software products today where it's enabled so much innovation that it's hard to pinpoint any one specific major contributor or use case. But we know the world is much more productive today because of the internet and we think it will be even more so by 2030 with the profiling of generative AI. And one of the things we look at to kind of understand that is cost declines. If you look at the cost to train these models and the cost to actually run them, they've been collapsing over the last three years. So back in 2020, the cost to train a GPT-3 model, this is kind of the foundation, the underpinning for a chat GPT cost \$4.6 million. And just two years later, by the end of 2022, you can train a performance equivalent model for under \$450,000. So it's a 10x cost decline in two years. And it's even cheaper to train that type of model today. And we think that's going to continue where these models become more and more accessible to more people over time as costs come down. And I think to bring this back to Bitcoin, I think it's interesting that cost declines enable a democratization of access and really shape that S-curve that Kathy talked about with technological evolution. And I think there's a really interesting role that Bitcoin can play in helping accelerate and proliferate that technology as costs come down even more, which is why I'm super excited to get to talk with all of you guys today about it. Let me ask this to Paul, because Paul, you're in the belly of the beast with stack work. I'm curious to know your comments based on Frank's framing and background there.

Yeah. I mean, the first thing that came to mind as Frank was talking was, how ironic is it that when I first started thinking about robots as Kathy was talking about, and I thought, okay, it's going to be folding laundry. We're going to see house painters go away, and then, bam, us knowledge workers that were all of us on this call got hit right away. Look at what's happening to finance. If you look at what's happening, I mean, we're, so just to give everyone context, so we started this company stack work several years ago, actually, and started paying people over the Bitcoin Lightning Network around the world to do micro tasks. Our big unlock doing this for an application we were running at the time was that if you try to hire people to do complex tasks, you really limit who can do that work. And so when we had churn and people would quit, we'd get another person, you're paying great wages at the time, data entry for \$20 to \$25 an hour is good wages for data entry. But people would get so bored they'd quit, we figured out, wait a minute, if we dissected these tasks that are really 50 decisions that are going on in the human brain and then broke them down into their smallest pieces and assigned those small pieces to humans and paid them over the Lightning Network, then suddenly we had almost unlimited demand for the tasks. People were just snapping them up as soon as we could post them. And then also, if that person wasn't there that day, then someone else could jump in. It's almost like Henry Ford with the Division of Labor. You can now do that with knowledge work. And with the Lightning Network, you can pay them anywhere in the world. So it felt really amazing to get messages from people saying, hey, you know, even at two hours a day of this work, I'm able to actually afford to rent my house, depending on where you live, that income can be major. And I said, hey, do you know, I mean, we had people sign up who didn't know what Bitcoin was, had no idea what was going on. So having this kind of be able to digitize work, send it anywhere in the world, digitize money, send it to the person doing the work, I think most people don't realize that Bitcoin and Lightning is also a data network as well. So being able to send data in the form of work, receive data in the form of an answer back, and then give them real unceasing money is pretty fascinating. So I think people are not just barely aware of that happening right now. So that's the context. But having human beings helping machine learning models, we've seen an incredible just wave of these models being used. So when we broke up those tasks into small pieces, some of those pieces were done by code. Two years ago, I would say 80% was done by humans, 20% done by code. The power law has flipped. It's now completely opposite. These models can do these simple tasks. I'm not saying that's a good or bad thing. I just talked about how great it is to pay people money. We're definitely on Team Human. But if you look at the forward progress of humanity here, these models are getting incredible. So just a wake-up call that is just everyone wake up, this is not, hey, one of those futures things that's happening down the road, this is like right now we're seeing it. I'm on Team Human as well. Yeah. I actually think this is really fascinating. We have a company in our venture fund. It's called Replit. And it's a platform for software developers. And the numbers are exactly what you just said, Paul, which is 80% of the work these developers do and is done by AI, right? Again, completely transforming the position and natural language processing, because we all speak the natural language, right? We understand that concept. We're all going to become programmers, right? We need to get to be better prompt

[Transcript] FYI - For Your Innovation / Why Bitcoin is the Currency for AI

engineers, right? But Replit has 25 million developers on its platform. There are only 25 million professional developers in the world. What's happening here, these are not professional developers. They're like you and me. Maybe you are programmers, but I am not. And even I, supposedly, will be able to do this. We love Replit. I've got the app right on my phone right here. Pay attention to their Bounties program, what I just talked about with breaking complex work down into small pieces. Think about what Bounties are. We have our own Bounties program as well. So you have bottoms up simple tasks. Hey, can you identify the stop sign in this image, right? That's what Tesla is doing billions of times a day in terms of labeling these images. But then also think about top down. What's the most valuable work out there today in terms of economic impact could arguably be programming? So this whole division of labor and breaking things into small pieces is being applied at the bottom of the stack or the top of the stack. And if you look at what Replit's doing with Bounties and the ability to code on their mobile phone, when that first came out, I was like, okay, this is a little bit like a pet project. No one's going to sit on the curb on their mobile phone and type code. Copilot comes out. Ghostwriter comes out. These language generation models for coding. Suddenly, you're not writing semicolons and dots and pound signs and all this stuff on your phone. You're more describing. You're pseudocoding. You're talking to a code generation model. And then getting functioning code out. And then you have automation for deployment. And it is true. I mean, I definitely think that we're going to see more personalized bespoke software and then we'll be more productive. So coders aren't going to go away. You're just going to have the demand for more and more code. What's that whole law when we digitize photography? When you digitize anything, you see an explosion of demand. The same thing is going to happen with logic and code. Absolutely. Yeah. If I can jump in here, I didn't know that you guys had invested in Replit. But I've been a long time Replit power user, which I think was funny when you brought that up. And so when we did the hackathon last month, I used Replit for the platform for it. And we had a ton of benefits for it. And one of them was building out Lang chain Bitcoin tools, which is this AI framework of Lang chain for using these large language models as a programmer where you don't really have to know everything that's going on under the hood. But using it off of Replit, I was able to deploy the tools that Rosebeef had made of basically giving AI agents Bitcoin so that they can go off and make purchases and then deploy that on Replit and then make a little tutorial around it. And then when people were involved in the hackathon, even though they don't necessarily know too much about how Bitcoin works, even though they don't know necessarily too much about how AI works, they're able to just deploy these agents, give them Bitcoin, and they can go off and do things. And so that's a little bit for bringing it to the Bitcoin element for this is that when we're talking about these large language models, there's a couple elements for it, which are really important. One of them is that while the costs of training them are going down, it's also every time you hit one of these things, you can think about it as setting a graphics card on fire, right? So it's like, I put an F something into GPT 3.5. And like, you know, it's relatively cheap, but still, like, I'm hitting some graphics card somewhere, I'm basically setting it on fire. And so it's not the same thing as when I just show up at like a server, like if I show up at Facebook.com, it doesn't really cost them too

much, right? Like you can kind of amortize that across the billions of people who are using the platform for it. But if every single time I showed up on Facebook.com, it had a GPT4 generation for me, well, that's two cents out the door, right? And so one of the cool things that we've kind of seen with people moving on to using Bitcoin and AI is that because the proliferation of these models has become so wide, is that it's really about the energy costs associated with it. When I have someone show up, if I have a free tier user and I want to give them access to these AI services, well, that's expensive. It's hard to amortize those costs across there. But if the user shows up and they can just do the micro payment with Bitcoin, like I think Lalu jumped in after this to talk about L402 specifically, is that when I show up, you can give me a Bitcoin invoice, I can pay the invoice. And so I've paid for the positive compute, just like the fraction of a cent that is going to be when I show up on your platform. And you know nothing about me except for the fact that I paid you. There's no sign up flow. I don't have to give you all my information for that one. And so for free tiers for using AI services, Bitcoin becomes a really interesting opportunity there. And Paul, I think kind of led the charge on this in terms of how does Bitcoin become more widely adopted? And so this convergence that you're talking about of that, there's a lot of stuff that people use Bitcoin for where it seems a little bit gimmicky, because it's like, hey, I can do that with credit cards or I can do that with dollars already, right? There's a bunch of really cool things that are now being opened up that become possible because Bitcoin is a better payments mechanism in this way. And so one of them, for example, is instant final settlement over the Lightning Network with no risk of a charge back. And so that's been something super valuable in working with different AI companies

where I can tell them that, hey, when you get paid for in this manner, there's zero risk of a charge back. And so I think Paul, he can jump back in when he comes back, he can talk a little bit about the charge backs for it. That's just a little bit in terms of like Bitcoin's conversions with AI there, where you get the instant final settlement. So when people use your compute, they're paying you as they use it, and you can get it down to a fraction of a cent. So every single time someone shows up, they're paying for exactly what they use. And so I'm always in the black when people use my services. And because it's a finally settled digital instrument, I could, for example, give some Bitcoin to a lane chain agent. And when it goes off and it does things, I've already given it the money, when it does its payments, those are final settlements. It's not like, hey, it's swiped my credit card 100 times, I go back later, and then I cancel like 30% of the charges or something like that, is that you can use this better payment mechanism for when you're using AI services. So I don't know, I'm a recipe for Paul, you want to jump in and spend on that? Yeah, that's a really good point. And I want to go back to the point around kind of like aligning costs, right? I think this is like one of the things that like, Cody and I did this hackathon that's when we sort of like got together and kind of like, you know, combined a bunch of ideas that we had going on. One thing we realized is that like, whenever people had these AI apps, basically, because they actually had their API key on the hook, they would, you know, if they went viral, they would basically have like a credit card to thousands of dollars, basically, right? So it's kind of like a way to send it where it's okay, you know, if you want to go up and go viral, basically, you need to be ready to like actually have money just to, you know, kind of like offer that stuff, right? So one cool thing about the way some of the cool tools that we created is that like, this actually allows the user to basically help, you know, the, the credit aligns with the costs, right? I can pay one Satoshi, basically, for a given query, basically, you know, at that

point, maybe, maybe the credit can actually either break even or monetize a little bit over time, right? Rather than like, have anything, you know, being free from the get go, this is basically a way to kind of like align the cost, you know, on that level, right? And Lage, this is something that I made, I think in like, you know, 2019 or so was something called L4 or two, right? So what it is, it's basically kind of like a way to sort of like embed a Bitcoin micro payment with regular API calls, right? And so people know the internet, they know something they call HTTP, right? It's kind of like a webpage typically, you know, I get like a webpage back, right? HTTP also has certain error codes, right? If you're, if you ever went to a webpage, maybe you had some URL typing correctly, maybe you got a four or four, which basically means not found, right? So in the very, very early in the internet, there's actually an error code called HTTP 402, which was payment required, right? But this is like, you know, in the very like late early nineties, basically, we didn't have Bitcoin, you know, SL was really thing, you didn't even really have e-commerce at that point, but they could they sort of have the foresight to put something in there for the future. And one of the ideas I have is, okay, well, what if that future is actually, you know, combining with Bitcoin at that point, right? So what HP 402 is, it's basically or L402 is basically a way that you can sort of like have like micro payments metered for regular API calls, right? So this means, for example, maybe some API call for like creating an AWS EC2 and stuff like that itself, right? Typically, I would have like some billing, no penetration on the back or something like that, I have my credit card, maybe I need to have additional account, things like that as well. While this is basically more just kind of like a direct tip for TAP, basically, so I pay some Bitcoin basically over the light network, which is like this global payment service that lets you have like very, very low cost payments. And then from that, I can actually then actually API is really cool thing about combination of something like, you know, the LLM platform APIs with lightning is that like typically many of these API is they charge like by the token, right? So maybe like you're typing in, you know, a particular word, that's like, you know, a few tokens, right? Maybe the tokens are a few cents, something like that, right? But lightning is very unique in that it can actually send very, very small amounts of money, right? So lightning can send even, you know, less than one set, maybe like a third of a set or even a tenth of a set, and basically they're very, very small micro payment on top of money. So now at this point, I can basically do a very much direct, you know, tip for TAP type basis, basically, right? This is really cool because now I can actually like pay as I go and actually stream this money to the AI bot or maybe all these other platforms, basically, this really cool because now number one, this allows things to go a lot more globally because now there's no necessary identity required, you know, necessarily, right? Maybe there's no credit card as well too. And the other cool thing now is, okay, well, you know, if there's no identity needed, there's no credit card, you know, all you need is Bitcoin, you know, these AI models, they're they themselves, you basically compose like, you know, billions of like, you can say like integers to floating point values, right? You know, their private key is just another, you know, integer value, basically, right? So now we can actually embed these, you know, this private key inside the AI model itself and actually let it hold Bitcoin and actually then interact with the greater world and actually to do things like maybe, you know, buy a GPU card, you know, once in the gun's replete. So this is really cool because now you have like once again, this convergence going on, right? And when I made this stuff back in 2019, you know, I was like, there was always like a sort of like a founding, you know, kind of like thesis for the for lightning on itself that will eventually have like machine to machine

payments, right? But back then, like, I didn't really know what the machines were, similar to what Paul was saying, he thought it was going to be folding, you know, his laundry or something like that, right? Now we're seeing these AI boss, I was like, oh, this is basically it, right? They can reason, they can plan, now we can give them tooling to basically interact with the world. So it's like one of the really cool convergence styles. Okay, like we should do something in this area, basically. And then Cody and I got together with Roy Janning in San Francisco, and they started to make some of these additional tools, which is really cool to see people out there, people excited about it. Can I ask a question here? It's very interesting to me that you're saying something called L4-2 is the enabler here that someone foresaw, because when I'm trying to describe to people, you know, peer to peer and what's going on here, and Bitcoin generally, I basically say, you know, in the early days of the internet, because it was just an information network, and in fact, in I don't know if you all wouldn't remember this, you weren't alive at this time, but I think in the 80s, it was illegal for consumers to use the internet. And then when we got a telecommunications act of some sort, that changed. But the thought was, this is an exchange of information, no one thought of commerce, no one thought of financial services. And so when I'm trying to explain what's happening now with blockchain technology is, I said, this is the layer they forgot to build in. But what you're saying is there's something else that someone with foresight put in to enable this. Is that right? Yeah, exactly, right? So, you know, the error code is the 402 error code, and our protocol is called L4-2 or Lightning HTTP 402, right? I think that's a really good point around kind of like it being an information network, basically. I think one of the biggest revelations with Bitcoin is that like, okay, well, money can now also be information, right? The money is actually just a sort of like, you know, 200 clicks bit integer, right? So I think that was like the big thing. And I think then as well, like, you know, we didn't necessarily have, you know, things like encryption on the web, which basically enables e-commerce, because before I would want to put my credit card over the wire, because otherwise you'd basically be in plain text. So once we actually had encryption on the internet, then they were able to actually make a payment system. But the payment system that they created wasn't internet native, right? There were these companies, you know, Visa, MasterCard, basically they were on the edges and also adding additional costs and things like that. Now we have Bitcoin, which is actually the internet native currency, and something that can actually be used very, very directly, right? And the cool thing about, you know, the Lightning network itself is that it's global and it's instant settlement and very, very low cost as well too, right? So now all of a sudden I can basically pay for a given token on an LLM or even anything, you know, smaller beyond that. So it's really cool kind of conversions there that once like you're saying someone had the force that, okay, we don't necessarily know what's going to be in the future. Let's add that error code there. And if you look at the specification, it says like, reserve for future use, right? Then I was like, no, no, the future is now basically. I would love to add just some insight from our research to maybe help summarize or visualize this for people. The traditional payment flow, what connects a buyer to a seller is more complicated than what you think when you just go up and swipe your card at a register and you get your cup of coffee. There's at least four parties involved between you and the seller. There's the payment service provider, either the Stripe online or the Square terminal in the kiosk in the coffee shop. There's the merchant's bank, so the bank of the person you're purchasing from. There's the card

[Transcript] FYI - For Your Innovation / Why Bitcoin is the Currency for AI

network, so the Visa, the Mastercard, the AmEx. And then there's your bank that's saying, yes, you have this money or this credit worthiness. And then the whole chain goes back and forth. And all this remounts to multiple things. One, a time delay in transactions confirming.

Two, settlement risk. And three, because of that risk cost. So we've estimated anywhere from on average like 250 basis points on a transaction or 2.5% to pay all these parties for processing these transactions. And what internet native money with Bitcoin enables is the buyer and seller to be connected directly, which is this internet native money that didn't exist in the 90s when the HTTP 402 concept was created, which really needed internet native money to work. And what these guys at Lightning Labs have been working on is basically using Bitcoin to fill that gap and connect buyers and sellers more directly and taking out those intermediaries and reducing fees and increasing the speed of settlement. And for anybody who's listening to things like, oh, well, it's so easy to still use a credit card. Why do I care about those intermediaries or the fees? Because the consumer usually doesn't see it. The merchant is paying for it or your bank is paying for it in some way. And getting denied access to these services, especially internationally happens. But even at ARC, we've had this issue. We have a corporate Amex that we used to pay for certain services. And I tried to sign up the entire research team for chat GPT subscriptions. And either Amex or OpenAI had a problem with it. And we literally got blocked and we couldn't pay for

it. And if that service was available for payment via Bitcoin and didn't need to go through the card networks, that really wouldn't have been a problem. And OpenAI could launch their services much more

broadly if there was a neutral internet native payment standard in Bitcoin. So the use case and the need is really there today, and it can dramatically simplify the payment flow.

I was just going to point out on for anything, we have a situation with these AI models where you have a digital good. I came from some work in telecom with voice over IP. You incur a cost and the product is gone. When someone gets a fake credit card and makes phone calls, if you ever in the 90s and aughts, if you ever drove around certain neighborhoods, you would see lines by pay phones. So why were 15 people lined up at a pay phone on a corner? I always stop and ask, what's going on here? Someone had a credit card that they got. It's not theirs. And those 15 people were making international phone calls on that credit card. And the phone company incurs those charges and you paid for that bandwidth for that phone call to tying up that copper line.

But the asymmetry is the product is gone. The person has completed the transaction and got the benefit of the product. When you type in the chat GPT-4 and ask a question, OpenAI is paying for that inference. As soon as you hit the return, that GPU cycle is gone. The opportunity cost of using that GPU is gone. Now, we have a mismatch here because the credit card system is operating on

a different reality. So the settlement description that Frank just described, OpenAI thinks they have a credit card on file. They believe they have incremented their ledger for your charge, for your prompt that they just paid for, that money's gone. But it's up to 60 days, 90 days to actually contest that charge. The money is then pulled from the merchant's bank account. This is very painful for any merchant that ever accepts credit cards. The money that you think you have, you actually don't have. When you sign up for that issuing card agreement, you're agreeing that the issuing bank can pull that money from your account. And that's when you contest charges and have fraud.

[Transcript] FYI - For Your Innovation / Why Bitcoin is the Currency for AI

So think about this for telecom, that the digital good is gone, the credit card liability still lingers. Now, think about that for these models and these viral OpenAI models that are being paid with credit cards. You go to Huggingface, pay with a credit card, go to Replet, pay with a credit card.

You're using compute that's forever gone, electricity, all that stuff. But the money isn't real yet. I hope this is going to get in braining in everyone's brain. Wait a minute, this is a huge mismatch. Now, insert Lightning Labs and what roast beef did with an L402. And you're moving Satoshi's over the moment that the good is used. And we're talking about hundreds of a penny. There's no county party risk. I cannot reach into Lightning Labs wallet and pull back the money. So this is an incredible step forward that I think any business owners out there who feel this pain, it's beyond just the financial pain. It's like the emotional pain. You put so much into your product. You think you sold something. You think you earned this money and then the system pulls it back from you. You can tell I'm still infuriated by it this day. And so when I see these model companies out there hosting these things, I'm like, hey, I've been there. I know what you guys are about to run into. You're going to be running into the biggest credit card. Look how easy it is to get credit cards on the dark web. So I think they're feeling a lot of heat from this and they're feeling a lot of pain. So that's kind of the why should you pay attention to this and do this? Right away, there's a dollars and cents reason to do it right now. And it's honestly the first developed economy use case for Lightning I've seen. Every other use case of Lightning that we deal with is an emerging market. It's the Global South. How do I get work? How do we get paid? That's the whole thing that isn't even on anyone's radar here. There's no one walking around a suburb of San Diego thinking about how do I earn Bitcoin right now? It's just not that common. But this is actually every company's deal with this. Go ahead. Yeah, I was just going to say it's not just dollars and cents, which you're really saying and I think what Frank was describing was you're lowering the risks of this payment, actually almost eliminating the risk of the, you are eliminating the risk of the payment. That's crazy. The reason we pay two and a half to four percent for or the merchant has to take that hit is because these, I call them middlemen. And I think on our last Bitcoin brainstorm, it was like light bulb went on. No, there's a lot of risk. There's a lot of risk cascading through the system. And there are seven steps between the merchant and the consumer. Those are seven risks that don't exist. And compound that on the supply chain. When we globalize things, I'm not just talking about one hop from the store to the buyer, from the store owner to the distributor, distributor, follow that supply chain all the way up. What's in between all those? Net 30, 60 day contest cycles. So your view of reality, your resolution on what your reality is, is incredibly imprecise because for that product to hit me, I have got six levels of net 30 stacked all the way back to some factory in China. There's county party risk along that entire chain. And so what we're talking about right now is we're starting with this little consumer B2C model with a person with a credit card making fake phone calls or ripping off open AI with a fake card. But if you actually extrapolate this all the way back out to the economy and what can happen is when you eliminate counterparty risk, suddenly you start going, why does my invoice all say net 30? We've accepted this whole thing of net 30 forever. Why? Because the payment system is so inefficient. I got to give myself 30 days to process a payment. That's absurd. And so you add globalization to that. And now you have this ridiculous bungee cord of train cars just slamming against each other, creating incredible inefficiencies in the economy. So that's off the AI

topic. But I would say AI is a very focused version of this problem. And then not to mention, if you think about all this impactful technology, ChatGPT4, my kids use it, it's amazing. Who has access to that? People with a credit card. Who can afford \$20 a month to pay? That's a very small subset of humanity. If this is truly powerful, are we really implicitly saying that in order to have access to this, these are the hurdles you have to get over to have access to that? We should all be very concerned if that's the default that we're just letting happen. Paul, are you suggesting in that case, the barrier to entry from the payment standpoint of, okay, you not only need to have access to the credit card, internet, but also be willing to pay \$20 a month? Are you also implying here that if you were to integrate some sort of micro payment service that call it streamed sats where you could basically pay on a single prompt, micro sense, instead of needing to pay the bulk and committing to a monthly subscription, is that what you're implying with something like the Lightning Network, where you have this internet native financial system where you could actually be much more efficient in how you're buying a product or a service? Not just implying, stating, and it exists today. Cody's doing it, Rosby's doing it. Let's raise a few micro payment use cases that we like to call it kind of streaming sats. Maybe some of our viewers aren't aware of. Yeah, if I can jump in on this one, because the stuff that Paul was talking about there, I specifically ran into when I was running this AI hackathon in the last couple months. The reason why, because I live in San Francisco, so when the generative AI boom started off and there were all these hackathons every single week, I was like, hey, this is amazing, but I would talk to my friends about it. They're like, I'm not in San Francisco, so I can't do that. I planned out with Paul and we got a bunch of other people within Bitcoin and AI involved of a global hackathon focused on the intersection of Bitcoin and AI. One of the first things that I ran into is that when I tried to get a global audience like Kala, which works on training Bitcoin developers in Africa, and there's a couple others like Vintium, which work on building training Bitcoin developers in South America, I would try to get them involved in the hackathon. And they said, okay, well, I can't get an open AI account because my credit cards won't get accepted. And the reason why is because their fraud rate is so high, so that when I'm someone from Togo or Nigeria or Kenya or something, like, I'm just out of the box, not able to sign up for the service. And so I was like, hit this dilemma, right? I was like, okay, well, I want to get all these people involved, but they can't get access to the tools in order to get involved. And so I built a service called Matador, which was based off of Rosebeef's aperture that he built, where it basically just takes my API key, and then anybody in the world can hit it. But the moment you try to hit it, instead of using an API key to hit open AI, you hit my service, and it's just a pass through. But the proof that you get to use that pass through is that when you hit me, I give you back an invoice. And so I've got like a demo of it. It's like `matador-ai.replet.app`. Like, I think it's currently up right now, and I'm rewriting it right now. But that was something that a bunch of people used already, because they didn't want to either couldn't sign up for the service because they couldn't afford it, or they just couldn't even get their credit card approved in order to do it. But they were able to use those tools because they were able to ping off my API key. And this is something super, super important when I'm talking to the different AI companies for this, because like I said before, is that, and Paul was like totally on top of this, of that for a lot of stuff with Bitcoin, where we're trying to say, hey, except Bitcoin versus dollars or whatever, right? Like dollars, there's so much infrastructure. It's got the regulatory clarity. It's got all that. It's a

lot easier for people for this. But if you can show up to an AI company and say, hey, the entire portion of your revenue that comes from people who pay using Bitcoin has zero risk of a charge back, right? That's super valuable. And that's really interesting to a lot of people who otherwise wouldn't be. And so that's like a very specific use case where I already saw it where there were like developers who wanted to use these tools who couldn't because they couldn't afford it or their credit cards couldn't get approved, who could use my service for it. And it's just a pass through, right? And ideally, I don't even want to run the pass through, like I just want to give like the 402 to the company doing it because, and this is something that just kind of the last moment before this is that I think that the thing that a lot of people are running into right now, where they can kind of see this most obviously is like everybody is scraping everybody else all the time for training data, right? So like Twitter has been talking about this, like the last couple of weeks, everybody's sort of talking about this of that because all everyone's scraping everybody else, all of your APIs are getting hit, all of your services are getting hit, you got people and they're just bots and they're not paying for it. But instead of giving you back an unauthorized error, right? Like why are you unauthorized to use my service? Well, because you haven't paid me, right? What's really powerful about the 402 error code is that when you hit my service, in the process of denying you, I'm also telling you how you can access it, right? And in exactly that amount that you would need to access it. So every time someone hits my service, instead of just saying no, I say if you pay, right? And so that's something that like I've been talking to some companies about this who do authorization and they're trying to identify humans versus bots. And for right now, their solution for the bot is they just deny the bot entirely and say you can never use my service, right? But just make it pay a little bit, right? And then all of a sudden, what's called, like I'm always in the black whenever use my service, right? So can I ask you this \$20 a month open AI? We've been wondering, you know, Meta is offering it's for free, right? It's foundation model. And we've been thinking, and others are as well, we've been thinking that this would commoditize this foundation model because all it is, everyone's doing the same thing. It's scraping the internet from the beginning of internet time. So is there, I'd like to, you know, this is something we're grappling with. Is this, is this commoditizing? Or is this going to be a very big business? So there's a couple elements for it. I think that you're right in terms of like the costs are coming down. It's being commoditized in these different ways and all that kind of stuff. I think just generally speaking, the specific element of it is that when I'm trying to offer a service, right? Like currently there's like a percentage chance when I have like a new customer who comes on that they're going to stick with my service or that they're going to be monetized or anything like that. What's really powerful about this is that it doesn't matter what the service is, I can open it up to anybody at any time. And I'm all, you're always paying for it as you use it. And so like this is something like just for example, one of the things that came out of the hackathon was this concept of Nostr data vending machines, where it's literally you just have an API that is listening to a social media network of Nostr. And anytime it sees anything on Nostr that it thinks it can provide value to, it sends an invoice, right? And so now it's not even having people sign up for the service. It's just looking for things that it could possibly help with and it's providing an invoice. But because the invoice is a final settlement, because I'm not running the compute unless I get paid, right? Is that it's always profitable. It just sits there

and just makes money in so far as people use it. And if people don't use it, then like it doesn't cost you anything. Well, I mean, it costs like, you know, you have like an \$18 a month box that you're running this on or something, right? So like to specifically answer your question, that one of that, the foundation models that people are using, yes, but then you still have to fine tune it over like a specific use case that you want to do and offer that as a service. And when you offer that as a service, what's the percentage chance that you're going to do that? I think Paul would be a good person to jump in here because he's kind of got like a whole workflow set up within stack of using these models and fine tuning them and that sort of thing. I'll describe our workflow on it is to answer the question, we completely believe that it's going to be, at least for enterprise point models, many, many models that are good at narrow solutions versus one large black box model, like a foundation model that's going to handle all your needs. If you think about it, just take a use case like answering a customer service email. So there's no company in the world, let's take Hertz rental car, that's going to take a customer service email, pass it to this great foundational model, let's say it's chat GPT six and then take the answer and send it back to the customer. Then yep, we've refunded your entire weekly rental, here's \$2,000, which is not going to happen. And so what you're going to have is small models that you've trained, we use Lama two now almost exclusively. We started off with, hey, does it work? We'll use chat GPT three, five. Ooh, not quite. Let's go to chat GPT four and pay for that. We'll use chat GPT four to do that particular task. Then we go, all right, let's find now that we kind of know what we want. Let's fine tune flanty five or remember that Falcon model that came out for a little while from the Middle East. And then right now Lama two is it seems to be beating most of them out on these specific tasks. Then you have a workflow that says, okay, first pull the topics out of the email, then pull customer data on this email address from our internal database that's retrieving information for it, then pass it to a third model to write the customer surface email in the style that this company would normally write it in. Then you have a human being check it at the end and go, wow, that saved me about 80 to 90% of the work. I made two fixes to it, signed a personal message about the weather so it doesn't look like a bot. And then I went ahead and hit send. So this is like the Centaur model. It's human with an AI assist. And I think that that's going to exist for the foreseeable future. And so that's why we focused on the entire thing is the workflow. Models are going to model. You're always going to need humans. And then we just kind of don't care which one is better necessarily, we just plug in the best solution for that particular step. And then whenever possible, we're paying with Bitcoin along the way. So hopefully that answers the question. Yeah. And one last thing to add on, Kathy, I think even if the cost of the actual model is completely collapses and it's open source and open source is the best model and you can download it for free, the unavoidable cost remains of the actual electricity costs and your hardware, whoever is bearing the hardware expense to actually run the model. So those costs will remain. I was going to say, as far as the view of these foundational models, I think number one, open source is very, very important here because we want to make sure all this is widely accessible. And I think we said before, this basically makes everybody a power user. It basically is going to increase the average intelligence of anyone in the world basically. So it's our duty to make sure this can be distributed as widely as possible. And you're taking foundational models. There's a lot of things that goes into the training of them. Maybe, you know, tens of millions of dollars to get the weights, but once you have

the weights, you still need like a pretty, you know, expensive GPU, maybe like six K plus to actually run this thing effectively, right? But the cool thing with something like L four or two is that like, let's say I put in that money, basically, and I have the GPU, I can actually resell that out to basically the internet. I basically run this proxy, this L four or two proxy in front of like, you know, some of this API model, I make it look exactly like the open AI API itself. I was like, now I can actually then monetize my GPU that I, you know, put some money down over time. Basically, I'm getting paid per query, which is very, very cool thing. Right. So I think like in the future, it would be the case that like, you know, I'll pay a different cost per token for, for the varying models, right? As we know, like the, you know, we can basically increase the parameter size of these models to some degree and they get better and better. But the bigger models basically take more memory, more electricity, more CPU. So maybe I'll pay, you know, one sense of query for a GPT three, but GPT seven or something like that. Maybe that's going to be 10 cents a query, right? This is really cool because I was like, now that we can like align costs very much directly and such that like, you know, I'm more even, we're even just right at the, the boundary, like the marginal cost, you know, per the token. So now we have these open models, we can use L four or two to basically distribute around the world and basically have people pay for them. And now we can actually distribute this a lot more widely. And I think as Frank was saying now, like this is really cool for these API platforms as well, because now they can launch globally from day one. They don't need to say, oh, we added Nigeria, we added, you know, China, we added, you know, South America or South America or something like that. They, they're basically hitting the launch button. It's global from day one. It's really cool because now everyone around the world can access this stuff and all these, you know, teams around the world in the developing countries that are, you know, hungry to learn and, you know, just further themselves. We now have this very, very cool outlet of both on the side of them, you know, being able to interact with these tools. But then also on the other end, people being able to provide these servers to them in a way that's very much open and very low costs and using all these open standards, open source models. So we have like a really cool open source technology. All of our tooling is open source, the model is open source. And this is basically now when all this, you know, the cost goes zero. Now we can actually build on top that and invade a lot more. Yeah. If I, if I can add a specific one for agents for this, and this is something that like, you know, this is something that grew on me because in the original L four or two specification, there was a requirement that you use a macaroon as a token, which is just a fancy cookie. It's like a super, something super fun that Google came up with a while ago that roast beef is like fell in love with. And, but basically what's super cool about these macaroons is like, it's like a API token. It's like, I give you a key. And what's cool about a macaroon is that you can take the key, add additional restrictions to it and give it to someone else. And they can use the same key that I gave to you initially. And this becomes really powerful with agents, right? And this is something that like people started working on as a way to make agents more production ready and make them more valuable is that I can, for example, give a macaroon that gives access to a Bitcoin node to an agent. And now the agent has a budget. And so the agent can self evaluate when I give it a task and say, do, can I do this myself? Do I need a human for

this? Do I need some other resource for this or something and use that budget in order to purchase access to the things that it needs or like purchase to post a bounty or something onto stack work so that some, this human solves the problem that the agent needs done so it can continue with what it was doing. I think Paul's made some moves in this direction as well, but that's something that like I found super cool of that you can give a budget to an agent and then if you ask it to do something that it can't do by itself, it can either purchase resources or post a bounty so that a human does it for it and then it can continue with what it was doing. And that's just a way because the way that you, the problem with agents and some of these large language models that people are running into is that it's really, really hard to tell it all of the tools that it has available while keeping it also specified on the specific task you're trying to do. But the way that you solve specialization within a marketing economy is money, right? Is that I get really good at one thing and I pay other people who are good at those other things to do those things. And so you give the agent some money, the agent can go do things and when it can't do the specific thing by itself, it pays for that component of it that it needs to have someone else do, right? To give an example of this that I've made up and would love to put to you guys to see if it accurately represents this. This is how I explain it to other people because I find this concept really fascinating is like you can go to chat GPT and say, how would I create a new ad campaign for my new product? And it'll probably do a good job of walking

you step by step. You come up with kind of your thesis and then you do your graphic design and I'm simplifying and then you buy ad inventory on Google and you deploy your campaign. But then if you say, go do it, it's actually going to be not able to do it because it's just within these walls of open AI's chat window. And the idea of an AI agent in general is a combination of models that you can basically form to act more autonomously and go engage in the real

world so they can go reach out to services and enact whatever your vision is. But the problem you run into very quickly is my model probably needs to actually be able to pay for things to get stuff done. My model is not an expert in graphic design and it can't create images. So maybe it needs to go contract out to an image model like a stable diffusion model. And then my model isn't Google AdWords. And so it needs to be able to pay Google to get ad inventory and deploy your campaign. And what this kind of combination of frameworks that Lightning Labs has built for using L402 basically would allow an agent to hold a Bitcoin balance and then interact with these services, the ones that allow token-gated payments, what we spent the last half hour talking about, and interact and be basically its own power user able to be much more autonomous and going about its work and basically fulfilling your task to a greater extent than it could have been possible without internet-native payments that these systems accept.

Is this already taking off so that the scenario that Frank just described, is it happening and is this spreading like wildfire or is this labor of love trying to educate? Fair question.

It is happening.

Yeah, it's happening. A lot of the tools are open source. For example, there's something called LangChain, which is this API framework or library to basically make agents, right?

What we did in concert with Cody and myself who were Lightning Labs, we basically gave the ability to actually hold Bitcoin itself, right? So this basically means you can hold the Bitcoin,

you can ask the questions, they can then start to use that. Then we gave the ability to actually navigate these L402 payment API gateways basically. So kind of like handling the error code, constructing their questions and sending it back as well too. So all the primitives are there now, and then so you can say now, what's the missing piece? These companies starting to actually run this proxy basically in front of the APIs, right? And it's something where it's like, what Cody was saying is more or less just like a pass-through. So I think it's something that I think has the most potential and I think it's when you get sort of the most sci-fi kind of future type of thing basically of this thing able to sort of hold the funds, maybe you can manage individuals, maybe you can even grow its funds over time. I think the most important thing is basically you're able to actually sandbox this thing by giving it a very set budget of money, right? It's not going to go away and run away with your credit card and have a few thousand on our credit card bill. It's going to run out of, you know, statutions at some point, right? But you can even recognize that potentially and then even maybe, you know, become its own sort of like a service itself. So I can imagine even agents in the future paying other agents basically for particular tasks, right? Maybe one task is very, very good at particular translation or, you know, data summary or like he was saying, kind of like your image generation basically, you can imagine the sort of like commerce web of these agents paying each other and even other humans as well because, you know, because all this is programmatic, you don't necessarily know if it's a human or an agent or dog, right? You know, as they say, you know, back in the day. So I think that's the cool part there. But I think it's still a little bit early and this is where we think it's going to be going. But at least today, like the barrier isn't necessarily there. It's just a matter of kind of like selling that use case and getting those integration points into place. Yeah, I can add some specific examples from and then let's go back to Paul because like, I know the example that he's going to bring up. It's a very exciting one and it's very interesting. But the of the last month when I was working with AI companies of trying to get them to add this of that they kind of get the arguments, they get the use case, the specific issue that they run into is that now I'm accepting Bitcoin on the balance sheet, right? And so you go to the developer who's very excited about adding this who kind of gets it and sees how it could be integrated and then he has to go to the back office and go ask one of the lawyers or ask one of the people involved in compliance of now we're going to be accepting Bitcoin, right? Of how do we handle it? How do we custody it? That's like a whole separate step for that one. But what's been really cool and this is something that I didn't even know people were doing this already, but I think that a bunch of Bitcoiners have identified this is that you use lightning addresses or you use something like that. So in the lightning addresses, I can accept Bitcoin payments to like Kodi at replit.app or something if I wanted to, right? I just set up a pass through for it. And so it hits a Bitcoin node, gets an invoice and you can pay it to that one. So people can pay Kodi at replit.app. But one of the cool things you can do is you can use one of these services that will swap that into dollars. And so the thing hitting the business's balance sheet is dollars, not Bitcoin. And so then at that point, one of the things this has been like kind of useful in terms of getting people to start moving in this direction is that the developer can just add accepting Bitcoin payments on the front end. But the thing hitting the business's balance sheet is dollars. And so therefore they don't have the added complexity of the capital gains of handling the swaps of handling

all that sort of thing. And so with that primitive and this stuff available, now you kind of adjust the developer by himself at the company can kind of add these things in without having to get all the checks that he otherwise might have to. And that sort of sets the table for getting people to add this in. So who's the intermediary there? Who does that conversion for? Yeah, if you wanted to do it this way, then you would have to use some sort of service like River or Strike or Ebex Mercado or like any of these guys who handle the banking on the back end where they can deposit to your bank account. So this would not be like a purely self custodial way of doing it. But you know, if you're not going to be using Bitcoin exclusively, you're going to have some sort of intermediary or custodian or something like that. And then like maybe eventually you get like stable assets like like Taro or anything like that. Or sorry, tap read assets, right? But let's go back, Paul, because you wanted to add that example of how you're using agents, right? Oh, no, I think you guys covered it well. But yeah, I sort of, I'll bring it up later. It's fine. But there are some real world examples of these things. I was going to I think speak to Frank's example of an ad campaign. So something that's similar that we were hired to do by a customer that's building celebrity websites. So we went through their entire process. This has created people, non-programmers and sort of saying, all right, we build hundreds of these sites for movie stars and whatnot a year. And they have a very clear process that starts with interacting with the artist, what are your favorite colors, all these different things, and they go all the way to a website launched on WordPress. So you end up in code and images. And so we took their process and we basically created these mini bounties and paying people Bitcoin over lightning, go out and get the latest Instagram posts and Spotify album covers and YouTube videos. So we'll pay people 50 cents, 25 cents per image related to that celebrity. And so instead of the \$150 designer going out to all these places and looking around, they just set the job in motion and people around the world crowdsource all those images, drop them into a bucket, then we have all these images. Sorry. So what do we do with them now? We have a model that does image to text. So then the model actually looks at all the images and describes what's in there. Yellow keeps coming up, purple keeps coming up, cars keep coming up, the word animate keeps coming up. So now you have all of these words, and then you can actually generate a prompt. So then what we do is we take the word word soup, and we generate a prompt that we then send to a fine tune version of stable XL. If we haven't mentioned him yet, Imad Mostak from stable, is it stable diffusion? I don't know the name of his company, but he's doing open source mobs. Stability AI, that's it. So stable XL is amazing. It's within shouting distance of mid journey. Pretty incredible. So you can pass that prompt to stable XL and then get back 20 comps that the designer would have taken forever to do the research, pull out the elements, generate these images. Now, are you going to send that to the artist? No, no, no. You're just getting these ideas, like all of the cognition that would be required to mix and match all those things, and you're getting to see some really creative comps. So then they take the comp, then they can feed it back through using another model control net to say, I love the pose here, but I love the anime look over there. So now control net on top of stable XL gives you the pose that they're looking for, but the style from this other image. This is what you guys were talking about with combinatorial technology. Someone created control net, someone

created stable XL, somebody created Bitcoin. All these things are working together now. You could not do what I just described without these tools. I mean, you would have been, this is the stuff where we would just been talking about these cool ideas that happen in the future. That's happening right now. There's no technical hurdles to that whole situation. What we want to see is, and this is a shout out to people who are called for startups, would be people creating the best models. Maybe you can fine tune stable XL better than you can, right? And so now you have your shingle out in front of your GPU, and you have this great model, you advertise it. Our machine pays your machine for the use of your model. At some point, someone's going to hack a Tesla to do this, and while it's sitting in the garage, it's actually cranking out inferences on top of models, but can be any computer, can be anywhere in the world. This stuff can happen is just, do you have GPUs? Do you have TPUs? Can you generate inferences? Can you run matrix math? That's the only question. Can you? Great, we'll pay you. And then we hired someone I didn't even know. I worked in El Salvador. I've never met this guy. Rafael is an awesome programmer, incredible self-taught machine learning programmer, wrote a transformer model that does specific handwriting recognition for our company that beat the pants off of Azure and also AWS's text track. So we got a bill for \$47,000 last month for text tract. That is so wrong. If you think about how badly we played that. Yeah, that's a big bill. I respect their ability to extract money from us. So we have replaced that completely with our own transformer because we don't need text tracks features. We need it to do this. And we're seeing this on the front lines every single day. And what I would love to see is Rafael's friend who we've never met create a model that we can use. And then we pay him via an L402 to use his model running on some unused PlayStation 5 or gaming PC in his basement. I don't care. Like, did I get the answer back? And he doesn't care because he got paid. We don't have to sign a contract. I don't have to do a vendor vetting process. It's like, hey, did I get the inference back? Was it better than text track? If yes, I'll pay you. Done. Let's move on. So we've talked about El Salvador. Do you think because they've deemed Bitcoin legal tender that you're going to be finding a lot more talent like this there? Is that happening? I hope so. That one was completely random. We just started working together. And I said, by the way, where are you? And he goes, oh, I'm in El Salvador. I don't know. No kidding. I've been 10 minutes from your house. So hope to meet him one day. But yes, I think there will be more talent. I think there will be more exposure for that talent to reach other parts of the economy. So it shouldn't matter. Your manifest destiny should not be based on where you're born and the poor monetary policy of the country you happen to be born in. It's just not right. Right. If anything, what this indicates is that the barriers to entry have just completely collapsed. Like listening to all you guys, it's like, it's just a hyper-efficient allocation of resources now when you've kind of embedded native economics onto the internet. And you don't need to ultimately trust anyone because the incentives are one in which if I'm getting paid, I don't care how you're getting there. And if I'm getting what I want, I will pay you. And so it's just very compelling dynamics. I think even when it comes to, you know, maybe I'll ask you guys, like, what is the, what is the knowledge worker like in 10 years look like now? Where it's like, you're almost your own startup. You're doing all these micro tasks. You're not really, so you're, you're sort of global in nature. So you're hopping from, you know, place to place. And it's like, I guess, what's the, if someone asks like a guy in El Salvador, what do you do? What is that? What is that answer? I'd be curious, like, what is the next generation of knowledge workers actually look like with this hyper-efficient allocation?

Yeah, I think you can say he's like managing, you know, programming agents, right? I think it's one of the biggest things that cogeneration is just so good now, particularly for things like, you know, Lama 2 as well. It's almost as if everyone has like their own personal, you know, fleet of interns or even like, you know, domain experts as part of, right? So I think it goes from the level of like the program of being able to do very, very quick prototyping, maybe some, you know, quick analysis and even like, oh, let me learn iOS development over the week and I can do that. But I think we'll also then have like these, you know, much more specialized models for particular tasks, maybe operating systems, writing Bitcoin in front of several other else. Those would be hosted elsewhere. In my ID, we have to, you know, pay them automatically using Bitcoin and Lightning, basically say, oh, use some code review, oh, do a security review, do a security audit, right? Things that cost so much money now because of, you know, all these additional, you know, risk, particularly in the Bitcoin cryptocurrency space, I think you'll be able to like, you know, farm those up more readily and just be able to kind of like have like a one man or one person, you know, one woman family team, basically, from any of these really qualifications, because the marginal cost is just, I like that. Yeah. It's like, I'm a full-time robot manager. Yeah. It's like, how do you get paid? Well, you know, sats are just streamed to me. Yeah. And you hold up your phone and it's just, you know, you just see the number going up. Yeah. So that's specific one that Rosebeef was just talking about. That was something, so during the hackathon last month, one of the things that came out, like I said before, is Noster Data Vending Machines where you just submit jobs that you want done, like I want a summarization or I want a transcription where there's just something I want done, but I don't necessarily care how it gets done and I don't know how to do it. But I'm saying I want this done and I express my willingness to pay of that, send me an invoice and I'll pay it, right? And so just as an example, I built a replet extension that does that where it takes your PR and then it publishes the code difference between like the code that you had previously and the code that you want to merge, it publishes that to the Noster network and asks, I want a code review of this, right? And then there's like just basically APIs that are just listening and they're saying, hey, I can do a code review because I'm like a fine-tuned version of like code llama or I'm a like GPT4 or something or whatever. And I will do that if you give me, if you like pay this Bitcoin invoice for me. And so that already works. And that's like a replet extension that I made with Ironclad Dev last month. But what's really cool about that is that for all sorts of these things, and this is where like to answer your knowledge worker question, right, is that like knowledge just becomes increasingly specialized and you become increasingly easy for you to, when you encounter something that you don't know how to do yourself, is that I just want this done. I don't know how to do it myself, but I can demonstrate my willingness to pay for that one. And I can instantly and finally get that task done for me, right? And so I think Paul, he is like the most unified kind of conception of like what bounty work looks like in the future. But like as an example, like if I wanted to pay someone like I need this machine learning thing done, I need like the best machine learning engineer to do this, I can pay that person 150 bucks in order to look at my project for three minutes, right? If I'm running an agent and the agent, I didn't want it to go do something for me asynchronously, I can say, Hey, I earn like 17 bucks an hour or something, right? Or whatever it is. And I can give the agents like \$10 and say, Hey, basically over the course of the next hour, like just figure this out for me. And that's like more economical than me having done it myself of that it just goes off and does its best to go do these things. Or if it recognizes it can't do the

tasks, then it just comes back with stuff. But just everyone just becomes much more productive and much more specialized. And you're able to sell your specialized services to a much wider audience, right? Fascinating. Can I just ask this a tiny detail, but that \$47,000 bill, how much did you collapse it to? Oh, yeah, yeah, yeah. So that's down to \$4,000. We'll get the full benefit in 24 hours, but we're talking the 10x improvement on cost. So mistake made on our part. Wow. Yeah. And this AI stuff gets out of control. I mean, we've hit every limit from every provider where they stop serving us, because the stuff just runs and it'll soak up all your money and all your budget if you don't watch it. So we have counterparty risk among ourselves. So I think the full like departments owing other departments is an issue. Hey, who spent my money? It's not like you can blame DevOps anymore. So you're going to see this really penetrate all the way in and just to speak to the bounty thing for a second. So just, I mean, I am more convicted about this than almost anything, just the nature of work. Look at what coding is today. We've talked about dissecting complex things down into small pieces. You have to predigest human thought to automate it. You're just exposing more surface area for automation. So that is a huge, huge, huge thing to think about. It applies to legal work. Look at e-discovery. Look at anything that requires someone to sit and think over and over again. By breaking that thought process up, you are exposing pieces of that thought process to automation. And as soon as you do, I don't think the job goes away. I think the usage of those parts of humanity that are needed more, blossom. It's going to be this weird thing where we're going to discover these inefficiencies in the way we do things that we just never even considered. And as soon as we start breaking it down and looking at it, we're going to go, why do we do it that way? What if that went away and magically became infinite? That's wild. So I think they'll be the positives that way and the negatives that people talk about are infinite identities, infinite content. The cost to create these things just went to basically zero. And so you have the scarcity of Bitcoin hitting the infinity of AI. These are two colliding forces. And I would put my money and my bad and obviously our whole company on the best training for this world has been people who have been born and raised in the Bitcoin adversarial private keys. It'll be keys to your money and keys to your data as well. So just Bitcoin or his understanding AI, it's just a one for one. You go, I get it. Yep, I've seen this with money. This is going to happen with identity and content. You're not going to know who to trust. We're already there. So it's an easy, easy transition, but now it's going to hit the mainstream and it's not going to be optional. You're going to be bombarded with it. You have a fixed attention span. So if you have infinite demands on your attention span and it gets better and better, then what do you do? You have to have a defense system, which is your personal AI. So, Yasin, you're talking about, I'm going to open up my phone. Work that you did a year ago for someone may be earning you passive income. So think about it. We think of money transfers as transactions because the counter parties, the banks that we require to do this, require us to lump it together. That 35 cents or whatever people are paying now, 2.9% plus 35 cents of transaction immediately exposes what can be a transaction. But when you get into the streaming of money and the streaming of value, you're going to think more like, what holes in the bucket are leaking money and what's pouring money at the top of the bucket? And you're just going to be thinking of it more like a plumbing system than this package system. So wallets are going to change. Your conception of your value in the world is going to change because the nature of the exchange of value has changed whether people

recognize it or not. As suspected, we covered an insane amount of ground in a little over 60 minutes. And I'm sure we can go on for another 60 easy. I think Cathy, Yasin, Frank, and everybody, I think we're going to need a version two of this discussion. That's for sure.

Before I wrap, maybe with some closing thoughts, let me open it up for this group.

Are there any, anyone want to jump in with any closing thoughts or points they wanted to make that they didn't have a chance to make? I think just kind of a last thing for it that we didn't really touch on for it. But one of the cool things that Bitcoin enables for this is this kind of spectrum of cascading model sizes. And so Paul's talked about this a little bit in terms of people running their personal models. But some of these models or whatever, if you want the most cutting edge stuff, you're going to need something a little bit beefier. And so having this sort of spectrum that you can go up to where it's, hey, maybe I'm running this locally, but I can't do it locally. And so instead of having to go hit GPT for something or like some big third party service, there's like a way for me to cascade up as we're going for this. And that's kind of the way that that's something that wouldn't necessarily be possible today, because you'd have to swipe a credit card. And when you swipe a credit card, then basically you're involving with all these third party payment processors, all that sort of stuff. So that's kind of like, you get some economies of scale of that. That's why people go to bar. That's why people go to open AI or whatever it is, right? But like, if you can just pay a tiny little bit of money, well, maybe I don't need to go all the way up to the top, I can just go there. And so there's sort of like, I think that one of the cool things Bitcoin enables with AI is that you can have people running different size models, and I can cascade my way up for them. And I'm not having to get locked in to these different credit card levels while I'm doing it. I can just make quick payments to each one of them. It's like, Hey, this one costs 0.0001 cents. So I'm going to pay for that one. But I actually need something a little bit better, but I don't need a full dollar. And so I'll go there. But now because of that, anybody with access to GPUs anywhere, right, will be able to run the things that they use personally and also sell it publicly. And so that's like a really cool change that's going to happen with Bitcoin and AI that wouldn't be possible with dollars. And so that's something super exciting of like a way that you get Bitcoin more widely adopted, because it's something unique, Bitcoin uniquely solves that can't be done with dollars. And so just as many of those things that we're going to kind of come out, those are going to be super important for like the adoption of Bitcoin. Yeah. And I guess for me, I think one thing was pointing out that like a lot of things that we talked about just now, you know, the whole L4 or two thing, the Lang chain Bitcoin, kind of like agents using Bitcoin, things like that, all this stuff is out right now. It's all open source. You can use all these open models as well. So I definitely call on any developers or kind of like, you know, Tinkerer snicking, you can basically use this stuff now and charge to basically start to make the future basically. And at this point, like we're saying, the playing field is just so level. And even if you don't necessarily have a GPU yourself, you know, you can basically, uh, you know, just acquire this stuff, uh, you know, kind of like tit for tat, you know, for two itself. And I think one of the cool thing is that like, you know, you can put this proxy, uh, you know, aperture, which is on our get help, get up, get up, get up, get up, get up, get up, get up, get up in front of any API, right? So if you want to kind of like have an agent be able to hit any of these APIs or use something programmatically, you can just do this over any API proxy. So it's a really cool thing in terms of like being able to just drop this infrastructure in and gaining all of this new functionality, you know, today at the ready.

[Transcript] FYI - For Your Innovation / Why Bitcoin is the Currency for AI

So that's why it's so exciting that we can continue to have all these really cool combinations and build, you know, what we think is going to be the future with all these open source technologies. Open source is just so magical. Uh, that's one of the things that got me into the coin in the first place. I'll give one kind of macro analogy that I like to use to answer your question and also tie back to something Paul mentioned at the very beginning of this conversation. You mentioned Henry Ford talking about specialization of labor. Another analogy we like to use, um, in the early 1900s is the advent of the assembly line, which dramatically increased productivity in physical labor, made physical labor much more productive. We're able to create more cars, we're able to produce more physical goods, and people worked less. The average hours worked declined at a faster rate during those say 20 years than throughout the rest of the 1900s. I think we're heading for a very similar period where generative AI is serving as the assembly line for knowledge workers, where it's accelerating knowledge work for the first time, where we're able to all produce dramatically more. And the average hours we work may go down just like it did with the advent of the assembly line in the early 1900s. But the thing we haven't talked about that always comes up in these conversations is will there be mass unemployment? And that's a common refrain when there's new automation tools throughout history. And we think just like with the assembly line, no, there won't be. If you look at the data from that time period, uh, the actual unemployment rate went down, meaning people were more employed. And GDP per capita went up all while we were working less hours per day on average. And I think you could see the same thing happen over the rest of this decade with generative AI, where we are able to produce more and maybe have a more relaxed work schedule all while more people are employed, because this technology can be distributed globally to tie in the Bitcoin component. I always have aha moments when we're talking about convergences. I actually had an aha moment around our first Bitcoin brainstorm as well. I guess what has hit me here, it's something Paul said at the end here, we're going to see, I think you said clash or the collision between, you know, the infinity associated with AI and the scarcity associated with Bitcoin. And so what does that mean? And here's where I come back to my economic roots, which is, wow, you could have a parabolic move in the price of Bitcoin. And I think that will happen. I mean, we're on record saying that. But it also brings me back to, okay, global monetary system, means of exchange, unit of account. And so I know we're going to have one of these brainstorms with Art Laffer, and he's going to be poking at all of this and talking about the quantity rule versus the price rule and, you know, and really wanting to understand the cryptography and everything. So I'm really happy we're doing these brainstorms. So thank you. Absolutely. And thank you to everybody. I know we touched on a number of topics. We touched on Noster, Langchain, Agents and so on. So if you're listening all the way to this point, be sure to check out the show notes. We're going to have links to a number of great resources for you to dig in even further. Again, I want to just thank all the guests for joining us. And hopefully you tune in next month when we do this again. Thank you.

ARK believes that the information presented is accurate and was obtained from sources that ARK believes to be reliable. However, ARK does not guarantee the accuracy or completeness of any information. And such information may be subject to change without notice from ARK. Historical results are not indications of future results. Certain of the statements contained in this podcast may be statements of future expectations and other forward looking statements that are

[Transcript] FYI - For Your Innovation / Why Bitcoin is the Currency for AI

based on ARK's current views and assumptions and involve known unknown risks and uncertainties that could cause actual results, performance or events that differ materially from those expressed or implied in such statements.