Medical News & Perspectives | AI IN MEDICINE

How Do Policymakers Regulate AI *and* Accommodate Innovation in Research and Medicine?

Melissa Suran, PhD, MSJ; Yulin Hswen, ScD, MPH

This conversation is part of a series of interviews in which JAMA *Editor in Chief Kirsten Bibbins-Domingo, PhD, MD, MAS, and expert guests explore issues surrounding the rapidly evolving intersection of artificial intelligence (AI) and medicine.*

What are the most recent advancements in establishing AI safeguards for clinical practice? In what way does AI intersect with democracy and its preservation? And how are the frameworks for regulating AI progressing and aligning across the US, UK, and EU? As the technology advances at lightning speed, such questions surrounding AI become more critical.

Multimedia

Medical News website

Alondra Nelson, PhD (**Video**), is focusing on effective guardrails that protect society from issues like data insecurity—but also encourage innovation in the laboratory and clinic. Nelson is the Harold F. Linder Professor at the Institute for Advanced Study in Princeton, New Jersey, where she studies the effects of scientific and technological advances on health and society. In 2023, she was included in *TIME* magazine's 100 most influential people in AI.

JAMA Editor in Chief Kirsten-Bibbins Domingo, PhD, MD, MAS, recently spoke with Nelson, who also served as deputy assistant to US President Joe Biden and was acting director of the White House Office of Science and Technology Policy (OSTP).

The following interview has been edited for clarity and length.

DR BIBBINS-DOMINGO: You have been a scholar of scientific advancements and their impact on society, often through an equity lens, long before we were talking about AI all the time. Could you tell us how you became interested in this field, and how you ended up going from an academic position to serving in the White House?

DR NELSON: Probably for the last 15 years of my career, I have been a scholar of human genetics. It was through following the emergence of the Human Genome Project, which was a big science project, but also what we would now call a big data project. It was a partnership with the NIH [the National Institutes of Health] and with the US Department of Energy because the computing needs were massive.

What the completion of the Human Genome Project did was create a data opportunity. Working in the field of the social implications of human genetics meant thinking about big data issues: we have all these data, now what do we do with them? How do you begin to analyze these data to offer the outcomes that these big scientific projects promised the public with regards to yields, better health care, and better diagnostics? So that's how I came to work on what we used to call "big data." And of course, big data are one of the enabling resources of AI. Data, to me, have always been about human beings, and so my perspective always has been through an equity and equality lens.

The initial bridge for me was that, as part of his State of the Union Address in 2015, former US President Barack Obama announced the Precision Medicine Initiative. In my research career, this was following the trajectory of human genome research and these sorts of big projects. I was interested in the precision medicine project and was invited to work in the Biden-Harris White House at the OSTP.

DR BIBBINS-DOMINGO: I love the idea that data are always about people. We're at about the year anniversary of ChatGPT's introduction, and it seems that there's no shortage of information in the news about what AI is going to mean for us in the future. What I'd love for you to talk about is your most recent experience in trying to create policy frameworks around this and to reflect a little bit about the Biden administration's recent executive order on artificial intelligence. What was the goal of this order, and what are its most important features?

DR NELSON: Going into the OSTP at the White House in 2021, it was clear that AI policy was going to be a big policy space. There was a national AI initiative office that sort of sits in OSTP. There were plans being made to create a task force and an initial strategy for something called the National AI Research Resource to see if the government can create a parallel ecosystem to help researchers in R and D [research and development] have the capacity they need to do innovative research in that space.

I come from the perspective that creating guardrails are opportunities for creativity and innovation, and that they don't

staunch or stop innovation. The creativity that we have as researchers often comes because we need to think anew, because we're trying to mitigate risks or think about new challenges. So I think when ChatGPT came on the scene about a year ago, it accelerated all of the work that already had been going on and that people had been talking about. It was also clear that after the last Congress, we got very close to having a federal data privacy law, but we didn't quite make it. So part of what you see reflected in this executive order is an attempt to work at the executive level—appreciating that we're not going to have legislation for guardrails, but that an executive order has the force of law even if it's not legislation.

And what's fascinating about this executive order is that it takes the broad scope and potential of AI seriously: it has things that are about employment. There's lots there about health care. There are interventions around both economic security and national security as well as privacy; algorithmic bias and discrimination as well. So it's really an executive order that appreciates that AI tools and systems are increasingly going to become the infrastructure for our lives and how we do research.

So how do we protect privacy or perhaps have regulations around even minimal use of data that one needs, particularly for medicine, so we don't scare off patients or health care workers from wanting to use the tools and systems because they feel threatened? How do you involve people in these processes, and how do we think about processes at the FDA [US Food and Drug Administration], HHS [US Department of Health and Human Services], and other places in the context of appreciating the substantive transformation that's going to happen with AI? So this executive order is really an attempt to use the whole-of-government approach to address the potential risks of AI and its potential benefits.

DR BIBBINS-DOMINGO: I like the way you said that the guardrails are not there to impede opportunities but rather to unlock many creative opportunities. How would this have looked if we had a different data privacy landscape? You started by saying that Congress didn't quite make it to data privacy laws.

DR NELSON: Right now, we have a kind of patchwork. Illinois, for example, has legislation around biometric data, some of which would apply to health care. California has a privacy law. Federal regulation would have created more standard rules of the road for all systems. I worked in the White House on something called the Blueprint for an AI Bill of Rights, which had principles for how we want to do this well. For example, systems should be effective, and they should be tested. We should be thinking iteratively about coming back to do risk assessment, auditing of systems, and the like. The executive order also includes a set of principles for how government in particular is going to engage in using AI tools and systems.

DR BIBBINS-DOMINGO: So what AI tools are you currently using, and what do you use them for?

DR NELSON: I'm using ChatGPT. As somebody who's interested in being a policy advisor, I'm playing around with it to understand what it can do and not do. I'm actually quite concerned about the 2024 elections. They're happening in many major democracies—from Indonesia to India to the United States. How accurately do [these tools] synthesize election information? You and I are from California; we have a whole history of referendums that are so long and have gotten longer. And if you pop those into ChatGPT, how does it synthesize and how accurate is it if you go through and read the whole thing? Does it summarize correctly? So I'm interested in playing around with its possibilities and potentials. And of course, I was intrigued when OpenAI announced that it's going to have an app store, so we're about to see a whole bunch of applications to layer on top of that.

DR BIBBINS-DOMINGO: Is there an AI tool that you would never use at this point?

DR NELSON: I think biometric tools. Folks in clinical practice know this very well: if you have someone's genome, you have someone's genome. You can't erase it; it doesn't go away. So tools that take genetic data, fingerprints, irises, and these sorts of permanent identity markers of people. Of course, we have a big problem with facial recognition technology in the United States and policing. So I would never use it, and I think there are some instances in which no one should ever use them.

DR BIBBINS-DOMINGO: I really enjoyed reading your editorial in *Science* about how it's easy to become cynical about the constant refrain concerning safety. Should we be worried about robots taking over or should we be more worried about everyday ways in which the application of these technologies are already shifting what we're doing—whether it's in health care or other aspects of our life? And how do we pitch the guardrails in a way that allows us to move forward while still recognizing that we need to have things in place now for everyday ways in which we are either not meeting our goals or opening other avenues that might have negative consequences?

DR NELSON: I think this is the moment that we find ourselves in, and I do think that the executive order attempts to do both. For example, it makes certain assertions about how there should be protections against algorithmic bias and discrimination; things that are causing harm to people right now, even as we are looking at future risks that we might have with regards to bioweapons and synthetic biology; the kinds of potential risks that the marriage of engineering and biological research potentially make possible. With that, I wrote the editorial in *Science* with my colleague Seth Lazar, who is a philosopher at the Australian National University. Partly, what we were trying to assert was that the AI safety conversation had shifted over the last year. There had been a conversation in years past about trust and safety. So there was an expectation that many of the particularly big tech organizations would have trust and safety teams trying to ensure that their products were not damaging the public and content moderation.

As we came to have a national conversation that included, at its worst, a bit more fear mongering about what AI might mean and the idea of safety, a couple of things happened: one was some of these very large companies shut down and reduced their trust and safety teams. So what they had imagined safety to be before somehow was not safety as it was previously imagined. And as we head into an election year, content moderation and challenges around the ability for AI to increase the scale and velocity of some of the things that we're already concerned about in social media still remain a challenge. So we still need trust and safety teams. At the same time, there was a new elite crop of safety research labs emerging and AI labs that initially wanted to create a pretty narrow lane of what safety meant for their work. And safety was, frankly, robots not killing us, and not leading to total

extinction. And I think we can imagine government, large organizations, and small organizations being able to think about the whole spectrum of safety and not having to make choices about whether things should be within that bucket of safety.

DR BIBBINS-DOMINGO: I really like that point you make in the editorial, that we don't want to leave it to just those who understand the technology or who are actively innovating the technology to set the priorities or to define the agenda for safety, given its impact on so many parts of our life. There need to be many other voices at the table. I think that applies particularly for the application of AI in clinical practice, where many people have pointed out that this isn't the time for clinicians and others who think about what's important in health care to stand on the sidelines and say, "We don't really understand everything behind the technology." We all have to have a role to play in setting these priorities. I'd love for you to reflect a little bit on how other countries like the UK and in the EU have thought about creating a regulatory framework or environment for AI. I've heard many people describe this as being different in the EU than in the US.

DR NELSON: I think that question changed after October 30, 2023, when President Biden signed this executive order. A lot of the conversation about the differences was what's the US going to do, if anything at all? And we were waiting on Congress, vs this robust multiyear system that had been built out by the EU. The EU had paused a bit, and over the course of the last year, it had gone from having a risk-based model—so there's AI that's high risk, medium risk, low risk, and there would be more regulations for things that are at different risk categories—to adding consideration of large language models and other kinds of AI, generative AI, that hadn't been imagined as the regulatory object when the process that led to the initial drafting of the EU AI Act was getting under way. The US doesn't have formal regulation, so the EU work is already more stringent. It's certainly multilateral: the EU is 27 member states, so creating some sort of synergy or harmonization among all of those different countries is significant and notable. But now, I think in response in part to President Biden's executive order, there is some concern in Europe around innovation.

And to the point that we were just talking about, is there a tipping point at which innovation regulation prevents innovation? So I think [the EU is] wrestling right now with that problem. I would say it will have the force of regulatory law when it's completed that's much stronger than anything that we have in the US. As things stand, it's very much making the case that there are different risk categories. The challenge that something like generative AI presents is that if it's a general-purpose tool, you don't know at the front end what the risk category is because it's not been used yet—and it might be used for a whole range of things.

For people who want to work in health care policy, tech policy, or their intersection, I think it presents a strategic conundrum. The FDA says this is a medical device, and if it designates something as a medical device, then all of these other laws, norms, and procedures fall into place. For example, a large language model: Is that a medical device or not? How do you think about that? And do we need to shift the world of generative AI from thinking less about the specific device to more about the outcomes that we want? So whether or not you're using an artificial tool or system across a wide spectrum—it could be predictive, it could be generative AI—the outcome that we want is for it to be safe, to be proven effective, and to have some transparency, particularly in clinical practice.

So I think the EU is one model. What we saw in the UK was another model. The UK is suggesting that they're going to do what they're calling "light-touch regulation," so not much regulation at all, and instead, they'll sort of dig into what they're calling "safety" with these AI safety institutes, which are private-public, sort of R and D. Safety needs to be this broader spectrum, and there are things that we need to make people's lives better right now using these tools.

DR BIBBINS-DOMINGO: You've been a scholar of new technologies, genotyping, and the direct-to-consumer marketing of technologies like genetic testing. Is AI something that requires us to think in different ways about setting up the guardrails? Or is this some variation on the same story as what we've learned from genetic testing?

DR NELSON: There is so much hype around AI, and you and I were around when the Human Genome Project genotyping hype cycle happened as well. There was lots of interesting research and shifts in diagnostics. To the extent that we might have a cure for something like sickle cell anemia, which is something that I've worked on as a sociologist, it is sort of proof of some of that hype. However, there was also a lot of hype that's not been proved out. But I've been encouraged that there are people who are not AI scientists willing to say this isn't magic, and there are things about this that are similar to things that we've had to deal with in the past. But there are things that are also new.

So much of this is being driven by commercial interests and released before we can fully understand it. Last November, part of the awesomeness of these tools was that they were cool, they were fun, and they were interesting—but it was also the case that no one knew these tools were going to be released and were going to go through all of these transformations. We now can anticipate, in part because of the disruptive business cycle in places like Silicon Valley, that there's going to be this constant churn of new technologies. We need to revisit the core values and the core outcomes that we want with new and emerging technologies and not try to make regulations and guardrails every time for every new tool and technology.

DR BIBBINS-DOMINGO: There's a tendency to get caught up in that: this is all new, and we're completely unmoored from anything that we've encountered before. And yet, in the end, it still comes down to—as you said—our values. And then I think in the clinical practice space, or even in the journal space, we're looking for improvements in clinical outcomes; we're looking for improvements in health. We're trying to ensure that the inequities that we know already exist are not worsened and hopefully are improved. So in the end, we have to come back to those basic values.

DR NELSON: That's the North Star of clinical practice in some ways, regardless of what the technology is. ■

Published Online: December 20, 2023. doi:10.1001/jama.2023.22625

Conflict of Interest Disclosures: Dr Nelson reported being a consultant to Ergo, an intelligence and advisory firm, on AI policy and governance.

Note: Source references are available through embedded hyperlinks in the article text online.