

My, Myself, and AI Podcast

Punk Rock, the Peace Movement, and Open-Source AI: The Mozilla Foundation's Mark Surman

Shervin Khodabandeh: How can open-source technology platforms keep AI trustworthy and safe? Find out on today's episode.

Mark Surman: I'm Mark Surman from Mozilla, and you're listening to *Me, Myself, and AI*.

Sam Ransbotham: Welcome to *Me, Myself, and AI*, a podcast on artificial intelligence in business. Each episode, we introduce you to someone innovating with AI. I'm Sam Ransbotham, professor of analytics at Boston College. I'm also the AI and business strategy guest editor at *MIT Sloan Management Review*.

Shervin Khodabandeh: And I'm Shervin Khodabandeh, senior partner with BCG and one of the leaders of our AI business. Together, *MIT SMR* and BCG have been researching and publishing on AI since 2017, interviewing hundreds of practitioners and surveying thousands of companies on what it takes to build and to deploy and scale AI capabilities and really transform the way organizations operate.

Sam Ransbotham: Hey, everyone. Today Shervin and I are pleased to be joined by Mark Surman, president of the Mozilla Foundation. Mark, thanks for taking the time to talk with us. Let's get started.

Mark Surman: Thanks, Sam and Shervin.

Shervin Khodabandeh: Mark, maybe let's start by hearing a little bit about the Mozilla Foundation and your role there. Could you describe that for us, please?

Mark Surman: Mozilla has been around for 25 years now — [March 2023 was] our 25th anniversary — really making sure that the internet is in the hands of the public; that how we build the internet is something that balances not just commercial interest but also public interest, personal interest; [and] that humans are in mind as we design technology. And in the first era, we focused on the web; we built Firefox. And right now we're really focused on making sure those values show up in the era of AI.

We really want things like human agency [and] accountability for how tech gets built to show up in the era of AI. And we talk a lot about that when we hear people talking about responsible AI, but it's not how we see what gets built for us. Often, there's just a rush to get stuff out the door. We've seen that really a lot in the past year with all the GPT-X and everything else — these things rolling out to billions of people without a lot of consideration for how they might impact people.

And so, really, that's what we're trying to do: is make sure that that changes — trying to make sure it changes through

advocacy; trying to make sure it changes through building new, open-source AI; and also slowly, by building AI into things like Firefox, but in a way that actually keeps people in mind, keeps them safe, [and] empowers them.

Sam Ransbotham: I think that keeping people in mind is big. You used the phrase “human agency” a lot. One of my personal pet peeves is when we talk about “artificial intelligence does X; artificial intelligence does Y.”
People use these tools to do something, and when we use phrases like “AI does something,” I think we sort of abdicate any responsibility, as if “oh, no, it’s the machine doing it.”

We’ve got to retain some of that agency here — that we’re in charge, or we can be, at least for a while, in charge. So, what are these sorts of initiatives that you’re talking about to, well, for example, with the foundation and trustworthy AI, and the progress you’ve made, and Mozilla Ventures? Tell us about some of these initiatives.

Mark Surman: Well, you know, I think that’s right: People need to be in charge, and it isn’t AI that does things. Sometimes, actually, it is companies that own big pieces of AI that do things, and so a lot of what we are working on is how you put AI in the hands of people or smaller companies or developers. One example of that is around open-source large language models.

You’re hearing a lot about open-source large language models, but how many of us have actually used them for something — used them to build a

personal assistant, used to them to help do sensitive research, used them in our work or our everyday lives? And so we’ve launched a company, Mozilla.ai, that’s about taking open-source large language models and making them user-friendly, making them trustworthy, letting you use them on your own personal data in a way that you control. You’ll see things coming out of Mozilla.ai early in this year.

Some of the other things we’re doing is looking at how you take the current wave of AI and roll it into something like in a browser, but not in a way that’s about selling you something — actually in a way that helps protect you or helps you make better choices. So over the course of the next few months, we’re rolling something called Fakespot, a company we bought, into Firefox, that uses AI to help you spot scams, help you spot fake ratings. And so it’s those kinds of things: taking the current wave of technology and putting it into the hands of people to make decisions for themselves.

Sam Ransbotham: So I like that, let’s say, push toward making the plumbing of artificial intelligence easy for people to use; otherwise it’s going to be dominated by the large technology giants who can afford to put these models in place. And you mentioned the large language models; obviously, they’re fascinating and they’re amazing, but they’re also developed and delivered by large technology companies that don’t have my personal objective function in mind. And maybe this model that you’ve talked about, when you get this plumbing available and open for other people, then we can start to have

artificial intelligence [whose] optimum optimization algorithm is about what Sam wants, not what a random technology company wants. How far are we from being able to do that?

Mark Surman: I think we're a ways away from being able to optimize the AI for us, but not as far as we think. So we're at a spot that I think is similar to when Linux came out, which it was an alternative to Microsoft, but it wasn't an alternative that most people could use. And then it got to the place where developers could use it, and then you started to see user-friendly desktop Linux. And we're at a spot where the core of that, more and more people are coming out with open-source large language models. How you deploy them, how you use them for what you want — that's hard right now.

I would say over the course of this year, you're going to see more and more people come out with stuff that lets developers use open-source large language models instead of turning to the big cloud-hosted models. And then, over the next couple of years, those are going to turn into things that all of us can use for what we might start to call open-source personal AI.

Sam Ransbotham: The analogy to Linux — I hadn't really thought about that — but it seems interesting from a couple perspectives. One, most of us don't actually even now use Linux on our desktops. So, you know, if that's the model, then I'm kind of worried that we're not going to have that penetration.

On the other hand, if I switch it around and think of how Linux runs the whole of

the internet, then I'm wildly optimistic. And so maybe your point [is] about, if we can get the developers the tools out there, then the market will decide about that objective function. If we have competing developers out there using these tools, then maybe the market will help with that.

Mark Surman: Well, you know, what happened with Linux was, it became the underpinning for Web 2.0: Linux, Apache, Firefox, and the web stack, right?

Sam Ransbotham: Mm-hmm.

Mark Surman: Open technology allowed developers to create alternatives to Microsoft, create whole new categories of software, of web services. I mean, you wouldn't have a lot of the social media that we use today had web standards not emerged. Those companies wouldn't have gotten off the ground had they not been able to set up cheap Linux servers and so on.

So I think that's exactly right: We have an opportunity to create a much more decentralized digital economy than the one we see emerging around the big, old companies and the new AI labs that they own. I think there's a chance for something that's kind of much more rich and open than that.

The point you make, though, about Linux is a good one, which is, you know, it really shifted things for developers. What shifted things for people really was the web, and it was that, if you think back even further to the mid-'90s, it's like all of a sudden, anybody could create a web page. All of a sudden,

anybody could have a digital presence, when really that was something that felt like becoming a publisher and was something that only rich people could do. And then we did swing back a little bit, where Microsoft tried to vacuum the whole of the web back into Windows, and by the time you get to the end of the '90s, 98% of browsers are Internet Explorer, and they're all oriented toward kind of tying into the Microsoft and Windows ecosystem.

And then you had Firefox come along in 2003, 2004, and again, it kind of swings back to the people. And I think that what we'll see in this AI era is, you move from a lot of open science, a lot of open research, like thinking about the transformer paper that actually led to large language models. That was an era where there was a whole openness in people sharing what they were doing in AI.

You now see a bunch of big companies, a bunch of big momentum, trying to close it down and grab it all for themselves. I don't think people are going to want to just live with that. Sure, those big companies will continue to exist, but I think you're going to see a swing back, like we did with Firefox and the open web, to more people wanting to control AI for themselves.

Sam Ransbotham: I think that's an optimistic view ... but I'm somewhat wary of coming across too anti-big technology company, because I do think that there's a huge role for all people in this ecosystem, but I think your final point there of worrying about some of the land grab is important. Even if the smaller language models don't end up

being huge or taking off, I think their presence really helps too because otherwise we have an unchecked development around these large technology firms, and they don't have to be dominant in the marketplace, I guess is what I'm saying, for there to be value from these sorts of initiatives.

Mark Surman: Absolutely. There's a real connection in how I see it between open source, which, you know, if it's working out well, can give a lot of people building blocks to create their own things, and open markets and competition. And really, that's what you want; you just want diversity. You don't want to close things down [and] the land grab turns into a few companies controlling how everything works.

And it's just finding that balance. One of the things we were really happy to see in the U.S. executive order on AI that came out late last year was this push to the FTC to think about competition in AI early on. And I think that's the thing that we didn't think about in the Web 2.0 era. You know, there were a lot of land grabs — really, arguably, a lot of anticompetitive behavior — so it's good to be looking at that early on in the AI era.

Sam Ransbotham: Maybe we did learn something from the last time. I've seen some recent papers out there looking at how much of AI is coming out of industry versus academia. That's something else I wonder about: Who has the resources to pull together these models? And perhaps the Mozilla Foundation and others are necessary here, because the idea that Sam alone at night in his dark room pulls together

a competing large language model is really unlikely, given the resources it takes. So, what's the model for Mozilla to support these?

Mark Surman: I think public options is how I think about them, in terms of people experimenting and trying things, and, as you say, not all of us have the resources. In fact, most of [us] don't have the resources to build our own AI systems or train our own models. And you see people like The Allen Institute out of Seattle talking about building a whole kind of pool of open-source large language models. You see community projects like EleutherAI, where people are putting their resources to train things. That's the kind of thing that Mozilla really wants to support and be a part of. So we're working with both Allen and Eleuther on this kind of stuff.

And then you see governments — and I was really happy to see this, both coming out of Europe and the U.S. last year — saying, “We're going to build public infrastructure that researchers and others can use.” That's a trend we hope to see continue, and it's very much one that goes alongside open source. Open source is about a set of public building blocks; we want to see those in AI. And then, kind of shared infrastructure or publicly funded research infrastructure so that people can play with that open source [is] also critical. And together, those things can drive some innovation that's different and maybe differently interesting than what's going to come out of big companies.

Shervin Khodabandeh: You mentioned that you were supporting Eleuther and

Allen. Is that through Mozilla Ventures? Is that how that's working? Or what's the mechanism for them?

Mark Surman: That's a good question, Shervin. We're actually working with those kinds of community partners in a bunch of different ways. So Mozilla.ai, which is our R&D lab, ... aims to take open source and turn it into stuff that people can use, basically — commercial stuff, noncommercial stuff that people can use to take control of AI themselves. We work closely with other open-source projects, just like we did in Firefox in the past. So people like Eleuther and Allen are people we collaborate with. We also, through Mozilla Ventures, fund a bunch of open-source AI companies. There's one called Flower, which is working on standards for what's called differential privacy.

And then we do a lot of grant-making and fellowships for people who are in the AI space, the trustworthy AI space, the open-source AI space — people like Deb Raji, who's a real pioneer in open-source auditing of AI systems.

Sam Ransbotham: You've described, I think, a very market-oriented approach so far, but you also mentioned the regulatory part. What do you think the role is for regulation in all this?

Mark Surman: We're early in the development of these technologies, and I don't just mean AI — I mean the internet; I mean like weaving the digital into our life. And we're probably going to be with the digital or what's next in terms of the digital for hundreds and hundreds and hundreds of years. I think it's a big shift in humanity. And when

new things come, you always kind of start with this era where you're not regulating the stuff because you didn't even know it existed, and then you don't know what it is.

I think we've now lived with the digital long enough that everybody agrees it's time to figure out what's the balance between the public interest and private interest. I just see us in that phase, and that phase means doing tech regulation. It means doing tech regulation, in my view, not in a rush, and carefully. And so in the kind of wave that is coming out, you see the AI Act coming out of the EU, you see the executive order, which hopefully turns into action this year, coming out of Washington. You see stuff in really every country around the world.

The key there is to tackle the big issues first and then to learn. So I think those big issues are privacy, competition, and really actually making sure that consumers are protected. So I think if you get consumer protection, competition, and privacy right, you'll have the basis of what you need to govern AI.

And then I think there's a bunch of stuff that we're earlier on, which is making sure that we connect, say, what are human rights or civil rights, and how do they connect to these technologies? I think that's something we want to make sure we put in policy frameworks, but I don't think we quite know yet how to build laws for that. Maybe, you know, what the EU has done around a risk framework is a good start, but I think it's going to take us years, probably decades, to figure out how to do that right.

And the main thing is that we build the capacity inside of governments, and the relationship between governments and industry and the public, to negotiate that over time, to adapt it, to understand we're going to live with something new. And as long as we balance the public interest and private interests, as I think we've tried to do in things like food safety or auto safety or things like that, we'll find the right path.

Sam Ransbotham: That's optimistic. I was a little struck by you saying, "Oh, we're early in the internet days," and, you know, it feels like it's been around forever. But no, we're still quite new on that. But then you paired it with privacy, which ... kind of bothered me a little bit as you said it, because I think about how ... maybe how poorly we are doing so far on that, and when we think about how much, let's say, lack of trustworthy infrastructure is in place in the internet to start with, and we're just pulling our teeth out to get the world move from HTTP to HTTPS, right?

And while we didn't need that secure, that "s" on there when it was just 16 computers at DARPA hooked together, we do when we're connecting the whole world. But we find that once these things get entrenched, then they're just brutal to pull back out and retrofit. So when you said "privacy," then I started to get worried, because if that's the analogy, then I'm worried that we're never going to be able to get back on top of this Pandora's box that we've opened. Give me more optimism there.

Mark Surman: Well, you know, the privacy one — I was specifically talking about privacy regulation, and maybe,

you know, I'm talking about privacy regulation and consumer data protection regulation in the U.S. It's clear: There are a lot of dimensions of privacy we aren't doing well, exactly as you said.

Sam Ransbotham: OK.

Mark Surman: And I think as we know that we're not doing well, one of the things we need to do is develop good consumer privacy regulations as a baseline for our social contract in the digital era. And you saw a first shot at that in the EU with the GDPR. I don't think that's really worked out. It's not very pragmatic, although maybe the principals are right.

You see an attempt at that in the last couple of years with the California Consumer Privacy Act, the CCPA, which is just one jurisdiction in the U.S. but actually has some better ideas — maybe ideas that haven't fully been picked up yet, like the idea of us being able to have data intermediaries, or data representatives, who are out there acting on our behalf to protect our interests, which, privacy's so complicated, wouldn't you like to be able to delegate it to somebody you trust? Maybe Mozilla in the future?

Sam Ransbotham: Mm-hmm.

Mark Surman: I think it's like moving into that topic from a policy perspective feels important, urgent, and it feels like the bedrock we need as we go forward in this digital era.

Sam Ransbotham: You're the president of the Mozilla Foundation. I'm guessing

that that wasn't your starting job. Tell us a little bit about your history and your career and how you got interested in these topics, and what your background is.

Mark Surman: When people ask me about that, my answer always — because it's true — is punk rock and the peace moment. When I was a teenager in the '80s, I was very much a punk rock kid, and punk rock was really tied into, or at least a branch of punk rock was tied into, the fact that we were in the middle of the Cold War. It was scary.

And the peace movement was there saying, "Less nukes." And so I was kind of into both of those things, cared about both those things. I cared about the music; I cared about the politics. And I happened to live in a very small town where I got to work in high school in a network TV station at night, running the shows, running the commercials. They kind of had a rule that we could play our own public service announcements. If the commercial time wasn't sold, you could decide if you were going to play a Red Cross commercial or one of the big PSAs of the time.

And I thought, "Why don't I make a commercial for my peace group? And just, you know, I've got this empty time; I can play it," and I did. It was very corny. It was the first video I produced.

I produced many more later. And I came in one day to play that public service announcement, and I couldn't find the tape. And I went to the station manager and I said, "Do you know where the tape is?" And he said, "Oh, well, the station

owner said we don't play local public service announcements."

And that was completely arbitrary, of course, and he didn't like my message of punk rock and the peace movement. And I guess that was an early — you know, pretty privileged, but early — lesson in censorship and how media ownership ties to censorship. And really, my whole career since that — you know, I went to film school — has been around focusing on people having their own voice through communications and through technology.

And when the internet came in the mid-'90s, I was like, "Oh, you know, this kind of activism filmmaking stuff that I had started to do, I bet you I can actually do a lot more with this internet thing." And I haven't turned back since.

Sam Ransbotham: So when you think about how Mozilla's organized, are two people thinking about AI, or seven people thinking about AI? Is somebody thinking about it in their lunch break? How are you getting these sorts of messages throughout the foundation?

Mark Surman: Mozilla obviously started out by thinking about the web. And that was the technology that defined the moment, defined what was going to happen for a decade or two, starting in the mid-'90s.

And a few years ago, a bunch of us got around saying, "Look, we can't just think about the web. This AI thing, data-driven computing, that is going to define the next few decades. And we need to take our values — openness, people having

agency, privacy — and make sure that those shape where the AI era goes."

And so we [wrote a paper about three, four years ago on what we saw as a vision for trustworthy AI and slowly started giving more grants, doing the philanthropic side. But over time, you know, kind of everybody across Mozilla started to say, "We need to do more to make sure that AI goes in a direction that somehow reflects the values that we have." And so we set up this AI R&D company, Mozilla.ai, we set up Mozilla Ventures; about two-thirds of the companies — I think there are 30 companies in there now — are focused on trustworthy AI. And gradually, in our core products, including Firefox, we're looking at how we layer in trustworthy AI.

So maybe it started four, five years ago with a few of us thinking about it on our lunch break, and now we're at the spot where, really, everybody across Mozilla is starting to think about "How do we play the role in the AI era that we played in the web era in terms of shifting the direction of things and decentralizing power?"

Shervin Khodabandeh: Mark, Mozilla Foundation is about to release a new report on trustworthy AI. Could you tell us a little bit about the focus of that?

Mark Surman: There's four things in that initial paper that we looked at. If you want trustworthy AI, if you want more agency, you want more accountability in the AI era, what are the things you look for? And we talked about shifting industry norms, how stuff gets built, shifting what the technology is and

the products are that people actually have available to them, shifting consumer demand, and then shifting the policy landscape.

And so we looked at all those things in this recent report and said, “How are we doing?” And interestingly enough, we’re doing OK on some of them and horribly on others. I mean, maybe that shouldn’t be surprising.

On the policy front, it’s better than we predicted. You know, three, four years ago, when we wrote that paper, we talked about just making sure that policy makers had the expertise to write good AI regulation. And you’ve really seen policy makers step to the fore. We haven’t solved it all, but you see more capacity; you see things like the AI Act and the executive order that came out of the White House last year.

So that’s promising. It’s not solved, but it’s promising. On the flip side, if you go to industry norms, we saw a trend a few years ago for more AI ethics people inside of big companies. That’s turned around. We’ve seen a lot of those teams get let go. The flip side of that, though, is you see people, through Mozilla Ventures, saying, “OK, if the big companies aren’t going to do it, I’m going to start my own.” And we see a lot more trustworthy-AI startups focused on auditing or even focused on social media that have a more human dimension.

And then I would say on that piece in the middle, like consumer demand and “Are the main products that we use and that we choose, reflecting a vision of AI that is more human, more trustworthy?”

I think it’s a real sweet spot maybe to focus on in 2024, because you’ve got a real public awareness that there’s something to worry about, care about in relation to AI, but you don’t yet have a way for people to act. Like, what are these products that are going to be different? And that’s something where we hope to fill the gap and we hope that startups will fill the gap, and that kind of nascent consumer desire for something different, that nascent consumer worry about AI, starts to get filled with products that people can trust.

Sam Ransbotham: I like that. I mean, yeah, you’re right: I do see the distinction between awareness and having those available, and certainly having them available and no one aware doesn’t do any good, but, like you say, I do think there’s more awareness growing. But, of course, we’re a bit like carnival barkers, where we promised something and excited a need for something, and now I think there needs to be a rapid filling of that need or else we’re going to have another sort of backlash behind that then.

So was there anything in that first paper that you feel like you missed that you really want this new report to address?

Mark Surman: I think we didn’t put enough focus on open source in that first report, and it really has become both a key opportunity and also a key battleground. The opportunity, as we’ve seen more concentration of power in AI — much more than we imagined three years ago, because we weren’t in the generative AI era yet — is that open source could be one element in pushing back on that concentration of power and

letting small players carve out a piece of the pie.

And of course that has to come with things like competition regulation and breaking down competition of power more directly, but open source feels really critical, as the land grab happens, to counteract it. And we didn't talk about that the first time. And at the same time, as the regulatory conversation gathers steam, you're seeing various players, and I think it's probably some pretty self-interested players, questioning open-source and AI safety, saying, "What if open source gets into the wrong hands and people use that open-source AI in scary ways?"

And of course that's a thing to worry about. Any technology can be used maliciously, and this likely is going to be very powerful technology. But time and time again, and we've seen that that is true of both proprietary and open-source approaches — that they can be misused. And, frankly, open-source approaches give us a way to scrutinize what's going on and to fix it faster than proprietary technology, in many cases.

Sam Ransbotham: We've got a segment where we ask you a series of questions. These are rapid-fire questions, so just think about the first thing that comes to your mind. What do you think is the biggest opportunity for artificial intelligence right now?

Mark Surman: I think the biggest opportunity for artificial intelligence is to take the digital world that's still kind of complicated to navigate and make it disappear and feel natural and a part of our lives in a way that it isn't yet, and do

it in a way that we have control and agency. So maybe the answer is [that] a lot of the web browsers, smartphones, interfaces we have today disappear and are replaced by personal agents — things that naturally we express ourselves through as we interact with all kinds of digital things, other people, other organizations.

Sam Ransbotham: I like the flavor of disappearing, and I hope it's a disappearing because we don't have to worry about it and not disappearing because we don't know we need to worry about it. And I guess that's some of the awareness you were just talking about.

What's the biggest misconception you think people have about artificial intelligence?

Mark Surman: Certainly, people think that AI is a thing. AI is not a thing. There isn't any "artificial intelligence." It's just an era of computing, a set of disciplines that are about using data to allow computer systems to be adaptive, to predict things, to make things feel like they're happening naturally. So I don't think we should look at these things to be artificially intelligent or intelligent in any way but rather as things that we should find ways to use and control and shape so that there's more ease in our lives.

Sam Ransbotham: That sounds great to me. I'm ready for more ease. What was the first career you wanted after you finished your punk rock career and, I guess, your no-nukes career?

Mark Surman: I definitely wanted to be a documentary filmmaker in the

beginning and wanted to change hearts and minds by telling the truth. And maybe that's still what I'm trying to do, in a different way.

Sam Ransbotham: Perhaps more powerful with software than film these days, I guess. When do we put too much artificial intelligence in things? When is there too much AI?

Mark Surman: There's too much AI when we're talking about things that make life-and-death decisions. There's too much AI when we're talking about stuff where we need a feeling of humanity, where we need kind of like human emotion in making good judgments or just actually in being connected to each other. So, you know, we really shouldn't be using AI to do risky, dangerous things that somebody needs to be held accountable for. That's still a place for people. We shouldn't be using AI when we're trying to create deep human connection and have it simulate that. That's what we are for each other.

Sam Ransbotham: So, what's one thing you wish artificial intelligence could do right now that it currently can't? What's a limitation we have?

Mark Surman: I wish artificial intelligence today could really just work for me. ... I could have an AI on my phone, on my laptop, even in the cloud, that I knew was really accountable to my interests and had the capabilities to interact with all the other automated systems around us in ways that I kind of train naturally and trust it over time.

Sam Ransbotham: You've talked a lot about this interplay between the technology giants and the market forces and how all these things come together. And I think it's really interesting to think about. We can have too much market, just the way we can have too much regulation, but one of the things that's really coming out from your discussion is this idea of balance and these forces working cohesively together to get us to a point that we want to be at, versus dominated by one of those. I appreciate you taking the time to talk with us today. Thanks for talking to us.

Mark Surman: Thanks, Sam and Shervin.

Shervin Khodabandeh: Thanks for listening, everyone. We've just completed Season 8 of our podcast. We'll be back on March 19 with new episodes and have a couple of bonus episodes for you coming this winter. We hope you can join us.

Allison Ryder: Thanks for listening to *Me, Myself, and AI*. We believe, like you, that the conversation about AI implementation doesn't start and stop with this podcast. That's why we've created a group on LinkedIn specifically for listeners like you. It's called AI for Leaders, and if you join us, you can chat with show creators and hosts, ask your own questions, share your insights, and gain access to valuable resources about AI implementation from *MIT SMR* and *BCG*. You can access it by visiting mitsmr.com/AIforLeaders. We'll put that link in the show notes, and we hope to see you there.