**Lisa:** This holiday season, please consider including Scholars Strategy Network in your plans for end-of-year giving. With your generous contribution, we can continue to provide No Jargon ad-free with no paywall so that the powerful research and perspectives of SSN scholars are accessible to as many people as possible.

**Lizzy:** No Jargon is just one part of what we do here at SSN. Mandana, our producer, as well as both of us hosts, are working day in and day out with researchers across the country to help them connect their research to some of the most pressing policy issues of our times.

By donating, you can help SSN cover the $100 cost of a one-on-one coaching session with scholars to help them draft a jargon-free explanation of their research that both the public and policymakers can easily access and digest. Giving is easy. Just visit our website at scholars.org/donate.

**Lisa:** Hi, I'm Lisa Hernandez.

**Lizzy:** And I'm Lizzy Ghedi-Ehrlich.

**Lisa:** And we are your hosts for Scholars Strategy Network's No Jargon. Each month, we will discuss an American policy problem with one of the nation's top researchers, without jargon. And this month, we are talking about efforts to regulate artificial intelligence, otherwise known as AI.

**Lizzy:** Hey there, policy pals. Welcome back to the Scholars Strategy Network's No Jargon podcast, where we break down complex research without the confusing lingo.

**Lisa:** That's right. Today's episode is all about demystifying the world of artificial intelligence. I'm, insert your name, and joining me is my partner in policy crime, co-host's name.

**Lizzy:** You know it! And buckle up, folks, because we've got a fantastic researcher in the studio. We're diving deep into the realm of AI and the pressing need for transparency. No technobabble, just straightforward insights.

**Lisa:** It's like AI for dummies, but, you know, in the most endearing way possible.

**Lizzy:** Absolutely. So grab your favorite cup of coffee -- from among your multiple cups of coffee that you have laid out for you, I suppose -- sit back and get ready for a jargon-free journey into the heart -- I gotta stop you, Lisa. I gotta stop. We gotta reveal what's happening here.

**Lisa:** Okay, so for the listeners, I have given ChatGPT a prompt that says we are two quirky policy associates who host the Scholars Strategy Network's No Jargon podcast and to write us a very short banter for the introduction of this podcast episode. I mean, honestly, not that far off from our usual answer.

**Lizzy:** I feel both seen and terrified, which is fine, you know, there's definitely worse ways to feel. for those who don't know, ChatGPT is a large language model, I believe, or that is what it works off of, meaning it has been fed thousands upon thousands of our words, so that it kind of acts in a way that a human brain might learning what to say and what not to say and what makes sense in context. and yeah, I, you know, I guess. Okay job here ChatGPT, thanks for helping us out with this banter. Policy pals that might be a keeper, we'll see.

**Lisa:** Hey, I like co-hosting with you, my policy pal. So, we're actually going to learn a lot more about ChatGPT and all the different forms of artificial intelligence, because we talked to an amazing, researcher, Jim Samuel, an associate professor and executive director of informatics at Rutgers University, New Brunswick. Professor Samuel's primary research includes a focus on human intelligence. And artificial intelligence interaction. Overarching themes in Professor Samuel's writings include applications of AI, AI bias and ethics and AI education.

Here's our conversation.

**Lisa:** Hi, Professor Samuel. Thanks for coming on No Jargon.

**Jim:** Good to be here, Lisa. Thank you.

**Lisa:** Well, you, definitely research a very popular topic at the moment. It seems like everyone is talking about AI these days. In fact, our No Jargon producer and I just left an AI for nonprofits training right before this conversation with you right now. Could you tell us what artificial intelligence is in its most basic form?

**Jim:** Just to take a step back about a year or two ago, when I would tell folks that I'm working on artificial intelligence, I would be looked at as though I'm

talking about something from outer space, but now it's a different world altogether.

Everyone is talking about AI, some are very well informed and others not so much, but there's a lot of conversation around what artificial intelligence can do for us and AI applications. there's not as much of a focus on what AI is. And it's very important to understand that to be able to kind of look into the future and project as to what kind of applications we will have, based on AI in the next few years.

Right now it's like the Wild West, anyone and everyone who can is creating technology and calling it artificial intelligence. Many of those technologies are indeed using elements of artificial intelligence and some are not, but it's still being branded as artificial intelligence. So it's, it's like the Wild West right now.

Artificial intelligence has been with us for quite some time. And so we have about 60 to 70 years of perspectives on artificial intelligence and one of my research topics is to try and bring these streams together and try to study the evolution of the concept of AI. The definition that I like to use most, it's a very simple one, and I think it brings together many of the other perspectives, is this: artificial intelligence is a set of technologies that mimic the functions and expressions of human intelligence, specifically cognition, logic, learning, and adaptivity.

So anytime we speak about artificial intelligence, we are talking about creating technologies that perform human intelligence functions. And within that, we are talking about the ability to see, so we are talking about cognition. And AI, some AIs are able to see, they're able to read text, they're able to identify images, so on and so forth. Some AIs are able to demonstrate logic, whether it's in the playing of sophisticated computer games or, superseding human intelligence in games like chess and Go. We find that AI can demonstrate and mimic human logic to an extent that it actually gets better than human intelligence in narrow areas.

Then we have the idea of learning, whereby AI learns from the data, which simply means that there are, what we call as an artificial intelligence application may consist of one or more or a cluster of models. So the learning concept is that the underlying model is developed based on some original data input.

And then as time passes by, it learns from experience. It provides output. That output is evaluated. That evaluation is provided as input into the model and the AI model keeps self-improving, or this cluster of models keeps improving itself.

And that's the idea of learning. And very closely connected to the idea of learning is the idea of adaptivity; they are linked to each other, but they are distinct.

When we speak about adaptivity, we are speaking about the ability of artificial intelligence to learn and also to make decisions, which means to implement some kind of change, whether the change is updating the weights of the variables within the model itself, or whether the change is in the form of making decisions. The adaptivity component is what ultimately plays out, and that's what makes AI so powerful.

**Lisa:** Right. So by AI having that adaptivity component, you as the user are feeding into the eye, or actually improving the AI models, is what I'm hearing.

**Jim:** Yes, absolutely. So, you know, it's, it's very interesting. So artificial intelligence, in some sense, it's like a, it's a two-way street. And that's exciting. That's interesting. It's amusing at this point of time. And hopefully, you know, it will remain a positive experience for individuals and, and human society as a whole.

So on one hand, when we speak about an AI application, a good, healthy, well-defined, well-developed AI application, the application interacts with human intelligence. And then, based on the feedback it receives, it learns and it adapts. And that ability to adapt means that we can develop AI applications which are highly customized.

It could be customized at the level of a group of people with common characteristics. It could also be customized to an individual person. So, for example, we could train an artificial agent, which recognizes human voice, and over a period of time, as it interacts with that one individual, it gets better at recognizing the nuances of that individual's voice commands.

**Lisa:** What I'm also hearing is it does sound very similar to the way that we engage with social media platforms and different types of technology. The more we use it, the more it learns about us. Is that the reason for the recent rise in artificial intelligence use or really what has been the motivator for this increase in interest in this field?

**Jim:** Lisa, I think you kind of touched upon a very interesting point. So let me take a step back into what's been happening over the past few years. We have been using machine learning and artificial intelligence technologies for quite some time. So, when we speak about applications like Alexa, when we speak

about recommender systems, even simple things like when we are typing into a chat message on our phones over the past few years, we have received prompts. All of these mechanisms have been self-learning mechanisms, which means as we continue to use these applications, they get better at understanding us.

Even Netflix, or some of the websites that we use for product purchases. All of these platforms, social media, and platforms like Amazon from where we purchase goods and other places where we purchase services, they have been collecting our data. They have been modeling our behavior patterns. And so, in that sense, we have been interacting with machine learning and artificial intelligence technologies for quite some time now through our smartphones, through our smart devices and through social media and through platforms like Netflix and Amazon.

But as I mentioned before, when I would speak about AI and machine learning over the past few years, sometimes I would get glazed looks and partly due to the fact that people didn't find it relevant, partly due to the fact that they thought it was too sophisticated. But what happened in between, somewhere between 2020, 2021, 2022, that time period is that we began to deal with generative artificial intelligence. We began to deal with, so, between 2014 and 2020, there were a couple of interesting developments in terms of AI technologies. One was GANs or generative adversarial networks, which in simple terms means we developed algorithms which were able to speak to other algorithms to accomplish a certain goal. So for GANs, the specific feature was we could take two images and we could map them such that they produced a third required image, which is according to our objectives.

Fast forward, we had the whole idea of we created another technology called transformers, which allowed us to take vast quantities of digital data. Which simply means things like text, let's say Wikipedia. So we can take the text of Wikipedia, we can break it down into unique words, and each word is represented by a number. And then these numbers are fed into the transformer. And the transformer develops for us, and I'm speaking in an oversimplified sense, but still the conceptually what this technology does is it develops a set of probabilistic associations between the words in Wikipedia. And then we bring these technologies along with other machine learning methods together, around 2022, we saw the, we began to experience more popularly the power of what we call as foundation models or large language models. The most popular one was GPT-3 based on which, around November 2022, we saw the release of Open AI's ChatGPT, and in many senses, I think from a public experience of artificial intelligence, this was the tipping point.

So what happened at this tipping point is that suddenly there was a powerful AI application, which was easy to access, easy to use. And which behaved in a way that surpassed certain human intelligence capabilities, at least for some people, not perfectly, but it was at least to start off with, it was amusing, which means now we had an application which anyone could interact with.

It would generate text that didn't exist before. It would do it instantly, and it would do it in a way that was quite useful. So now, not only was it accessible, easy to use, but there's actually now commercial value in it. And subsequently, they improved, ChatGPT quite rapidly through training and that brought AI to the forefront.

Suddenly, everyone woke up and said, we need to understand this technology because the potential is huge, the derivative applications are huge, so on and so forth. So to fully answer that question, I think from a conceptual perspective, I mentioned two very important things that everyone needs to understand.

One is a foundation model. A foundation model is simply a very large baseline artificial intelligence model. You could take all the text, not literally, but let's say sufficiently large quantities of text, Wikipedia and hundreds of thousands of research papers and other text of your choice. So, for example, the AI model that was developed by Elon Musk's AI initiative. They used social media data. They use Twitter data to train their model. So you could, you could train it on any data. You could also train it on visual data. That is, you could collect all the images in the world and train a foundation model with visual data.

Or you could create what's called a multimodal data. That's a foundation model that uses both text and visual image data. This model serves as a general intelligence base. Based on this foundation model, we can now develop a wide range, thousands, tens of thousands of AI applications, all based on a foundation model.

So, companies like Google, Facebook, OpenAI, all of them have their own foundation models. The foundation models, which use only text data, they are called large language models. So, ChatGPT was built on GPT 3 initially and GPT 3X, 3 point something. So GPT 3 is the large language model. The large language model was used to build ChatGPT. And that's the AI application that we interact with.

So today, if my only goal is to create an AI application like a chatbot, then I would go and try to identify one of the existing foundation models, large

language models, such as Facebook's, that's Meta's, 70 billion parameter chat model, and I would use that as a baseline to create my chatbot.

**Lisa:** Okay. And then you mentioned sort of the data portion of this and really since the users are feeding some of these models, then what happens to the data that it is collecting? Are there any concerns as far as the kind of data collection, the rapid data collection that these AI models are able to do?

**Jim:** That opens up a very different direction of questions altogether. So, any artificial intelligence application or foundation model, it's only as good as the data it gets. So data is collected by an AI model or AI systems at various stages. The first stage is when you're building the baseline foundation model and we do not know what kind of data has been, we know we have a--we have some kind of an indication.

So, for example, most of these foundation models tend to use all of Wikipedia. It's public information and so that's fed in. But in addition to that, there's a lot more data that's fed into these large language models and it's led to some controversies. So, for example in the case of vision models, there have been complaints by some companies who own the images that their images were used as raw material for building the large vision models. Now a corporation would be able to raise that as a question, but what we do not know, for example, is has individual data been used?

For example, in the past, when in the pre-artificial intelligence era, before artificial intelligence became popular, when people signed away their rights to data or even made their data open source, they did it in an environment where they were not aware of the fact that artificial intelligence is going to use their data to create models which will mimic their intelligence.

In other words, their approvals were not for AI usage. Their approvals were for general analytics and for general public consumption. So we have an ethical question out there as to, so is this kind of grab-all data approach being used by AI companies, is that ethical? Once this main model is developed, then comes the question of Interaction with the model.

I think when a human interacts with an AI application and uses the feedback data, that's a different kind of data, and there are a couple of other nuances out there. So number one, yes, data. Even in that phase, we are talking about user data being collected by the AI companies and being fed into their systems to improve the performance of the AI models. But what happens out there is the AI application is now already created. Most likely, when you use the AIs, you

have already signed off some rights away, where you have agreed that whatever interaction you have with the AI will be recorded. Whatever feedback you provide will be also recorded and used as input.

So in other words, these companies have already, most likely, I've not looked at all the terms and conditions, but most likely it's common sense that they would build that into their terms of usage.

**Lisa:** Right. And I'm hearing a concern is whether or not, you know, using information that's been extracted is one thing, but how do we differentiate that from using someone's likeness, from using people as data, essentially. I'm wondering if you have any advice on how companies that are creating AI models can really provide transparency.

**Jim:** I think my suggestions in this area are quite radical. But not uncommon. I think there are many people who are calling for a simple but radical approach to artificial intelligence. That is, make all data input into artificial intelligence completely transparent. Declare 100 percent of the data that is going into an AI model.

And it's very simple. The logic is very simple. Any product out there, we see that it has a product label, which by law requires 100 percent of the ingredients to be listed. If you buy medicine, if you buy a food product, whatever you buy, the consumer has a right to know what are the ingredients that have gone into that product or service. In this case, we just don't have visibility into what data has been used to build these models.

So that's one side of what I, what I understand your question was. The other side of the question is in terms of the likeness of people and in terms of AI companies building products that mimic human likeness. That's a separate area of. And I think AI companies are going to be very careful when it comes to building exact copies. I think that, can they do it? Yes. Whether it is visual. So for example, it would be possible for artificial intelligence companies to take images and videos of an actor or an actress. Yes. And then use that to create an entirely new video as though it were that actor who is acting it out. I don't think that's going to fly, not without them having an agreement with that particular actress or actor.

But what will happen is, or what will and what is, what I've already seen happening is that these technologies are used to create entirely new models, so to say. And they can then use these artificial agents and replace human models,

human actors, and actresses. And at present I believe that we are seeing some of, some regulations, some laws being put into place as early-stage guidelines.

They're more in the form of guidelines, agreements, strategies, than in the form of well-informed laws. So I think that's, again, an open area. So we are, we are at the cusp of the AI revolution. So as these products grow in strength, as we see more and more capabilities being rolled out, I think, that's when we will be able to have better visibility of the AI applications landscape and how we would need to work with it.

**Lisa:** And as far as the laws that are creating regulations around artificial intelligence, are there any aspects of artificial intelligence use that you see that are being addressed or targeted through policy at the moment?

**Jim:** I think there's been a huge rush over the past, especially over the past year. You know, as I said, sometimes policies and regulations, they follow public perception, public awareness, and also business interest. As business interest in artificial intelligence has spiked, and so has public concern, we have seen a flurry of regulatory and policy and guidelines initiatives.

At this point in time, I think the intentions are good. I think, you know, I was one of the persons who called for AI regulations early on. You know, so we call for AI regulations, we got AI regulations now. But what we call for and what we are seeing now, I think there's a huge gap.

I think the current regulations are well intentioned, but there are a couple of challenges. And the first thing is the motivation for AI regulation. Why are we regulating artificial intelligence? This is a very critical question. If we are going to follow a narrative of AI is going to destroy all human beings and therefore we need to regulate AI, then that's the wrong motivation.

Artificial intelligence is just a technology. A foundation model, even the best one, it's just about as dangerous as statistics is. Statistics by itself is not going to destroy the human race. Similarly, artificial intelligence by itself is not going to destroy the human race. It is not, of itself, a threat to the human race.

People could use statistics or even our knowledge of biology for that matter to create threats to the human race. And in a similar manner, just as you could use biology or statistics, you could also use artificial intelligence to create threats to the human race. But by itself, it's not a threat. So that should not be a motivation for regulation.

Our motivation for regulation should be based on the fact that artificial intelligence technologies are very powerful technologies. For example, they can be used to replace human labor. These technologies are built based on data collected from human beings, and then based on that data, we are modeling intelligence, and then we are using that intelligence to replace the people from whom we collected the data, and that creates ethical questions.

So I think the motivation for artificial intelligence regulations should be to give power to the people, and to give power to the people, the goal should be to regulate the large companies who currently own these artificial intelligences. They must be compelled to be transparent. I think that's one of the first steps because transparency is critical.

We need to know what's going into these artificial intelligence models. We need to know how is it that they are tweaking and fine tuning these models, otherwise these models can be used for mass manipulation of human minds and, and, and that's not something, that's not the kind of society we want to see developing, but now it's a possibility.

**Lisa:** Well, you mentioned the companies that own these AI models. So right now, as far as the sourcing for AI, it is not, is it not acceptable for the public to make their own artificial intelligence model, or are companies, sort of the owners behind that code?

**Jim:** So the foundation models are owned, that's these large language models, are owned by large corporations, for the most part. There are, there are a few which are, semi open source, not completely open source, but Facebook, Meta, for example, makes their models available for us to work with, but we, it's not completely open source because for example, they have not revealed all the data that's gone into their model and they have not revealed the complete process which they used to create the model.

The model itself is available for download. GPT, in contrast, it's completely black box, we are able to use ChatGPT, but we don't see the weights, for example, of the underlying GPT model, as we see with Meta's models. There is a resource barrier for small companies and for ordinary people to create these large language models. These large language models require high performance computers, that is supercomputers, for us to train them. They require specialized expertise in code. They require large quantities of data to be collected and processed using these supercomputers. That is out of reach for most small companies, and that's why there's a resource barrier.

Once, once these large models are built, then I think any company could leverage those. Even small companies could take those models and leverage them. So, to create these large models, there are barriers, but to create the applications based on those large models, I think there's not as many barriers and pretty much anyone with a keen mind and an interest in working in that space could do so.

**Lisa:** I definitely appreciate you walking everyone off the ledge of this AI-pocalypse that everyone has in mind right now, especially with hysteria around what exactly is AI going to be used for moving forward. And I want the listeners to know that you do have a policy brief, within our own website of the Scholars Strategy Network that really delves into these different models. and especially this need for transparency and regulation, that you've mentioned here. And I do want to ask you, do you have any closing thoughts here on the future of artificial intelligence as a whole?

**Jim:** Well, artificial intelligence is here to stay. Just as we got electricity and electricity is not going anywhere, just as we have elements of the industrial revolution that came in, so also artificial intelligence is here to stay. Our lives, pre-2022 and post-2022, it's going to be quite different.

We're going to see an increasing adoption of artificial intelligences by individuals, companies, governments, and human society at large. So for the future, what I see is individuals being augmented by artificial intelligence who were able to perform a lot more work than they would otherwise be able to. Or another perspective or another way to look at is AI augmented humans doing work much quicker, which means if I needed to put in 40 hours of effort into a small project with AI, I may be able to do that in 10 hours, maybe less.

The other part is there is a big area which has not been addressed. And that is, we are developing technologies that will replace human intelligence, which simply means a lot of people are going to be out of work. Bill Gates recently said that we may be able to look forward to three-day work weeks.

So that simply means that 40 percent of our labor will not be required. Apart from that, I believe there are jobs which can be replaced 100 percent by AIs. So we are talking about millions and potentially tens of millions of jobs being taken over by AI. Then the question arises, what happens to those people who will be replaced? Will the government bring in some kind of a UBI, Universal Basic Income, or some kind of a scheme to support them.

So the future is complex. On one side, there's going to be tremendous benefits in every area, in medical care, in personal experience of use of AIs for a broad range of activities, smart cities, smart everything.

And then on the other side, we have a few challenges. The risks associated with AI technologies. The risk of bad players using these advanced technologies. The risk of job loss. So, it's going to be interesting. It's too early to predict what will happen. But what we can say is we are entering into a time of potential for great good and potential for great evil.

How it plays out, a lot depends upon all of us humans, as human society, taking collective responsibility. And the last one thing that I want to emphasize more than anything else is we need AI education, and I'm not talking about the use of AI technologies in the process of education. I'm talking about actually teaching what artificial intelligence is and what it can do at every level, right from school, college, professional education. Everyone needs to study what AI is.

**Lisa:** Well, we definitely appreciate your education today on artificial intelligence. And thank you for that overview of what our AI future might look like.

**Jim:** Thank you, Lisa. I appreciate it.

**Lisa:** Thanks again, Professor Samuel. And thanks for listening. For more on Professor Samuel's work, check out our show notes at scholars.org/nojargon. No Jargon is the podcast of the Scholars Strategy Network, a nationwide organization that connects journalists, policymakers, and civic leaders with America's top researchers to improve policy and strengthen democracy.

The producer of our show is Mandana Mohsenzadegan. Our audio engineer is Peter Linane. If you like the show, please subscribe and rate us on Apple Podcasts or wherever you get your shows. You can give us feedback on Twitter or X @NoJargonPodcast, or at our email address, nojargon@scholars.org.