***Open Minds … from Creative Commons***

# Episode 19: Sam Williams of Arweave

[00:00:00] **Ony:** Welcome to Open Minds from Creative Commons. I'm Ony Anukem CC's communications and campaigns manager. Our Open Minds podcast is a series of conversations with people working on the issues we're involved with and the subject we're excited about. Each episode, we sit down with amazing guests who are working to make the internet and our global culture more open and collaborative. On this episode, CC's chief operating officer Anna Tumadóttir sits down for an interesting conversation with Sam Williams, the co-founder and CEO of Arweave, the company that created the Arweave protocol, a permanent archive of human knowledge and experiences on a blockchain. Arweave recently introduced CC licenses as its first licensing standard on the Arweave application. Sam has been immersed in open source since he was a kid and started learning to code when he was nine. So it's no surprise that now he's passionate about building innovative software and solving complex problems in computer networking and has built extensive experience in real-world mechanism design and implementation. When he's not working on Arweave, Sam actively participates in the decentralized webspace as a technical advisor and mentor of blockchain projects. Well, that's enough for me. Let's hand over to Anna and Sam.

[00:01:43] **Anna:** Welcome Sam, it is lovely to see you. So at this point, we've given our listeners a little bit of insight into who you are and what you're doing now, but I would like us to go back in time a bit, to when you were first introduced to the world of open source and what your background is there.

[00:01:57] **Sam:** Well, first of all, thank you so much for having me on. This is a really exciting opportunity. Well, I don't remember when I first encountered open source software so long ago that I think... it's one of those things that's so early in my childhood memories that, you know, I'm pretty sure I can't recognize the first time. I remember the first time I engaged with a really large piece of open source software, which was installing openSUSE, it was called, which is a Linux distribution on my first, you could barely call it a laptop by today's standards, but it was at the time a laptop with this crazy serial cable to a CD drive. And so I had to install this thing off, I think, seven different CDs that I'd burnt. And I was fascinated with Linux for a very long time. I thought it was amazing to have an operating system where all of the components were open and editable. That was the thing that got me really going about it. It's like having a playground as a computer, if you will. And that was super fun. And I must have been. And nine years old, then something like that.

[00:02:57] **Anna:** Oh, wow. So you got started early. All right. Well then it's fair then that you don't remember when you were introduced to it. Sounds like it was like I learned to read and open source was there.

[00:03:05] **Sam:** Exactly. I remember doing visual basic five was my first programming language, which was horrendous, and then a friend of mine introduced me to Python, really early version of Python. He was building video games with it, very, you know, hard to call them video games. We're pretty close to video games. Yeah. And I suppose technically that's probably the first open source software I came across.

[00:03:26] **Anna:** That's a lovely story. And I'm entertained to hear it happened when you were nine years old. So you put other nine year olds on notice. But okay. So before we dive in too deep on things, can you briefly describe Arweave, your current work, and sort of like the problems that you see that it solves, and so on.

[00:03:41] **Sam:** Yeah, a hundred percent. So, Arweave is a permanent information storage system. That is what it attempts to build. The simple problem we're trying to solve is sort of twofold. We want to build an open, irrevocable, unforgeable archive of history that is replicated in so many places on earth. It's impossible for any person to alter. It is sort of history beyond the control of any individual or group. Collectively owned history, you could call it. And on the sort of, you say corollary to that, it's not just history, but also knowledge, making knowledge available to all people to access across all time, and making it so that, you know, you could say the long tail of knowledge doesn't degrade. And so in order to build this system. So I was watching blockchains since well, since Bitcoin, obviously, but like 2010, 2011, that type of period. And I was fascinated. I mean, before that I was interested in BitTorrent, and this was this idea that you can have an amorphous system for moving information around basically. And then Bitcoin sort of, you say, improved on that, and say well, what if we have consensus as part of this system. That was kind of interesting, and then basically what these nodes were building was this ledger, right? A ledger of history. And you could in fact embed tiny amounts of information inside them and you send that transaction. And then your piece of information was replicated in, well, even in the early days, it was thousands of machines. Now, I guess, tens of thousands of machines around the world. That's pretty cool. I thought. But it wasn't until later in about 2016, I was watching the way the world was tending and I was just finishing my PhD. So I never quite finished because I started this project. But you know, I was seeing what was happening in the world and I thought, wow, well, it's possible that at least for the first time in my life, we are facing the rise of authoritarianism in the west. And I didn't wanna live in a. In a world like that. I thought that freedom and democracy were pretty special things. I profoundly believe that to be true. And so I thought, well, how about that ledger thing? Right. Like, so, so I read a lot about how authoritarian regimes develop, what is required and I realized that one of the fundamental things is that you have to gain control of the information space at some point. Oh, well, like nicely, I mean, it's fiction, but it's one of those beautiful sort of meta truths like it's fiction, but it is to some extent, more true than tales of just one thing. It's a pattern that repeats itself over and over again, expressed as a parable, if you will. And he said that basically by controlling the present, you continue the ability to control access to the past or information about the past. And subsequently you change the way people act in the future . And you know, this happened all over the world where authoritarian regimes developed. It's by no means a silver bullet. It's not gonna solve the entire problem, but I figured, ha, that's actually something, I'm pretty sure I know how to solve. And so I got a team of people together and we went about solving it. So we made basically what is a blockchain which you can fit arbitrary amounts of data inside. So just scaled so they can fit, that 160 bytes or whatever it was in Bitcoin can now be two to the power, 256 bytes. So quite a lot larger. And then we built this endowment system to pay for it, 'cause once you've solved the first problem of, you know, how do you store data? Well, you get this other problem, which is, well, you're gonna store that data forever. How you're gonna pay for it. Simple answer for us. We figured. Well, you put in enough storage cost to pay for 200 years worth of storage. And as the cost of storage declines, that 200 years expands out over and over again. And so as long as the storage cost decline rate stays about 0.5%, you end up with more storage, purchasing power at the end of any given year than you did at the start. And so we define success as basically leapfrogging this generation of technology, making it so that this data set lives until, you know, whatever some bright person, 300 years from now creates which sort of subsumes it, you know, comes along and so well, that's what we've been doing. And now, we're five years into this journey. We just celebrated the fourth anniversary of the network two weeks ago and we're coming up on somewhere over 500 million pieces of data stored now.

[00:07:54] **Anna:** That's impressive. And congratulations on the anniversary. That is a...

[00:07:58] **Sam:** Thank you.

[00:07:59] **Anna:** An exciting accomplishment. Yeah, must feel good. So it's interesting in terms of sort of where the idea came from with political systems and the rise of authoritarianism. I remember when we first met you were telling me about the pro-democracy activists in Hong Kong and they were archiving the apple daily articles. Right. Yeah. And so since then, I've seen a couple of other reports around Arweave being used to store this politically significant content? So like Russia's war on Ukraine, for instance.

[00:08:23] **Sam:** Absolutely.

[00:08:23] **Anna:** And so I'm curious to hear your perspective as to why Arweave specifically is the useful solution for storage of these materials? Like I understand like the sort of immutability and for the long haul and so on but why could they not just be stored elsewhere and be safe?

[00:08:37] **Sam:** Yeah, well, I mean, and being safe is the critical reason. So when you throw a piece of Arweave, a piece of data in Arweave that's replicated in the base sense 45 places across the world. Okay. But in the practical sense, because of the block reward right now, it's actually stored around a thousand places around the world with no centralized controller. So when we were starting out, we were looking at you. How do archives fail? The answer is essentially, the library of Alexandria problem, right? So either there's fire flood or the government comes and destroys the data or another government becomes a you know, invader essentially. Well, you know, blockchains are extremely attractive for solving this problem because they allow you to maintain ledgers of information where there is no leader anywhere in the world. And the system is highly redundant. A second component is this endowment. We don't think that you can trust a company to run such a system because inevitably the company, at some point will, either get greedy and start to try and make money off the top of the money that has been put in the endowment, or they'll just make a mistake. And so if you run it as a foundation, perhaps not a company, but that foundation is run by humans and humans make mistakes. And so we looked at this and we said, well, what if we just, would you say devolve that power to math and math alone and allow that to run that system. And we think that's much, much more viable over, you know, century use time frames rather than yeah human run institutions.

[00:10:04] **Anna:** Oh, that makes sense. And permanent storage sort of as a concept makes sense. But then, you know, I think about the user in this I'm like, what about real time retrieval? Like for our generation, for accessing the things that are happening in real time? Like, how does it work? How does a normal person take advantage of the permaweb?

[00:10:21] **Sam:** Yes. So real time retrieval is done thanks to, well I told you I was kind of a fan of BitTorrent back in the day. We basically took BitTorrent, modified it slightly to make it so that you're not just rewarded for giving. Bandwidth, but plausibly other things as well. And that system can adapt. It's kind of like an adaptation of the BitTorrent mechanism design. Yes. But that bandwidth swapping is essentially the base way that you get access to data in the Arweave network. On top of that, we have a system whereby users can come to a gateway, we call it and anyone can run one of these systems and they expose the data in people's browsers. This is really exciting. Now you have this sort of what we call the permaweb, a web of knowledge that is interconnected and doesn't get forgotten and anyone can go to a different gateway and get access to it in the browser on their mobile, wherever it happens to be. So it's almost like a modification to the base protocols of the web. HTTP allow it to have memory. So you don't really need an archival service. It is an archival service in itself. One way to think of it is if the web and the internet connected people over space. The idea of Arweave is to connect people over time, as well as space. So you put a piece of information into it, and then it goes into this sort of permanent knowledge commons, if you will. And that is a very exciting thing for both of us. I think yes. And then when, once it's there, anyone can access it in the future. You can speak to people in the future and you can read what people thought in the past. It's the basic idea.

[00:11:48] **Anna:** So these discovery mechanisms, these are gateways that you enter through existing browser technology, essentially.

[00:11:54] **Sam:** Exactly. It plugs totally into the existing systems. We think that the web is basically made up of. What we call the web is really just a stack of these technologies, and Arweave is just another layer on top of that very deep stack. So Arweave nodes actually speak to each other through HTTP, interestingly. And so you can point your browser at an Arweave node and say, Hey, get me the price of storing this many bytes. It'll just return it and it'll render in the browser. Yeah it's all kind of intermingled with the webs that currently exist today. It's just adding long term memory essentially.

[00:12:27] **Anna:** Got it. So. Imagine I'm somebody who's not very technical actually. Don't imagine that I'm not very technical. So if somebody is sharing content and it's licensed free use and their idea is that like, oh, like Anna, to be able to discover this content or people like her who are doing her kinds of work, like the mechanisms, like, is there something that makes sense to a regular person to enter where this lives and how would. How would I find what I'm looking for? Like for instance, if I wanted to look for footage of war crimes that's been released under an open license because people want to get the word out there, or I want to look at the archives of the apple data or anything else that might have been stored in this way.

[00:13:09] **Sam:** Well, because it all intermingles with the webs that currently exist, we get this strange effect that Google actually indexes it. So if you just search and you'll find the content from it yeah, that that's very helpful, but one of the things I'm so excited about with the work that we've been doing together actually is this idea that so data on the permaweb is on, on, Arweave that sort of protocol underneath it is tagged. We saw that basically when you have a ledger, when you're trying to archive things, you want to not just store the data itself, but you want to store arbitrary, like tags that you can set as well as you know, who is the person who uploaded it? When did they upload it? These types of basic pieces of information. And we realized by working with you guys that actually what we can do is we can make it so that the license data itself can be atomically associated. So sort of. Combined into that what we call it, transaction ID. So the identifier of the piece of information, such that when you transmit that piece of information, you are necessarily transmitting the license information to, and this is what's so exciting about the, well, should we jump into maybe the standard we built together?

[00:14:15] **Anna:** I was gonna say so yeah this dovetails with the fact that Arweave recently introduced the support for CC licensing and so...

[00:14:22] **Sam:** Exactly.

[00:14:23] **Anna:** And the thing that we would really like to explore is the impact that that might have on people who are, archiving knowledge or are, cultural heritage practitioners, digitizing collections within the open community. Like what does this mean for them?

[00:14:36] **Sam:** Yeah. So this means that basically when you publish a piece of information to the web, now, if you put it on the. You can put the license data itself, you can say, well, CC license you want. And so on into the upload. And then when someone uses that link, so maybe they embed your image on their website or something other people can see, oh, right. Click view, open image and new tab, and then see that identifier, put it into a block Explorer and Arweave, and they can see, oh, that's CC licensed unit. So, and which is really powerful because we hope that we can avoid with this system. Ending up in the track of the current web, which is like, everything is basically unlicensed by default. Nobody's really thinking it's through. And we're, you know, there's 500 million pieces of data on this permanent web we built. Now, but it's still very early days. So our hope is that we can sort of encourage people to set the license type of the data as we go, and then it will become the standard behavior. So we don't end up with a web of unlicensed data. We end up with a web where almost everything is properly licensed and you can kind of understand how you can reuse it and remix it in the future.

[00:15:38] **Anna:** So attribution and provenance are just baked into these pieces of data on the premise.

[00:15:43] **Sam:** That's exactly right. Yep.

[00:15:45] **Anna:** Right. Well, I mean, it's interesting, and it's also interesting, particularly in light of sort of, sort of cc's strategic shift. Like it's not just sharing for sharing sake. It's not that everything must be open, but we're very much focused on something that we call better sharing. And so. There's a number of ways that one can talk about that, but like sharing to improve the world based on practices that empower and enable everyone. So, not everything should be openly licensed, for instance, there are culturally sensitive materials that maybe shouldn't be, there are, there is traditional knowledge that maybe held that really shouldn't even be subject to copyright. But here we are in this world where that just happens automatically. Right. And so thinking about how Arweave would make sharing better? Sort of more broadly within the commons and sort of for the purpose of the public. Good.

[00:16:34] **Sam:** Well, I mean, what we're attempting to build is an open knowledge commons. If you will, like. We really want it to be such that if there's any human out there that thinks that a piece of information is useful and should be sort of available to other people in the future and licensing is self-evidently part of that. You don't think, this will be useful to people in the future without thinking, well, how should it be useful? Yeah. You need to explain that as part of it. But we want to make it so that that ability kinda superpower, if you will, of storing the information for essentially indefinite periods of time is available to all people for the lowest possible cost. That's another interesting thing about the network. It doesn't try and extract value at all. It offers the service at the minimal cost that we think it's essentially risk tolerance. We say 0.5% storage cost decline. You could imagine it other ways, but we think 0.5% is pretty conservative. Yes. And so the network doesn't try and extract value from people it tries to instead just make this utility available.

[00:17:33] **Anna:** Got it. So it's a sort of a facilitation of an open knowledge commons. Is there, are there no cases where you're seeing, or is it even possible to privately store things on the permaweb or is that just not, is that not part of the equation?

[00:17:45] **Sam:** You can store ciphertext if you want to. That's totally possible. There's no, like, you know, would you say true private data? It's all on a public ledger, but if it's sufficiently encrypted cipher text, there's an expectation in society that values... that is secure. And frankly, it's actually a very interesting question right? Do you think that all encryption or we will ever get encryption algorithms, which will never be broken? There's a school of thought that says yes. And we are already there. There's a school of thought that says probably not. I think I'm on the 60% probably not camp, but yeah, I think I'm also in a minority.

[00:18:21] **Anna:** interesting. So this system is built for openness.

[00:18:24] **Sam:** Yes.

[00:18:24] **Anna:** Like added core. Yes. Yeah. So. This is gonna be a fun one for you because I frequently hear about Arweave and Filecoin being described as the same thing I know. And you chuckle, I'm aware of the fact that they're different, but can you talk us through key differences? Like why those differences matter and what problem one is solving versus the other and so on.

[00:18:46] **Sam:** Yeah, sure. I mean, two fundamental things: Arweave is about permanent storage and temporary storage. So we see that there is this new ability that you can achieve with a decentralized system that where, you know, you don't have to trust the person that's running the endowment and you don't have this centralized control of the archive you think you can use that to that's centrally enable permanent storage in a way that wasn't possible before. I think Filecoins focus is more on being a cheaper version of S3, because they see the arguments very different. They're basically saying, okay, well, there's all of this, you know, open hard drive space in the world. Maybe we could make use of it and offer cheaper storage in a decentralized way. That's interesting. It's just a very different model than Arweave. The other component I would say is that data on Filecoin is totally private by default, so it's all encrypted. On Arweave it's the exact opposite. It basically started with Bitcoin and said, what are the minimum viable changes to Bitcoin that have to be made, which is completely open ledger in order to make it so that you can do this long term data storage. And that's how we ended up with the design where everything is open by default.

[00:19:51] **Anna:** That is an excellent explanation of the differences there. Are there other organizations or technologies in this sort of general storage space that you find interesting or are doing interesting things?

[00:20:03] **Sam:** Yeah, certainly that are doing interesting things. I mean, I'm kind of curious about the temporary storage systems, so there's also Storj and Sia in that area. And I was a I mean, I'm not sure if you can call $15 or something, an investment per se, but I was an early contributor to storage, which was the first attempt as far as I'm aware, at least to build a properly decentralized storage system. But it's focused on temporary storage. Yeah, I think that there is something interesting happening there. My personal view of blockchains is that they're Large slow databases that are kind of expensive. So you have to have a really good reason to use one. And we see that it's well, you use it for the things that are not possible within the centralized space. So you can't trust a company to do this. So, trusting math is much better and not having a single centralized point of failure is a really important thing for a system like this, like a 10 X improvement of the old archiving systems. So, that's where we saw that there was this niche to use this blockchain technology. But in temporary storage I think it remains to be seen whether that idea that you can make use of all of this spare hard drive space in the world. It remains to be seen if that is actually possible long term, but it would be cool if it was for sure.

[00:21:11] **Anna:** Yes. So, you're sort of weaving into this, but let's just talk about decentralization as a concept. So why you think it's necessary, why you think it's important and also. Is there a world where it's balanced with centralization as well, like that there are offsetting technologies or use cases.

[00:21:28] **Sam:** Yeah, that's very interesting. So, I think maybe this is, we've already kind of given this away, but I don't see decentralization as an end in itself is just a means. And it's a means for which we pay our high cost too. Like, It's not a cheap thing because you have to have basic... What decentralization essentially means is literally non-centralization. So replication redundancy is a necessary part of that else. When something stops working, when the central thing stops working, the whole thing stops working. And so you pay a real cost for doing this. I don't think that it's a sensible solution to every problem in any sense. But for extremely long term archival data storage, I think it's actually a perfect fit. If you will, same with running an endowment structure that you can't trust any humans to touch. I think Bitcoin's also like this Bitcoin says, at which the Bitcoin isn't just burning energy with this work, but that's a separate conversation.

[00:22:25] **Anna:** Oh yes.

[00:22:27] **Sam:** Right. But you know the idea that you can have a trustless mint in Bitcoin that's really powerful and it requires decentralization basically anywhere you have. The ability for someone to print money. Unfortunately, history shows they do, and they do it en masse until it pricks. And so Bitcoin is a cool solution to that problem that requires decentralization. I wouldn't insert it where it's not necessary. And you could even argue and you were saying, is there some mix between the two? Well, I think, you know, systems like Arweave and Bitcoin are actually sort of like that. So. Bitcoin for example is controlled by, there was an interesting DARPA study that was published yesterday, I think about potential vulnerabilities with blockchains and they argued well decentralized kind of, but you only need to coerce four people in order to change the way that the Bitcoin protocol processes new transactions. Oh, decentralized. Right. But that's also not the important thing. because as I see it, it doesn't really matter what those four people are doing. I mean, as long as they continue to process transactions, and this is a technical conversation, that's interesting, but maybe we shouldn't get into it now. But putting that aside, the important thing is no matter how much hashing power they have, they can't they can't print more than 20 more million tokens. And in the case of Arweave, they can't rewind time and remove data or alter data or anything like that. So math enforces this stuff, and then we have. Decentralization or close to decentralization to make sure that we don't have this single centralized point of failure problem. So I think that actually everything that we believe is or is claimed to be decentralized is quasi-decentralized. Yeah, relative to the dream of, everybody's running this currency node on their phone and there's like millions of them. What tens of millions of them, that's just not realistic. It's unfortunately not how the world works. But I also don't think it's necessary to get utility out of these systems.

[00:24:15] **Anna:** Right? Well, that's a very, that's a very practical way of looking at it. And, I'm glad you brought up some of the things that we wish might be different because within our community, right. And beyond the community, there are very mixed feelings. And very varied levels of understanding about blockchain technology broadly. So I do appreciate you sort of speaking through it in sort of practical terms, but I mean, we see folks who are huge advocates and then we have folks who really don't understand, and then we have others who just say, it's one giant scam. Like there's no utility. There's…what is it? It's a solution and search of a problem, I think is the phrase that folks like to toss around. And so. I would be curious to hear your thoughts 'cause you are in this space, like on these existing narratives, like what's your advice for people who are apprehensive or who are overwhelmed, but who may actually be technologically curious?

[00:25:02] **Sam:** So I, I can say that I really empathize with all of those. Arguments apart from the last one, maybe we'll get there in a second, but I would say that the important thing here is not to throw out the baby with the bath water, right? Like there is so much nonsense in the blockchain space. And it's endless, and it keeps happening in these cycles. There's a cycle in 2011 and in 2015, then in 2017, then in 2021, If like hype and nonsense and noise but, and I can understand why that would put people off and it puts me off frankly, I despise working around this stuff all day because I really care about the thing we're doing. And I just don't really care about speculation, frankly but here's my sort of what I've come to understand of what's happening here, basically there is a core technology here that is useful and I'll get to that in a second. But that technology itself is to some extent, at least partially financial. And the thing is that new technologies, they always produce bubbles. Unfortunately, this is the gonna kill the I won't quote it 'cause I'll get it wrong. But there's this well known cycle basically where you get a new technology. Everybody gets overly excited about it in the short term, huge amounts of money flows into it. And then the bubble pops and everyone thinks the technology is dead. And then sure enough, five years later actually has the technology that everyone's using. It just happened a bit slower than they were expecting. But in the case of blockchains, I think because it is its core, a financial technology, this has just made the bubble cycle component of it go wild. Which has been very unfortunate in some senses. And it's buried, what I think is the real need here, which is look, blockchains, allow us to transact as people, without the need to trust each other completely. And there are certain things, there are many things that really doesn't matter for, but there are certain things where it's unbelievably important. One, I would say, is a currency. You give anyone the control to mint the currency. They'll just do it. They'll go wild. It's sad, but true. The US used to be attached to the gold standard, right? But then the person goes on TV one day and says, oh, whoa we said these things were redeemable, but it turns out we printed too many. And we're doing a temporary pause. The redemption of these notes. You. For their underlying assets and what you know, that was 50 years ago, they never unpaused it and now they've printed 33 times as many tokens.

[00:27:28] **Anna:** Yeah, it's been... the printing of money is pretty wild in modern society.

[00:27:34] **Sam:** It's happened all through history. Basically, you can't get away from the fact that if you give someone the power to print money, they're just gonna do it. And I would argue that the same is true in information. If you give people the power to censor and alter information, they will do it. And so those are two areas that I think there's a very strong argument to say, well, if we can delegate this to math, which doesn't change then we can add something that is useful for society. That is the baby that please don't throw it out. like, I understand the rest is a whole lot of noise. I really empathize with that.

[00:28:08] **Anna:** It's interesting to think about that. You, you wouldn't require that trust 'cause I think through like what if scenario where somebody. Places something on the permaweb, but then alters it, like someone alters it and puts a new version of it on there, for instance. So it's like two distinct pieces of data or like, exactly. How does that work? Are you able to connect? They wouldn't be.

[00:28:31] **Sam:** I mean, you could connect them. There is an interesting argument for sort of Meetic tracing, if you will, where if someone modifies something on the permaweb, they can say, this is a modification of this old piece of data or multiple old pieces of data. And this is just in that metadata I was describing, which can be very cool, but you know, if they modified a piece of data on the permaweb and they uploaded it again by default, it has no relation to the original piece of data.

[00:28:55] **Anna:** Got it. And so there is the risk. Humans, because this is what humans do, will misuse this technology I would assume.

[00:29:02] **Sam:** Well, I think that's true for essentially all technologies and certainly this one.

[00:29:05] **Anna:** Yeah but the goal here is not that .

[00:29:08] **Sam:** Oh yeah, absolutely. And, we've put in place systems that are, there is no top down censorship in the network and that is something we've worked very hard to bake in if you will. But that. There are systems that essentially expose the storers of data to their local laws. So some of these other data storage systems that are decentralized, they have that data encrypted. And that there's kind of a problem in terms of responsibility there, because well, if it's encrypted data, you don't know what you're storing. Right?

[00:29:37] **Anna:** And it could be illegal where you are.

[00:29:39] **Sam:** Right. It could be illegal. Whereas in Arweave's case, well, you can just say, I'm not gonna store encrypted data. It's totally within your rights. No one is going to punish you or anything like that for doing that. And then the data that's plain text, you can just look inside it. So we built a system called shepherd, which is basically a sort of generalized architecture where you can plug in classifiers and look for content that you do or don't want to store. And this allows you to basically adjudicate whether this is something that is against your morals or against the laws of the country you live in, and then don't store. And that sort of exposes the network to, you could say the meta of the information governance of the world, essentially.

[00:30:19] **Anna:** So if I am a Storer and something comes my way and I object to it on moral grounds is that recorded somewhere? Does the community see that response?

[00:30:31] **Sam:** They it's not recorded anywhere per se, but you can always ask someone what they're storing, so you can go in and say, Hey, do you have this TX ID, which is important because now there's a means of essentially legal recourse against the stores. So if you know that there's some piece of data, that's illegal, you can go, who's storing this, just pull the whole network and then you can find out and then they can be held accountable. And so now they're heavily incentivized of course, beyond the morals of the requirement to do so to make sure that they're not storing anything that is deemed unhealthy in their society is essential.

[00:31:02] **Anna:** Got it. So there are these safeguards for morality or legality sort of relying on humans essentially to intervene in what is stored and where.

[00:31:12] **Sam:** And what we would hope is of course, that people live in, you know, at least some good subsection of people live in democratic countries, where there is a same legal system for changing that idea of what is reasonable content or what's not. It's actually very much like web servers. So if you host a piece of content on a web server, you are legally responsible for that piece of content. Well, as I described before, Arweave nodes are web servers. And so all of those legal responsibilities and systems sort of percolate through to this new level, if that makes sense.

[00:31:43] **Anna:** Yes. No, that does actually make sense. And that's a really interesting, that's a really interesting side of it. Okay. So you talked about shepherd a bit. Are there other new things that you're working on or within the Arweave ecosystem that are particularly interesting?

[00:31:55] **Sam:** Well, I mean many

[00:31:57] **Anna:** but how much time do you have,

[00:31:59] **Sam:** right, right. I dunno. We just got done with the Open Web Foundry demo day. So the Foundry is a series we run, it's kind of like a startup incubator for people building applications on top of a permaweb. And there are people building all sorts of exciting stuff there. Channeling the system for storing sort of personal legacy items. But all the way through to really interesting stuff like storing. Would you say academic accomplishment certificates, right? Because then you have this open ledger that anyone can access and anyone can contribute to. It's kind of like an, you get the benefits of an open protocol versus this closed system where, you have to sign up with this company who gets to decide who can, certificate certain things and not Yeah, and this is an open protocol. Anyone can come along and say, well, I certify you as an expert and such. And whether you people choose to trust me or not is their business, but it's open. And that's a pretty interesting use of the technology, but there's frankly, so many more things it's hard to list off.

[00:32:52] **Anna:** Oh, that's neat. So is this an incubator specifically for the permaweb then?

[00:32:56] **Sam:** Exactly. Yeah. For building permaweb applications. So this was a little bit of a. A sideways discovery, you could say. But when we built the permaweb, we shipped it, we got thinking about, well, it serves all of this information and webpages, huh? What happens if you put web apps inside it? And then we thought, oh, well, they're permanent. That's kind of cool that you pay. Once you store the app forever, it never goes anywhere. And then we realized, oh, well that means from the user's perspective, that if there's a version of the application that they like, the developer can't take it away. This is not theoretical. This is really useful. Because in the Web 2 world, the game is basically like offer service for free for everyone. And then at some point, once you capture the market, start essentially trying to extract as much value as possible as you can from the captive audience. Well, you can't do that with the permaweb app. You get these sort of what you see is what you get forever. Guarantees, which is pretty cool. And then there's this other side of it, which is, huh. So one of the things the permaweb does is it changes the relationship of control between users or people and information. It's this neutral data lake where nobody can change a piece of information after it's been uploaded. And that allows you to build these web applications that have no centralized controller, which doesn't mean they're anarchistic, but it means that their content policies are baked in from the start. And so everybody can understand the deal they are making when they start to use an application. And no centralized controller can come along and say, whoa, actually, we've just decided to completely change that, which is a really powerful thing. We think it totally changes the power relationships on the web fundamentally.

[00:34:37] **Anna:** It does. It makes me think of a number of these technologies that have been free to start and then all of a sudden it's like, oh we did say this was gonna be free forever, but actually if you're gonna use more than this amount of space, so if you're gonna do this differently, or now we're gonna start collecting your data. Exactly. It also makes me think of the old arcade games that you can play on the Internet Archive, which is important that some of these things stay the way they were because yeah we don't want to lose, we don't want to lose that era.

[00:35:02] **Sam:** The thing that did always comes to mind for me most is two examples. The first is Medium. So if you know, used Medium back like five years ago as your blogging platform, well, you. You were signed up and they were like, here's a free blog. And, but you thought that was kind of cool. Yeah. You shared that link with so many people. And then one day they came along, they said, whoa, actually, we're gonna monetize your content and put it behind a paywall and you don't get to choose. And now all of your readers have to pay us to access your content, which is totally crazy. And same way with Gmail, when you start using Gmail at the beginning, you. Well, if you are like me, I obviously did not read the terms and conditions. So I was like, yep, sure. let's do it. But that was 14 years ago or something. And over that time, they completely changed the deal. So at the time it was basically, here's email for free and now it's, oh, here's email for free, but we're gonna target ads based on the content, the messages you send and receive. And by the way, you might actually sell access to that data to, well, frankly, whoever wants to buy. It's a totally different game. Yeah. These permaweb applications basically subvert that, or put a bull walk in place such that that can't really happen. Now the quality of the application must increase over time, not decrease. Yeah. And otherwise users will just use non version.

[00:36:18] **Anna:** Exactly. Oh that's a fascinating approach. Yeah. There's so many examples of this recently. I've read some article about Vimeo changing their, what was it? It was like changing the pricing for bandwidth. So high volume, like people who had built up audiences were suddenly paying a lot more for those audiences to consume their content. That made some waves in the community. So, before we wrap, I feel like we've covered a lot of ground. But sort of in the broader open movement, where CC sits and you have these open ecosystems on the web, like, are there any trends there that are particularly catching your eye? Anything that you wish would happen in this open ecosystem?

[00:36:54] **Sam:** The trend I see is people hating blockchains and I was just talking, please try and look past the shouting about whatever the latest ape type NFT is, or the latest coin that's totally gonna pump. I get it. That's annoying. And it's zero sum games in a lot of ways, but I really see a lot of people that, if we could cut through that noise, a lot of people that would find real interest and value in the stuff that's being created, totally discarding it because of the things around it. And I find that's so sad to watch and I've seen this cycle actually three times in depth so far. So the first time was like 2015 ish, 2014, 2015 Bitcoin had an excitement, bubble first one or the first main... Well, it's a second major one technically, but the one that really people started to talk about the first one. It was so niche, no one even noticed essentially.

[00:37:48] **Anna:** Yeah.

[00:37:48] **Sam:** Yes, but there was this hype bubble and then you saw, people like, ah, actually blockchains are terrible when the prices came down and then 2017, 2018 blockchains are terrible. And it's like, okay, not really. It's just like all things in life, some good, some bad. And it requires the person to engage and to sort through it. But I promise I'll give you, you know, some tiny number of tokens, if you can really give me a good argument why the entire thing is completely useless. I promise you. It's not completely useless. There is some gold somewhere buried in here.

[00:38:19] **Anna:** That sounds like a good challenge to our listeners. Looking for that. Yes. That utility. Well, this has been fascinating. Thank you for delving so deep into some of these concepts that I know are sort of complicated and foreign to many folks. It's been a real delight.

[00:38:33] **Sam:** Thank you very much for the opportunity. We're huge fans of Creative Commons. So this has been amazing.

[00:38:37] **Anna:** We appreciate that.