EPISODE 545

[INTRODUCTION]

[0:00:00.3] JM: Data streams about the weather can be used to predict how soybean futures are going to change in price. Satellite data streams can take pictures of the number of cars on the road and judge how traffic patterns are changing. Search engines can aggregate data from different queries and determine what people are most interested. Data streams define how the world is changing over time. Technology companies process these data streams and make decisions based on that stream. The most direct example of this might be financial trading companies, which use all kinds of data streams to predict economic price changes.

When Henry Pihkala worked on algorithmic trading systems, he saw how useful these data streams are, and he decided to build products around data streaming. Eventually, Henry started working on Streamr a platform for data streams to be bought and sold on top of the Ethereum network. Streamr is an adaptation of technology that Henry worked on before he started working on the decentralized version. The original technology is a user-interface for connecting data streams and building applications on top of them and he acquired several customers for that platform when it was the original centralized platform version of what Streamr eventually became. But today, the vision for Streamr is something decentralized. However, the application infrastructure is still mostly centralized. Henry and his team are working on building out the decentralized version.

Streamr has raised an ICO worth around €25 million. Most startups would not raise this amount of money before series B, much less before they have a product with a large user base. In this episode Henry discusses why they raised so much money and explains why ICO's are different than equity raises. The investors who participated in the Streamr ICO received the data coin token. Henry also explained why it makes sense for this ecosystem to have its own token.

[SPONSOR MESSAGE]

[0:02:10.7] JM: We are running an experiment to find out if software engineering daily listeners are above average engineers. At triplebyte.com/sedaily, you can take a quiz to help us gather

data. I took the quiz and it covered a wide range of topics; general programming ability, a little security, a little system design. It was a nice short test to measure how my practical engineering skills have changed since I started this podcast. I will admit, although I've gotten better at talking about software engineering, I have definitely gotten worse at actually writing code and doing software engineering myself.

But if you want to check out that quiz yourself and help us gather data, you can take that quiz at triplebyte.com/sedaily and in a few weeks we're going to take a look at the results and we're to find out if SE Daily listeners are above average. And if you're looking for a job, Triplebyte is a great place to start your search, fast tracking you at hundreds of top tech companies. Triplebyte takes engineers seriously and does not waste their time. I recommend checking it out at tripledbyte.com/sedaily. That's triplebyte.com/sedaily.

Thank you Triplebyte for being a sponsor.

[INTERVIEW]

[0:03:41.2] JM: Henry Pihkala, you are the CEO of Streamr. Welcome to Software Engineering Daily.

[0:03:45.7] HP: Thanks, Jeff. Happy to be here.

[0:03:47.8] JM: So before the show we were just talking about music production, and it's funny because I think you are the second or third person who I've talked to in the Ethereum ecosystem quiz who is running a company built on top of Ethereum, but also is a hobbyist, electronic, musician. The other one that comes to mind is Luis Cuende, who runs Aragon. I don't know. It's interesting that there seems to be some overlap between the musical creativity and the Ethereum-based creativity.

[0:04:19.1] HP: Yeah, I think it's something that's been known for quite some time, that music and somehow math go together, tend to be found in the same people. But I don't know if that relates as far as blockchains and Ethereum and so on. But that's an interesting notion.

[0:04:37.1] JM: You also used to work in high-frequency trading. So I worked briefly at an options trading place right out of school. So I wrote some code for an options trading place, and I found that there was actually a lot of musicians there too. I think there definitely is this love of systems and mathematics, and creativity that I guess you can find an applied area for that in both music and in trading. I mean, do find that to be the case in trading as well?

[0:05:10.5] HP: Yeah, perhaps. I mean, in trading and in tech general and programming and maybe math and in music, like all of these things you kind of get to witness the results of your own work in quiet a concrete way, maybe. It's a bit like handicraft or some artisan stuff. You do things with your hands and your brain and get something real out of it. In trading it maybe appears as just profits are or something, like applying a cool algorithm or machine learning or something. In building software, it appears as beautiful, innovative UIs and backends. On the blockchain, it appears as new possibilities and inspiring ideas of a future decentralized world. So all of these are maybe somehow connected.

[0:06:05.5] JM: How did you make your way from the high-frequency trading world into working on blockchain related stuff?

[0:06:13.2] HP: Well it's actually quite straightforward. I mean, I've always been working with real-time data, all the way from finance. Then from algorithmic trading, I went to build a kind of a more generic real-time data platform which was kind of the Streamr 1.0 version and run in the cloud based on centralized technology, building solutions on top of that for some enterprises and startups and other companies.

Then some time, one and a half years ago, I had my second kind of — I saw the light for cryptos for the second time. I mean, I got interested about cryptos back in 2011 back when we were doing the trading thing with my co-fonder Nick, who was also a cofounder at Streamr, and we got super excited about Bitcoin, but it was so early that we didn't go all in on that and instead decided to continue on our trading efforts there, but not in crypto though but rather in traditional stock markets. We kind of left the crypto out for quite a while, but then we got excited about it again, thanks to Ethereum and Smart Contracts and the kind of ideas about decentralizing the way on building new infrastructure for application development and all the promises that it

brings with it, and we kind of saw a possibility there to apply the background that we had in real-time data and analytics and data monetization in a new way in this decentralized space.

We began to think like what would the package look like? What would be the offering? What should we build? What use cases would it supports? Who would be interested And so on, and then we kind of embarked on the new, very interesting, very inspiring route there.

[0:08:15.7] JM: So I think I'm getting an understanding for the history for how you came to Streamr. So what you're working on Streamr is a platform for data streams, for people to buy and sell data streams in a decentralized economy built on Ethereum, and it sounds like the way that you got there is you were in finance, and in finance there's all these different data streams. So you've got like people that have weather data that's available, and then you have traders that want to trade on that weather data. May be they want to buy or sell corn futures or soy futures based on that data. As well as all kinds of other proprietary data streams that are hard to acquire and hard to link up to a high speed connection too.

So I can imagine how you ended up in building a centralized platform. Is that the trajectory that you took to get to that centralized platform where you're connecting different data streams?

[0:09:13.3] HP: Yeah, it is. For sure, the one thing that I learned in trading is that real-time data can have huge value depending on how it gets used and how it gets combined with other data and how it gets made available to those people that find it useful and valuable in their own fields. From that idea, we start to build that centralized platform for data distribution and analytics. That's also the starting point that we have today. So instead of starting from scratch, we started with the centralized platform, and our roadmap consists of kind of iteratively decentralizing the different layers and components that take part in that system. So we are happy about that, and that's one strategy to take that you can — From day one, you can deliver something that's functional into the community instead of raising money and then disappearing for three years into a cave to build some software, and then maybe emerge with something that works. So we find this to be a very productive route and then allows us to have meaningful conversations with partners, data providers and our users, of course.

[0:10:33.2] JM: So your centralized platform that you started before Ethereum, you're taking that and incrementally decentralizing it.

[0:10:42.8] HP: Exactly. Exactly. Of course, much of that will need to be rebuilt from scratch to support the kind of decentralized use cases, but since it's a modular and layered architecture, we can do that bit by bit and build kind of in between versions which are hybrid, and hybrid infrastructure where part of the data is centralized and part of the data is on the blockchain and we can have things like permissions, access control, payments happening on the block chain while still using parts of the centralized stack to deliver excellent user experiences and so on.

[0:11:24.2] JM: In the centralized version that you have today, who are the people that are exchanging dollars and data? Who are the kinds of customers and data providers that are making exchanges?

[0:11:34.9] HP: So the marketplace is something that we are currently building, so nobody just yet. What we have currently up and running is the data delivery infrastructure, which is basically like massively scalable publish subscribe messaging, and we have our analytics engine that is also based on centralized technology and employs this kind of visual programming environment for prototyping and getting your hands dirty with the data enabling people to build prototypes, build automation on top of real-time data, connect data to smart contracts and build interactions with the Ethereum blockchain and so on.

So it's mainly the data traffic that's happening on the platform. At the moment it's for mainly internal purposes of the users, and we're currently setting up the marketplace. We've said that we will launch it in production this year, and actually the work is quite advanced already. So that will be an interesting milestone to reach for sure.

[0:12:38.8] JM: In the centralized model that you started with, when you went from finance to building this centralized — This cloud infrastructure for people to use data streams, what was that product? Did that product have customers or were you really early on when you pivoted to focusing on making it decentralized?

[0:13:00.3] HP: It wasn't early on, it wasn't established either. We had a pretty long RND sprint there, and actually the roots of that software already go back to the algorithmic trading stuff. So it's kind of been in the works for a long time. The productization, we started in 2014. We had major customers in there, but of course we kind of pivoted in a crucial moment, and I think that was a good move to make, because we were mostly building solutions. Like we were the ones to implement solutions to our customers and we hadn't yet reached the kind of stage of productization that we could have others take our platform, build something cool on top of that for their customers. So that would've been the next step.

But then again, the promise of decentralization was so strong, it was so inspiring that we suddenly got this huge drive to go there and try to build something that's actually not just a tool and competing with the existing cloud services, but actually build something that's completely new in the world and disruptive enables new business cases and new data monetization patterns for everyone. So we kind of just ditched the idea of going with the centralized solution and rather escaped two squares ahead to the future and started to build a future version, aStreamr 2.0 if you will.

[0:14:28.7] JM: Okay. So I think now I'm understanding the full timeline. So you started building the technology behind Streamr when you were at an algorithmic trading company, and now it's just technology to visualize and connected data streams to each other and to act on those data streams with trades. Then you said, "Let's productize this and offer it to, perhaps, other trading companies." Then other trading companies can link together data streams and use them, but it wasn't exactly this marketplace for data streams. It was really just a UI and a way to connect data streams to each other and act on them. Is that accurate?

[0:15:08.4] HP: Yeah, but not just for trading. I mean, in trading, we saw that the similar kind of pattern will present itself in other business verticals as well. Like in trading, you ingest data from the stock market then apply some model that makes decisions about buying and selling stocks. But all the fields where you can get this kind of real-time data, especially thanks to IoT with the sudden explosion of measurement devices being spread out into the world, suddenly companies, organizations, even individuals have such a huge amount of real-time data at their disposal that they can suddenly build automation if they're given the right tools.

One thing that's always interested me personally is how to bring the data closer to people who are the kind of domain experts and who know what the data means and kind of extract meaningful information out of the data. That's why we built the visual programming environment to make it easier for these domain experts to find value in the data and build automation.

[0:16:16.9] JM: I totally understand the vision that you're going for. You want a marketplace for data streams for people to be able to publish data streams and purchase subscriptions to those data streams and to be connected to one another, and you've got this nice UI where people could build those connections. One question I have is why does this need to be decentralized? Why couldn't you just have a marketplace for data streams running on AWS, for example?

[0:16:49.3] HP: Sure. We could, for sure, and perhaps there even are some, but that's not as good as selling point, because there are trust issues. Centralized marketplaces always take a cut on the sales that's happening there. We wouldn't be any better than if Google or Facebook or whoever, or Amazon themselves set up a data marketplace, which they could.

Currently they are getting most of the data in the world one way or another, and that's kind of disturbing that a handful of these large companies get most of the world's data. Only by building a decentralized infrastructure and a decentralized marketplace can we kind of break free from this pattern of some single entity having control on all the data that machines, organizations, applications and people produce.

That also makes the system kind of work without us. I mean, we are a startup and we face this problem back in the old line of business when we were doing kind of cloud analytics, that big enterprise don't play well together with small startups, because they have the vendor risks there, the small startup might just disappear and then they're left with nothing. Instead, building this decentralized technology and infrastructure which is open source. It's not run by us, but it's eventually run by the community. This makes it immune to the vendor risk to some extent. There's no obstacle for even the big boys to jump in and participate in that ecosystem, because there's less risk involved. Plus, the network effects of having a decentralized marketplace are in order of magnitude better than a centralized marketplace. Because, I mean, if Google had a data marketplace, who would go out there and tell their friends and rave about it about the Google marketplace? But if there's like a shared marketplace, that's not really Streamr's

marketplace, even though we are developing it, but it's the community's marketplace. So the participants in that ecosystem will be incentivized to make it better and help make it better and provide data in there and consume data in there. So I think there's an advantage as well.

[SPONSOR MESSAGE]

[0:19:29.4] JM: If you love Software Engineering Daily, I think you'll love the Google Cloud Platform Podcast. It's a podcast about Google Cloud products, how they're built and how you can use them. Really, it's all about the changes that are going on in software engineering as told from the point of view of Google engineers.

You can find the Google Cloud Platform Podcast at gcppodcast.com. You'll hear from Google is like Vince Cerf and all about TensorFlow and Firebase and BigQuery from the high-level use cases to the low-level implementation. The GCP Podcast has all of that covered.

Find the Google Cloud Platform Podcast at gcppodcast.com and subscribe to it wherever you get your podcasts.

[INTERVIEW CONTINUED]

[0:20:21.0] JM: Okay. I'm a little confused there, because — So let's take GitHub, For example. GitHub hub is a centralized place where people publish open-source code, and I have an incentive to make the GitHub ecosystem better, because I derive value from the GitHub ecosystem. So GitHub has network effects.

[0:20:41.9] HP: It does, of course.

[0:20:43.7] JM: So if somebody were to tell me, "Hey, I want to build GitHub on the Ethereum blockchain." I would say to them, "I mean, that's great, because you can't censor me," for example. I guess that would be an argument to decentralize the underlying infrastructure. But as far as the use case itself, GitHub on centralized infrastructure does a pretty good job and it also seems like a data stream economy on top of centralized infrastructure. It's not clear to me why that running on decentralized infrastructure, that running on Ethereum is that much better unless

you're talking about sonsorable data streams. I guess what I understand is if you you're on this trajectory to gradually move from centralization to decentralization, why not first prove out the data stream marketplace on centralized infrastructure? Because there's nothing — Like I don't know of any centralized infrastructure data stream economy other than there are specific companies that have specific APIs, like the Yelp API. You consider the Yelp API a stream that is on a marketplace kind of. Yeah, why not try to build this on a centralized infrastructure first to prove that it's actually something that people want?

[0:22:06.5] HP: That's indeed what we are doing. So we'll first launch the marketplace that will be based on a hybrid infrastructure. So we are able to leverage the existing network that we have to build the marketplace, but we can use the blockchain for certain things, such as identity and payments, because they are far superior to what exists in the kind of centralized world. Instead of inputting your credit card details, you can just press a button to make a transaction, and that will be great.

But that is what we are doing, and building the decentralized infrastructure will be a long-term project. It's not easy by any means, and it will take a few years to accomplish. So building applications on top, improving the use cases, starting to build the kind of real-time data economy and community around it will be possible much sooner by leveraging the centralized technology in there.

Regarding your GitHub example, I think that's a very good comparison there. There's two issues that you can you find in their given that they have indeed successfully built a working ecosystem in there. But one thing is do you trust GitHub? Probably you do, but how can you really know that GitHub doesn't like look at your private repositories or whatever? So you're placing a huge amounts of trust in GitHub, and thanks to their good PR, they have profiled themselves as a trustworthy company that doesn't do this things. Maybe if the repositories were run by — I don't know, Facebook or whoever, you might have some issues in placing your code in there. This is, to some extent, sold by the decentralized technology.

Another thing there is how do you — Can you have an economic incentive on GitHub? Can you actually make — For example, can you have a stake in GitHub? If it's a publicly traded company, then of course you can. But if it's not, then it's very hard for the average user to like a

product and have some kind of economic stake in that ecosystem. That's almost impossible. But thanks to the kind of tokenization all of new decentralized applications, this is also possible to those interested in participating in the ecosystem in that way.

So you have the token, which represents obviously the value in the data itself as it's being used as a payments, means of payment on the marketplace, but it also represents the value in the overall ecosystem.

[0:24:57.7] JM: The GitHub trust example, I mean, you could build a version of GitHub that runs on centralized infrastructure that would have assurances of trust.

[0:25:12.7] HP: Sure. It could happen to end-to-end encryption, for example, happening. Yeah.

[0:25:14.6] JM: Yeah, exactly. So you don't necessarily need decentralized technology to build a transparent company, right?

[0:25:24.3] HP: Yeah, for sure. For sure. Luckily, transparent companies have been built before, and hopefully will be built in the future as well. So I'm not saying that. I'm saying maybe that the decentralization gives a new opportunity to build projects that are trustless and transparent by default. So traditional companies need to go through a whole lot of efforts to become transparent and to do their PR correctly to appear as transparent and lovable by their users, and instead the decentralized products somehow have a different starting point. They start from open source. They start from transparency. They start from trying to benefit the community and not just turn a profit, which is eventually what all the traditional companies are looking to do. So there's some kind of fundamental differences there, but sure, a similar end result, if you accept the need to trust, can be established for sure.

[0:26:37.3] JM: Okay. You mentioned the identity and payments system on top of Ethereum being better than the centralized alternatives. Can you contrast the centralized options for identity and payments with those available on Ethereum?

[0:26:55.1] HP: Sure. Of course, here I might have to make the note that this is very young and immature technology when we're speaking about the blockchain. I think one place where it falls

short today is the usability side. So we can truly compare the usability of paying on the blockchain versus paying on the credit card, because it's not really a fair comparison in that sense, because the other one has decades of history and development behind it and the other one is brand-new. But the promise is amazing. You can make payments without any middlemen directly from one, say, data producer to a a data consumer. No setting up of APIs to ingest the data or take in the credit card payments, taking care about your credit card details, security when it's stored somewhere, and so on.

On the blockchain, you don't have these problems. The ideal use case is that you go on a marketplace, you have a buy button there, you press that and you get the product, and this is achievable with centralized technology. For example, you go on to the app store and you see some apps and you press buy and everything just happens. But this is due to the various middlemen in there, operating the system and getting a cut, like Apple takes 30% of the app sales, then there's the credit card company that takes a few percent and all of these chain continues on and on. Now we can have direct interaction between the buyers, which also means they don't even need to be human beings. They can be machines. Like there could be a machine that's producing some data, like a car, say, driving around and measuring something, like the road condition or cell network, signal strength or something like this and selling that data to other interested parties, like the cell stations or the network operator or the smart city that can automatically detect the condition if some roads are in bad condition or things like this.

This can happen automatically and give rise to data economy of machines, because there is no longer the need for the human middleman or the human credit card holder to enter those details. It's rather based on a private key, which is your identity and your wallets and your value you, and you can use that to interact with the system and other machines, or people, or organizations, or whoever are participating in that ecosystem. It kind of brings down the barrier of automation and makes it somehow more direct for everyone. Does it make sense?

[0:29:52.3] JM: Well, kind of. So a few counter arguments there. So in terms of data itself. So if I wanted to purchase data about cars that are driving around, I think I would want to purchase that data from Lyft or from Waymo or from Uber, and it doesn't seem problem — Like if they're willing to share it on a centralized platform — First of all, that doesn't seem problematic, because if like let's say there are trading companies that want to monetize that data stream, that

doesn't infringe on Uber or Waymo's profit opportunity. So again, like if there were a desire to have access to that kind of data stream, I'm still not quite convinced that this couldn't exist on centralized infrastructure.

But just to also go in on your payments point — So in the vision that you've portrayed, you talk negatively about middlemen, but one of the advantages of all the middlemen in the traditional payment system is that if somebody defrauds you as a customer, there is recourse. So if somebody were to go on Streamr, under the Streamr economy and to start to publish data that said, "Hey, this is data about a collection of soybean farms around the world." You can consume that data and you can trade on it, and if I start consuming that data stream and I start trading on it and then I'm like, "Why are my trades all going the wrong direction?" and then I inspect the data further and I compare it to some other data sources, I'm like, "This data is totally bonk. None of it makes sense. None of it is accurate." If I were doing that on a centralized payment system, I would be able to get a refund. But if I'm doing that on the Ethereum blockchain, well, sorry, there's no refunds on Ethereum. Maybe there will be infrastructure for that someday, but then we're basically — First of all, that's really far off, and second of all, that starts to look a lot more like traditional baking infrastructures. It's not clear to me that the payment system for this kind of use case is actually superior. I actually think this is a use case where you want middlemen.

[0:32:03.4] HP: Yeah, in some sense, but — Okay. So the new technology enables directness, but it doesn't mean that it has to be direct. It just enables that. So there can still be opportunities like insurances or something like this can be built on top of that, kind of adding a middleman which can have like varying levels of trustlessness happening on there. It's like a different kind of ground to build in.

In the traditional systems you're limited by having to trust somebody and not take the responsibility so much on yourself. In the blockchain world, the kind of default starting point is that all the responsibility is on you, and if something bad happens, you lose the money. But it doesn't need to be like that. We can build mechanisms that can be trustless but still guarantee you that if you get screwed over, you can get a refund. Data producers, they can place a steak to guarantee the quality of the data, and if they commit a fraud, they will lose that stake. Of course, the kind of dispute resolution is always a complicated thing to implement, but there are

some patterns will work there. As I said, there can be trust providers or kind of middlemen who are basing on this trustless technology to implement the good parts of what we have in the centralized economy.

So I do recognize the problems there, but as we know, it's quite immature at the moment and it will probably look quite different in, say, 5 to 10 years. Most of these problems will probably be overcome in one way or another.

[0:33:57.3] JM: So maybe this platform does exist on centralized technology. I mean, I guess there are data providers that will provide you with a high speed data feed of information about soybean farms or cars driving around the world.

[0:34:11.3] HP: Sure. These are all silos.

[0:34:15.0] JM: They're all silos. Okay.

[0:34:16.1] HP: Yeah. The providers need to set up their own systems and try to attract the consumers of the data to their little silo. So the kind of threshold of starting to provide data is large, because you have to set up the infrastructure to do so, plus the marketing efforts to attract users for your data are considerable as well, as opposed to a kind of central but not centralized, but central place where people could get all the data that they need and easily also provide data. This makes providing data easy enough even for individuals. Individuals would never set up an API and credit card payments to sell their, for example, personal data, or health data, or location data, whatever. But this kind of makes it possible. So there could be an application ecosystem that measures something about its users, enables them to earn on that.

So today, like my database going to Google and Facebook and I'm not getting money out of it. I'm getting maybe an application that I can use, but Google and Facebook and the other big ones, they are the ones who are able to monetize the data, because they have operations of size large enough to overcome these thresholds and they're not even sharing that data. Usually, your question earlier, about why wouldn't you buy that data from ways or whoever. They tend not to share the data, because they get some business advantage out of keeping the data in their own silo and using it internally. What we are super interested in is how to enable the

transformation there. How to make it profitable for those who collect data to find the external use and value of the data, which is currently unknown.

Companies collect data in their internal silos, but by doing so they also place a lid, they place a cap on the value that can be extracted out of that data, and the volume outside that silo is completely unknown, and it's hard to reach without overcoming those thresholds.

[SPONSOR MESSAGE]

[0:36:42.4] JM: Software Engineering Daily is brought to you by Consensys. Do you think blockchain technology is only used for cryptocurrency? Think again. Consensys develops tools and infrastructure to enable a decentralized future built on Ethereum; the most advanced blockchain development platform.

Consensys has hundreds of Web3 developers that are building decentralized applications focusing on world changing ideas, like creating a system for self-sovereign identity, managing supply chains, developing a more efficient electricity provider, and much more. So, listeners, why continue to build the internet of today when you can build the internet of the future on the blockchain?

Consensys is actively hiring talented software developers to help build the decentralized web. Learn more about Consensys projects and open-source jobs at consensys.net/sedaily. That's C-O-N-S-E-N-S-Y-S.net/sedaily. Consensys.net/sedaily.

Thanks again, Consensys.

[INTERVIEW CONTINUED]

[0:37:58.9] JM: So are you suggesting that individuals are going to share their data on Streamr, because they are going to be paid monetary rewards in contrast to how they share their data with Facebook and Google and get a search engine or a social network?

[0:38:14.6] HP: Yeah, absolutely. It's, of course, not sure what the killer applications will be. It's probably not going to be a search engine or a social media, but imagine that you could earn money by driving around in your car. I mean, typically, that costs you money because you have to pay for the gas. But perhaps the rewards — Like given that a data economy that's established enough, you could have, for example, rental cars that are free to use, because the costs of renting the car are offset by the data that you can produce by going around and collecting all the data that's out there in the world.

[0:38:55.4] JM: But again, we've had the technology to do that for probably a decade or more, and I have not seen a data stream economy come up. So why would it happen now that it's on Ethereum?

[0:39:08.9] HP: I don't think it's about Ethereum, it's about how the kind of mind shift needs to happen to words sharing data and enabling the direct monetization there. It simply too difficult in the old world to enable people to earn like tenths of a dollar because there's overhead in every transaction that's being made there. Of course, the same applies today to Ethereum, the kind of transaction costs. There are relatively high, but this is something that will be solved in the future for sure.

We can go down to smaller and smaller payments, micro-payments and this will also enable people to monetize their data in a more kind of fine-grained way. Today, data is only valuable if it's large. If the ecosystem is larger, or like you run a stock market and you're selling that data to traders. That's like a huge, that's like an ocean liner, but we can enable like little boats to floats on the same data ocean as well.

[0:40:19.5] JM: To be clear, you're saying that you think that the reason that my Fitbit does not publish data to a publicly accessible data stream, the reason that my car does not publish data to a publicly accessible data stream, it is because of transaction costs. It's because of a mindset issue. Is that what you're saying?

[0:40:40.9] HP: I'm saying that probably the reason is because they want to lock you in into that ecosystem. So your Fitbit is not sharing the data or making it easily accessible to outside, because they want you to buy another Fitbit product and kind of stay in the family. But let's say

you had two competing risk computers. You have Fitbits, which doesn't let you share the data or monetize data, and there's another wristband that earns you $3 just for wearing it. So which one would you choose?

The centralized approach is a kind of losing thought there, because it can be overrun by those who actually enable the end users to monetize their own data. This is what I believe in. But it will take some time for these products to appear for the technology to be mature enough. The business models will change, but kind of keeping things secret cannot — It's simply cannot be the winning solution, like in 100 years, it has vanished off this earth, and this is my prediction. I hope I live to see it.

[0:41:54.8] JM: Yeah. Well, it makes sense to have the choice. I mean, to have the — Why wouldn't you have some way where you have the choice to share your data and receive payment, direct payment for it, rather than to receive a service that is making money through advertising. I would agree with you that that experiment has not really been run. It's interesting.

[0:42:16.3] HP: It has been. Wouldn't that be superior? If the product is the same, they both look nice and they have the same features and so on. So if it comes down to that selection, the other one enables you to earn money. That's amazing.

[0:42:31.7] JM: I mean, that's not clear to me. That's not clear to me, but I then again I'm kind of a fan of being surveilled and being served useful ads, and I think for me that's less of a big deal than it is for a lot of people. But that's a totally different conversation.

[0:42:44.4] HP: Sure.

[0:42:46.3] JM: Let's talk about ICO's. So the Streamr ICO raised about €25 million through the ICO. Can you describe the ICO process?

[0:42:56.7] HP: Sure. It was quite an effort for sure. So it started out with the kind of idea. Most of last year, we spent on thinking about this stuff. We made a whitepaper that came out in May last year. Before that, we had done our kind of first public pitch of the idea at AdCON in February in Paris, and we got a lot of positive feedback on that and decided to move forward.

So whitepaper in May, then we had several stages of the actual token launch. So there were some private product presales or pre-contributions for the token launch. Then we had a public pre-contribution around in September. Finally, the crowd contribution around in October last year and ended up, as you said, raising 30 million Swiss Francs, which is approximately equivalent to 25, 26 million euros. It involved a lot of community building, just getting feedback, getting users, crystallizing the idea and make everything smoother and trying to find the right track for the project.

[0:44:19.2] JM: Yeah. The coin is called the Datacoin. How does the Datacoin token fit into the Streamr economy? What is the Datacoin token do?

[0:44:29.6] HP: Sure. So it has two main purposes there. So it acts as a means of payment on the marketplace. So you can have data streams for which licenses can be bought using this Datacoin token. It's also essential usage element on the Streamr network. Because we have decentralization and there are community members who run the nodes, there's an incentivization structure which is kind of comparable to mining in Bitcoit, for example, or Ethereum. But instead of solving this artificial and hashing problem in there, in our network, the people who run the nodes, they contribute bandwidth and storage on to the Streamr network to produce the delivery service of data on the network. So they're doing work. They're contributing resources to the network and they're able to earn the token in reward, rewards for doing so.

[0:45:36.5] JM: Why do you need to issue token for that? Why not just have Ethereum smart contracts facilitating the logic of this system and use ether to pay for it?

[0:45:52.3] HP: Sure. That's a great question and an important one to get asked. The first reason is that by using our own token, we can separate the kind of holders and people who use the token from other users on the Ethereum network. So if we used ether, we can we can't really track which ether is a part of our ecosystem and community and which is not. So they kind of get mixed in there, which is okay. But for some purposes, for example, we are building this reputation system into our network which enables the more reputable notes to get more responsibility in the network or on the marketplace. The more reputable sellers can gain more

visibility and appear to be more trustworthy by participating in the token ecosystem. So we can track things like how much data they earned or spent.

[0:46:52.8] JM: Well, but every smart contract has a database associated with it. You could just keep track of who is paying what and when. You don't necessarily need a token to track people.

[0:47:03.3] HP: Sure, when it gets transferred, then it gets mixed up. When it goes to centralized places, like exchange, then the trace gets lost. So the token enables us to do a bit more — It allows us a bit more flexibility. It also has features that ether doesn't have. We can have like allowances granted to smart contracts. Technically, there are some differences.

Of course, it also enables the project to fund itself, and there's no denying that in that. So that's why the ICO is also important and a magnificent new way to fund open source projects, which so far have been dependent on donations or doing some commercial offerings on the side. So that's completely disruptive. I think the kind of app tokens will find their place as well. In our ecosystem, it just enables more flexibility compared and more kind of clear-cut control. We could have like plasma side chains operating, the data token ecosystem or micro-payments by keeping it kind of separated in there and stuff like that in the future.

It's a more future-proof approach. It gives us more flexibility today. Plus, it allows us to fund the project, which is kind of a starting point, without which it would happen, right?

[0:48:35.5] JM: Yeah. I mean, I think the fundraising aspect of it, it makes complete sense. And even if you could just use ether plus smart contracts to basically — You have colored coins on Ethereum that facilitate all of your functionality. I do think it is defensible to say, actually, even though we could do it in an equivalent fashion, we're going to issue a token because we want to raise some money and we want to have people that have a stake in the system. We want to have our own token. I think that's justifiable, and then we get to the question of why are these ICO's — Why do they have such a premium associated with them compared to companies that are at an equivalent stage that are raising a series A, or series B, or seed round. There's clearly a premium associated with these tokens. One argument for that premium is that you have a liquidity premium. When you buy a token, the token is more liquid and therefore when you put

— When you invest in a token, it makes sense for you to be paying more for that token, because you can exit from that earlier. There is more liquidity associated with it.

So then we get to the question of; is the current market liquidity premium, or at least the liquidity premium when Streamr specifically raised, was it appropriate given the stage of the company? So like most companies don't raise $25 million or €25 million until they're at a stage where they need to scale dramatically. Like $25 million like a series B or a series C round in traditional companies, and you're raising that without really having customers, without having users.

[0:50:23.5] HP: That's not really true. I mean, we did have customers and users. If you remember, we have a working product to start with. So we are not really just starting with a whitepaper, an idea. We have something that's already halfway productized and being used out there. So maybe that's one differentiating factor for us, but I do agree with you that the market is probably in some kind of hype cycle. The liquidity premium is one component there, but another thing to consider is that the ICO project, they only get to raise funds once, basically. A typical startup goes through a range of fundraising round starting from angel rounds, then going to series, A, B, C. They can kind of turn the boats as they go and decide how much money they need at each stage. So they get more visibility into what's happening.

In some sense, the ICO projects, they raise all the money they're going to need upfront, which is rather difficult. So there could be — I think a fruitful approach —

[0:51:37.4] JM: You could raise equity rounds, right? You could still do a classic series A and series B, because that's a different pool of equity than the token.

[0:51:47.0] HP: You could, but what is the equity worth. I mean, why would it be worth anything in this case? So it's interesting how the value of the company or the ecosystem is kind of in the token and not the actual stock of the company, and it's unclear what the stock might be worth. So it's much harder for the companies to raise equity funding. Plus, not all of them are companies in the first place. They are their foundations. They are dows. So the kind of pool of entities operating in this space is much more diverse than the traditional kind of of startup community.

[0:52:23.8] JM: Okay. So just to zoom in on the question of customers, like how many customers you have and who those customers are. From my understanding of what you said earlier, the customers that you have, what they are currently purchasing or what they are bought in on is a way to visualize and use proprietary data streams for trading systems. That's a very different use case than what the Streamr ICO is raising on. The Streamr ICO is raising on this vision of a decentralized world where individuals are giving up their data and cars are giving up their data and so on. Am I correct here? That's a very different customer set than what you actually have today.

[0:53:06.4] HP: It is. The modern-day Streamr is not for trading. So there's a misconception there. So it's a broad generic platform for any kind of data, whether its origin is financial, IoT, social media, smart cities, whatever. So it's agnostic.

[0:53:24.8] JM: Oh! So you do you do have customers like that, that are doing smart cities and stuff on —

[0:53:28.6] HP: Yeah. Well, we used to, but before the pivots. Basically, once we started on the new track, we kind of kicked out the old customers, because we were doing consulting and solution delivery for them and that was not where our heart was. So it made more sense for us to kind of concentrate on the product building, which is what we really want to do and feel that — At least a successful path. So we kind of took a little time out from serving customers to turning the ship into a spaceship and go where we want to go.

[0:54:03.1] JM: Okay. So what's like the unlocking schedule for the money raised in the ICO and who gets the money raised in the ICO? Does that money get allocated to founders? How does it get allocated to the company? Where is that money allocated and what's the unlocking schedule or the vesting schedule for that money?

[0:54:23.7] HP: Yeah. So all of the raised funds goes to the company, and it doesn't have an enforced unlocking scheme. So it's at full discretion of the company. We have communicated how the money will be spent and some approximate time tables and we are sticking to those for sure. So we tried to go for full transparency. So if there is anything questionable about the use

of the money, then we are happy to elaborate it, for sure. We are four months into the project, almost five months, and things are moving quickly forward and exactly as planned.

[0:55:00.9] JM: Cool. So does that mean like — Do you publicize what the salaries of people in the company are or how many tokens they're getting or the vesting schedule for those tokens?

[0:55:11.3] HP: Yeah. So the token distribution is indeed known, not to the individual level, but I think that is quite easy to deduce by looking at the blockchain, like who is who, if somebody cares. It would be possible to publish like balance sheets and quarterly reports. I mean, if you think about public listed companies, they have regulations that enforce them to have a certain level of transparency into the financials, and I think that would be excellent for the crypto projects as well.

Of course, setting that up requires some resources, which is always a way from driving the project forward, but I think we do need in this space more regulation that we could be good for everyone, and more transparency, and all of these help fight fraud, for example. So the money raised cannot disappear into a black hole. There needs to be results delivered. Many projects are open source, in which case anyone interested can see what's happening, that actually things are being done. Products are being released and so on.

In our case, you can go and sign up and start using the product right away. It's up and running. It's out there. So I hope all these kind of reduce the risk of participating in the Streamr ecosystem. We welcome anyone to drop by our offices and say hi. We're actual real people here, even though that gets sometimes forgotten in the crypto community, which is somehow so anonymous and so faceless and how there's actual flesh and blood behind the code that's being written. But, luckily, the meet up community and all of that is buzzing. Here in [inaudible 0:57:11.2], we have a little crypto valley and interaction between projects, and that's great as well, one of the best reasons to be here.

[0:57:20.7] JM: Okay. Henry, thanks for coming on the show and talking about Streamr. It's been really great to learn about it.

[0:57:25.8] HP: Thanks, Jeff. Thanks for having me.

[END OF INTERVIEW]

[0:57:31.3] JM: If you are on call and you get paged at 2 a.m., are you sure you have all the data you need at your fingertips? Are you worried that you're going to be surprised by things that you missed, errors or even security vulnerabilities because you don't have the right visibility into your application? You shouldn't be worried. You have worked hard to build an amazing modern application for your customers. You've been worrying over the details and dotting every I and crossing every T. You deserve an analytics tool that was built to those same standards, an analytics tool that will be there for you when you needed the most.

Sumo Logic is a cloud native machine data analytics service that helps you run and secure your modern application. If you are feeling the pain of managing your own log, event and performance metrics data, check out sumologic.com/sedaily.

Even if you have your tools already, it's worth checking out Sumo Logic and seeing if you can leverage your data even more effectively with real-time dashboards and monitoring and improved observability. To improve the uptime of your application and keep your day-to-day run time more secure, check out sumologic.com/sedaily for a free 30-day trial of Sumo Logic.

Find out how Sumo Logic can improve your productivity and your application observability whenever you run your applications. That's sumologic.com/sedaily.

Thank you to Sumo Logic for being a sponsor of software Engineering Daily.

[END]