Illustration by Shonagh Rae

"AI is a tool, and its values are human values."

Computer scientist and "godmother of AI" Fei-Fei Li explains why artificial intelligence and public life are at an inflection point—and contemplates how to unleash positive changes while mitigating risks.

Fei-Fei Li has been called the godmother of AI for her pioneering work in computer vision and image recognition. Li invented ImageNet, a foundational large-scale dataset that has contributed to key developments in deep learning and artificial intelligence. She previously served as chief scientist of AI at Google Cloud and as a member of the National Artificial Intelligence Research Resource Task Force for the White House Office of Science and Technology Policy and the National Science Foundation.

Li is currently the Sequoia Professor of Computer Science at Stanford University, where she cofounded and codirects the Institute for Human-Centered AI. She also cofounded the national nonprofit AI4ALL, which aims to increase inclusion and diversity in AI education. Li is a member of the National Academy of Engineering and the National Academy of Medicine, and her recent book is *The Worlds I See: Curiosity, Exploration, and Discovery at the Dawn of AI.*

In an interview with *Issues* editor Sara Frueh, Li shares her thoughts on how to keep AI centered on human well-being, the ethical responsibilities of AI scientists and developers, and whether there are limits to the human qualities AI can attain.

What drew you into AI? How did it happen, and what appealed to you about it?

Li: It was a pure intellectual curiosity that developed around 25 years ago. And the audacity of a curious question, which is: What is intelligence, and can we make intelligent machines? That was just so much fun to ponder.

My original entry point into science was physics. I was an undergrad in physics at Princeton. And physics is a way of thinking about big and fundamental questions. One fun aspect of being a physics student is that you learn about the physical world, the atomic world.

The question of intelligence is a contrast to that. It's so much more nebulous. Maybe one day we will prove that it's all just physically realized intelligence, but before that happens, it's just a whole different way of asking those fundamental questions. That was just fascinating. And of all the aspects of intelligence, visual intelligence is a cornerstone of intelligence for animals and humans. The pixel world is so rich and mathematically infinite. To make sense of it, to be able to understand it, to be able to live within it, and to do things in it is just so fascinating to me.

Where are we at in the development of AI? Do you see us as being at a crossroads or inflection point, and if so, what kind?

Li: We're absolutely at a very interesting time. Are we at an inflection point? The short answer is yes, but the longer answer is that technologies and our society will go through many inflection points. I don't want to overhype this by saying this is the singular one.

So it is an inflection point for several reasons. One is the power of new AI models. AI as a field is relatively young—it's 60, maybe 70 years old by now. It's young enough that it's only come of age to the public recently. And suddenly we've got these powerful models like large language models—and that itself is an inflection point.

The second reason it's an inflection point is the public has awakened to AI. We've gone through a few earlier, smaller inflection points, like when AlphaGo beat a human Go player in 2016, but AlphaGo didn't change public life. You can sit here and watch a computer play a Go master, but it doesn't make your life different. ChatGPT changed that—whether you're asking a question or trying to compose an email or translate a language. And now we have other generative AI creating art and all that. That just fundamentally changed people, and that public awakening is an inflection point.

And the third is socioeconomic. You combine the technology with the public awakening, and suddenly many of the doings of society are going to be impacted by this powerful technology. And that has profound impacts on business, socioeconomic structure, and labor, and there will be intended and unintended consequences—including for democracy.

Thinking about where we go from here—you cofounded and lead the Institute for Human-Centered AI (HAI) at Stanford. What does it mean to develop AI in a human-centered way?

Li: It means recognizing AI is a tool. And tools don't have independent values—their values are human values. That means we need to be responsible developers as well as governors of this technology—which requires a framework. The human-centered framework is anchored in a shared commitment that AI should improve the human condition—and it consists of concentric rings of responsibility and impact, from individuals to community to society as a whole.

For example, human centeredness for the individual recognizes that this technology can empower or harm human dignity, can enhance or take away human jobs and opportunity, and can enhance or replace human creativity.

And then you look at community. This technology can help communities. But this technology can also exacerbate the bias or the challenges among different communities. It can become a tool to harm communities. So that's another level.

And then society—this technology can unleash incredible, civilizational-scale positive changes like curing diseases, discovering drugs, finding new materials, creating climate solutions. Even last year's fusion milestone was very much empowered by AI and machine learning. In the meantime, it can really create risks to society and to democracy, like disinformation and painful labor market change.

A lot of people, especially in Silicon Valley, talk about increased productivity. As a technologist, I absolutely believe in increased productivity, but that doesn't automatically translate into shared prosperity. And that's a societal level issue. So no matter if you look at the individual, community, or society, a human-centered approach to AI is important.

Are there policies or incentives that could be implemented to ensure that AI is developed in ways that enhance human benefits and minimize risks?

Li: I think education is critical. I worry that the United States hasn't embraced effective education for our population—whether it's K–12 or continuing education.

Research Resource, which would provide the public sector and the academic world with desperately needed computing and data resources to do more AI research and discovery. And that's part of policy as well.

And of course there are policies that need to look into the harms and unintended consequences of AI, especially in areas like health care, education, manufacturing, and finance.

You mentioned that you've been advocating for the National AI Research Resource (NAIRR). An NSF-led pilot of NAIRR has just begun, and legislation has been introduced in Congress—the Create AI Act—that would establish it at full scale. How would that shape the development of AI in a way that benefits people?

Li: The goal is to resource our public sector. NAIRR is a vision for a national infrastructure for AI research that democratizes the tools needed to advance discovery and innovation. The goal is to create a public resource that enables academic and nonprofit AI researchers to access the tools they need—including data, computing power, and training.

And so let's look at what public sector means, not just in terms of AI, but fundamentally to our country and to

As a technologist, I absolutely believe in increased productivity, but that doesn't automatically translate into shared prosperity.

A lot of people are fearful of this technology. There is a lack of public education on what this is. And I cringe when I read about AI in the news because it either lacks technical accuracy or it is going after eyeballs. The less proper education there is, the more despair and anxiety it creates for our society. And that's just not helpful.

For example, take children and learning. We're hearing about some schoolteachers absolutely banning AI. But we also see some children starting to use AI in a responsible way and learning to take advantage of this tool. And the difference between those who understand how to use AI and those who do not is going to have extremely profound downstream effects.

And of course, skillset education is also important. It's been how many decades since we entered the computing age? Yet I don't think US K–12 computing education is adequate. And that will also affect the future.

Thoughtful policies are important, but by policy I don't mean regulation exclusively. Policy can effectively incentivize and actually help to create a healthier ecosystem. I have been advocating for the National AI

our civilization. The public sector produces public goods in several forms. The first form is knowledge expansion and discovery in the long arc of civilizational progress, whether it's printing books or writing Beethoven's Sixth Symphony or curing diseases.

The second public good is talent. The public sector is shouldering the education of students and continued skilling of the public. And resourcing the public sector well means investing in the future of these talents.

And last but not least, the public sector is what the public should be able to trust when there is a need to assess, evaluate, or explain something. For example, I don't know exactly how ibuprofen works; most people don't. Yet we trust ibuprofen to be used in certain conditions. It's because there have been both public- and private-sector studies and assessments and evaluations and standardizations of how to use these drugs. And that is a very important process, so that by and large our public trusts using medications like ibuprofen.

We need the public sector to play that evaluative role in AI. For example, HAI has been comparing large language models in an objective way, but we're so resource-limited. We wish we

could do an even better job, but we need to resource the public sector to do that.

You're working on AI for health care. People think about AI as being used for drug discovery, but you're thinking about it in terms of the human experience. How do you think AI can improve the human experience in our fractured, frustrating health care system? And how did your own experience shape your vision for that?

Li: I've been involved in AI health care for a dozen years—really motivated by my personal journey of taking care of an ailing parent for the past three decades. And now two ailing parents. I've been at the front and center of caring—not just providing moral support, but playing the role of home nurse, translator, case manager, advocate, and all that. So I've seen that so much about health care is not just drug names and treatment plans and X-ray machines. Health care is people caring for people. Health care is ensuring patients are safe, are getting adequate, timely care, and are having a dignified care process.

involved in using AI to develop a drug that is illegal and harmful for people, for example. Most people won't. So there's a lot, from individual values to professional norms to laws, where we have responsibility.

But I also feel we have a little bit of extra responsibility at this stage of AI because it's new. We have a responsibility in communication and education. This is why HAI does so much work with the policy world, with the business world, with the ecosystem, because if we can use our resources to communicate and educate about this technology in a responsible way, it's so much better than people reading misinformation that creates anxiety or irresponsible expectations of utopia. I guess it's individual and optional, but it is a legit responsibility we can take.

When you think about AI's future, what worries you the most, and what gives you hope?

Li: It's not AI's future, it's humanity's future. We don't talk about electricity's future, we don't talk about steam's future. At the end of the day, it is our future, our species' future, and our civilization's future—in the context of AI.

So the dangers and the hopes of our future rely on

The difference between those who understand how to use AI and those who do not is going to have extremely profound downstream effects.

And I learned we are not resourced for that. There are just not enough humans doing this work, and nurses are so in demand. And care for the elderly is even worse.

That makes me think that AI can assist with care—seeing, hearing, triaging, and alerting. Depending on the situation, for example, it could be a pair of eyes watching a patient fall and alerting a person. It could be software running in the background and constantly watching for changes of lab results. It could be a conversation engine or software that answers patient questions. There are many forms of AIs that can help in the care delivery aspect of health care.

What are the ethical responsibilities of engineers and scientists like you who are directly involved in developing AI?

Li: I think there is absolutely individual responsibility in terms of how we are developing the technology. There are professional norms. There are laws. There's also the reflection of our own ethical value system. I will not be

people. I'm always more hopeful because I have hope in people. But when I get down or low, it's also because of people, not because of this technology. It's people's lack of responsibility, people's distortion of what this technology is, and also, frankly, the unfair role power and money play that is instigated or enhanced by this technology.

But then the positive side is the same. The students, the future generation, the people who are trying to do good, the doctors using AI to cure diseases, the biologists using AI to protect species, the agriculture companies using AI to innovate on farming. That's the hope I have for AI.

Are there aspects of human intelligence that you think will always be beyond the capabilities of AI?

Li: I naturally think about compassion and love. I think this is what defines us as human—possibly one of the most unique things about humans. Computers embody our values. But humans have the ability to love and feel compassion. Right now, it's not clear there is a mathematical path toward that.