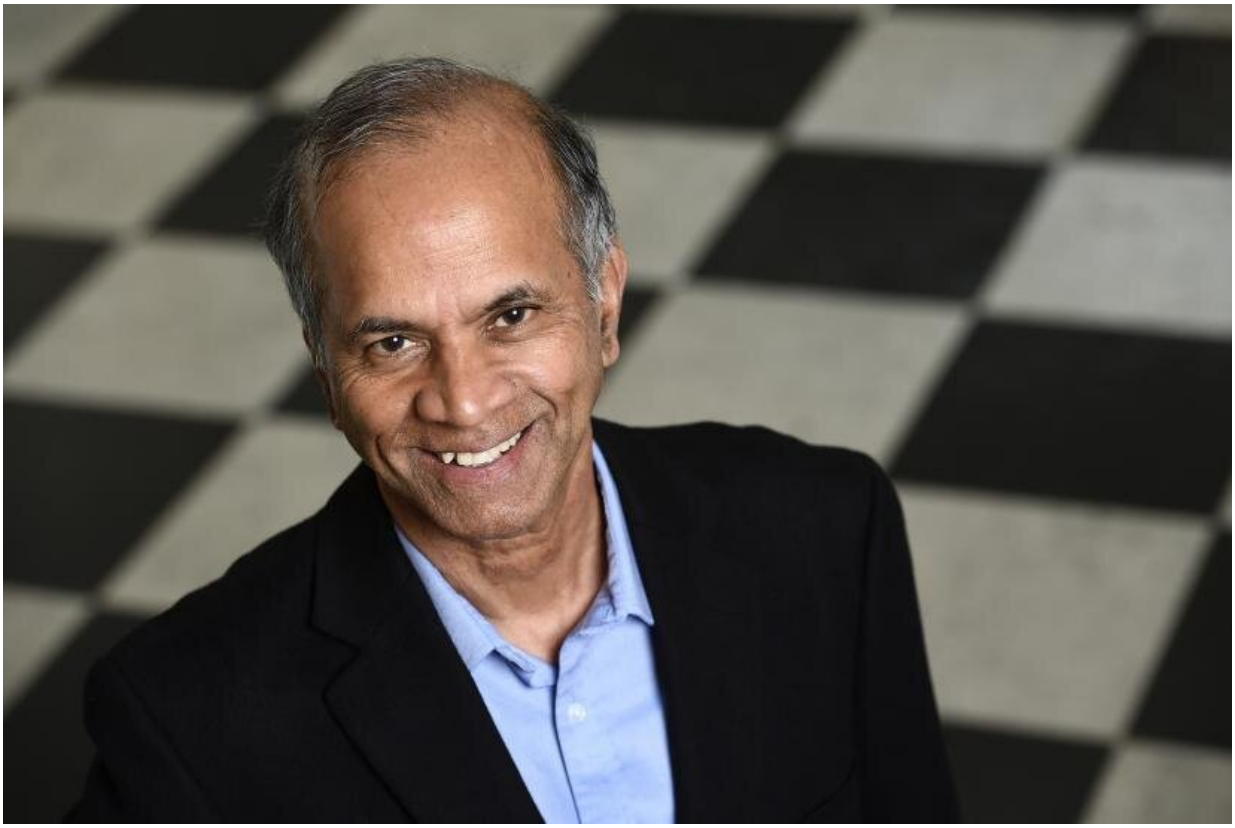
Q&A: Can we trust AI?

March 6 2023, by Katie Pearce



Rama Chellappa. Credit: Will Kirk / Johns Hopkins University

In the field of artificial intelligence, there are winters and there are springs—barren stretches followed by exhilarating bursts of innovation and funding. Right now, we find ourselves definitively in the midst of an AI spring, says Johns Hopkins engineer Rama Chellappa, a veteran of the industry for more than four decades.

"AI has this crazy life where it goes through cycles," says Chellappa, a Bloomberg Distinguished Professor in electrical, computer, and biomedical engineering. "Some of the algorithms we're using today have been around since the 1960s, but what's happening now is due to the explosion of data we have available to drive our systems."

Chellappa says he respects all of the apprehensions surrounding AI—privacy and hacking concerns, ethical quandaries, stark evidence of its biases, and of course "the Hollywood scenario of robots coming in and killing us all," he says with a laugh.

The reason he wrote his book, "Can We Trust AI?" with co-author Eric Niller, he says, was to address these challenges while also pointing to the overall "net positive" he believes AI can bring to human society. "This technology should only help us better ourselves and improve our quality of life," he says.

Published by JHU Press in November 2022, the book tours the field of AI from its post-World War II origins through the computer revolution of 1960s and 1970s up to our current "AI spring," as the global market for AI enterprises is predicted to top $228 billion by 2026. Chellappa recounts his own experiences in the early days of machine learning and computer vision, shares wisdom from expert colleagues, and looks ahead to promising uses in medicine, self-driving cars, and public safety, among other arenas.

Freshly elected to the National Academy of Engineering, Chellappa, who is affiliated with the Center for Imaging Science, the Center for Language and Speech Processing, the Institute for Assured Autonomy, and the mathematical Institute for Data Science, recently spoke with the Hub about his book, his current work in AI, and what he thinks of Alexa.

In your book you say AI is at 'the toddler phase' right

now. What do you mean by that?

Of course, in many applications AI can go further than a toddler, but I'm referring to its reasoning capabilities. Interestingly, AI in its early years was driven by domain knowledge and inference capabilities. Current AI designs, which are mostly driven by "big data," appear to have short-changed the reasoning and domain knowledge needed for making decisions.

It takes a while for humans to develop common-sense reasoning, right? You learn from experience. Two year olds learn from examples—you show them what a cup looks like, they can identify other cups. AI can do that but will need a lot more data. If you ask a two year old where the cup is and you put a towel over it, the kid will likely pull out the towel and say, "There it is!" AI may not necessarily do this with the same efficiency yet. Toddlers can also imagine, make inferences and comparisons. While generative AI models are able to synthesize new images, videos, text, and language, humans are much better at imagining "what if" scenarios.

AI systems are getting smarter, though, with what we call self-supervised learning, and they're learning to make these kind of connections.

How do you use AI in your own daily life? For example, do you use Alexa?

I think those devices are interesting, but really any information I need I can just find on the web. One of my former students gave me a Google Home and I kind of played around with it, playing music and all that. My son has Alexa and we ask it questions.

But in terms of AI I'm actually using, they're the systems that are now

integrated naturally into all of our web and smartphone experiences. For example, when you buy a book on Amazon and it recommends similar books you may like, or when you watch three movies of the same genre on Netflix and it picks up the pattern. The Google map I use all the time is based on a classical AI search algorithm. Some people are wary of this kind of thing, saying, "Oh, no, it's tracking me and controlling me," but really they're just suggestions. They're not forcing you to watch something or get out your credit card.

What type of AI projects are you and your team working on right now at Hopkins?

We're tackling AI from a lot of different angles. We're working on protecting AI from adversarial attacks. We're designing algorithms that will ensure biases are reduced. We're working on AI-driven applications for various facets of medicine—interacting with the Center for Autism and a cancer unit, for example, and working on problems related to pathology, healthy aging, and eye care.

We're also looking at some traditional computer vision problems—making face recognition or human recognition work at up to a thousand meters. And it's not just faces, we're looking to improve recognition of the body and gait, such as the way people walk.

Another new project I'll be involved in is modeling various sites of the world and visualizing them at various times of the day or different seasons, which could be useful for rescue missions or things of that nature.

In collaboration with the Applied Physics Laboratory, we're exploring the effectiveness of synthetic data—artificially generated data that imitate real-world scenarios—in designing AI systems. Synthetic data

doesn't need to be annotated as we generate it, and it largely reduces privacy concerns. More importantly, it will enable AI to imagine better and generalize well to new environments and situations. As I like to say, sensors can only capture what has happened; synthetic data can reflect "what if" scenarios, leading to generalizable AI systems. As a lifelong practitioner of generative approaches to computer vision, I'm elated by this possibility. However, we can't let that imagination go wild and affect the integrity of our AI systems.

Can you talk more about how AI can help with aging? I know you're part of the AI & Tech Collaboratory for Aging Research at Hopkins.

This group is funded by the National Institute on Aging, bringing together clinical and engineering expertise. The basic premise is to put useful AI devices and systems into the hands of elderly people and their caretakers, to improve quality of life and safety and longevity. For example, we're envisioning robots that could interact with patients with cognitive impairments, dementia, or Alzheimer's, and help them go about their daily business. We're exploring using Alexa to administer cognitive tests at home, and using Apple Watches to provide alerts of possible falls or wandering.

This is one of the big reasons I came to Hopkins, with the idea of exploring AI and machine learning for medicine, and I'm really excited by the collaborations happening.

Since your book came out, we've seen the explosion of ChatGPT. What do you think of the hype around that and where do you see that tech heading?

ChatGPT is the latest development in what's known as large language models. While ChatGPT marks an important milestone, we should be careful of how it's being used. There are many examples now of how ChatGPT just makes stuff up! That's worrisome. I hope with more training, we'll see improvement.

Why do you believe, as you say in your book, that AI will have a net positive impact on society?

As humans we have to make a lot of decisions and AI can help take the load off.

Let me give an example from medicine. We'll reach a point where AI can look comprehensively at all of our electronic health records and diagnostic images, all the medications we've taken, and can even process our patient-to-doctor conversations. And that will build a very personalized profile that can help us monitor our health, alert doctors to our individual needs, and predict any potential issues in the future.

Let's look at self-driving cars. The fully automated car is talked about as the ultimate goal, and we're short of that, but in the meantime, we're developing so many different features to help us—cameras showing us what's behind us when we reverse, alerts of cars in adjacent lanes, etc. These features will ultimately lead to fewer accidents and fewer deaths.

So I think AI can be kind of a friend at your side, and it's left to us to determine how much help we need from it. You know, sometimes our friends in real life don't behave well. AI likewise can make mistakes—it's an algorithm and can only do what we've taught it to do.

This idea that AI can run amok and become a monster or terminator, that's mostly Hollywood stuff but I also know some serious AI people

who also find that type of scenario credible. Personally, with the technologies that are known to us right now, I just don't see anything like that happening.

So to answer the question your title poses, 'Can We Trust AI?,' your answer is—

It's a yes. (laughs) It's a qualified yes.

I like to say, let's not talk about humans versus AI. Let's talk about humans and AI. We can identify problems and work to improve them. We can make AI work for us and with us.

Provided by Johns Hopkins University