Remarks about AI

Daniela Rus, CSAIL

When I tell people that I am a robotics researcher, I tend to get one of two reactions. Some people get nervous. They make jokes about Skynet, and ask when the robots will take over their jobs. They worry about autonomous vehicles because they don't trust computers to avoid crashes as well as they could themselves. The other group gets excited. This is the group that buys Roombas to vacuum their living room, and dreams of the day when their car will drive them home from work. They can't wait to hear about the latest research, and discuss all the ways that robotics and artificial intelligence will change our lives and our world for the better.

As the Head of the Computer Science and Artificial Intelligence Laboratory at MIT, I'm sure you won't be surprised to hear that I'm in the latter group. But I also know it's an important part of my job to understand the fears of that first group, to listen, to ask questions, and to give them some perspective on why I see things differently. That starts with understanding that AI is just a tool. It's an incredibly powerful one; but-—like other tools—it isn't inherently good or bad. It is what we choose to do with it. And I believe we can do some truly incredible things.

On a global scale, AI will help us generate better insights into addressing some of our biggest challenges: understanding climate change by collecting and analyzing data from vast wireless sensor networks that monitor the oceans, the greenhouse climate, and the plant condition; improving governance by data-driven decision making; eliminating hunger by monitoring, matching and re-routing supply and demand, and predicting and responding to natural disasters using cyber-physical sensors. It will help us democratize education through MOOC offerings that are adaptive to student progress, and ensure that every child gets access to the skills needed to get a good job and build a great life. It may even help those kids turn their childhood dreams into reality, as Iron Man stops being a comic book character and becomes a technological possibility.

On an individual level, AI will offer opportunities to make our lives safer, more convenient, and more satisfying. That means automated cars that can drive us to and from work, or prevent life-threatening accidents when our teenagers are at the wheel. It means customized healthcare, built using knowledge gleaned from enormous amounts of data. And counter to common knowledge, it means more satisfying jobs, not less, as the productivity gains from AI and robotics free us up from monotonous tasks and let us focus on the creative, social, and high-end tasks that computers are incapable of.

All these things—and so much more!--become possible when we direct the power of computing to fix the things that humans haven't been able to fix on our own. We're already seeing some of that work happening today. Advances are happening in three different but overlapping fields: robotics, machine learning, and artificial intelligence. Robotics puts computing into motion and gives machines autonomy. AI adds intelligence, giving machines the ability to reason. Machine learning cuts across both robotics and AI, and enables machines to learn, improve, and make predictions. Progress is being made quickly in each of these fields, so let me pause for a moment and tell you what's already happening today.

Robots have already become our partners in both industrial and domestic settings. They work side-by-side with people in assembly plants, building cars and many other goods. They help surgeons perform difficult procedures, improving outcomes and reducing scars. They mow our lawns, vacuum our floors, and even milk our cows. Altogether, nearly five and a half million robots were sold in 2015, and that number is expected to rise to more than 42 million over the next four years as we find new and better ways to use these machines.

Researchers are already making strides in improving what these robots can do. Most car manufacturers have announced self-driving car projects that will employ sensors to give vehicles a much better sense of road conditions than we can get with the naked eye. This technology will significantly reduce road fatalities through safe driving before the end of the decade. Other industry sectors will also benefit from this technology. For example, we're looking at new medical applications. My team at CSAIL is testing an ingestible robot that enables incision-free surgery. The patient swallows the robot it in the form of an ice pill that dissolves into the body, and can then be controlled by surgeons as they do useful work. As these robots improve, I believe we'll be able to offer new surgical alternatives that are less invasive, less painful, and have much lower risk of infection. Looking further down the road, this technology could intersect with advances in genetics and sequencing, and even lead to the ability to swallow a pill containing the right cells and biological agents to self-heal, or grow a new internal organ. Medical advances are only one potential area of application. We're also building robots with the strength and agility to replace humans in emergency situations and dangerous terrains, possibly even in space.

In addition to the physical capabilities being demonstrated by robots, we're also seeing rapid advances in how much information machines can process. Machine learning refers to a process that starts with a body of data and then tries to derive a rule or procedure that explains the data, or predicts future data.

These algorithms have potential applications in any field that has to process a lot of data. Again, medicine is a great example. Researchers have developed a predictive model that looks at thousands of data points and models different subtypes of lymphoma to diagnose cancer. Machines today can look at more radiology scans than a radiologist will see in their entire lives. Researchers have also developed a digital pen that measures movement 80 times a second to detect and diagnose dementia and Parkinson's disease more accurately than what we have today. They're even investigating a method of manipulating metabolic pathways that may hold the secret to reversing human obesity.

Finally, let me mention a few advances in artificial intelligence. AI is already being used in fields ranging from transportation to the environment, criminal justice to economic inclusion. AI-enabled traffic management applications have reduced wait times, energy use, and emissions by as much as 25% in certain locations. Autonomous sailboats and watercraft carry sensors around the oceans, collecting data on changes in Arctic ice and marine ecosystems. Universities are applying AI and data science to develop programs to address poverty and homelessness.

Researchers are continuing to push the boundaries of what this technology can do. For example, a new imaging technique lets you reach in and "touch" objects in YouTube. This could be used by filmmakers for entertainment purposes, but also by architects to determine if buildings are structurally sound. For those worried about cyber-security, we've also seen major progress with an AI-based system able to comb through data, detect suspicious activity, and predict up to 85% of

cyber-attacks. Researchers are even building algorithms that mitigate fake news problems by flagging articles for review.

These advances represent just a few of the ways that computers can help us tackle both everyday tasks and major global challenges. Traffic safety, medical diagnoses and treatment, environmental protection, the list goes on and on. When I read about advances like those I just described, I feel very optimistic about what technology means for the future of humanity.

But I'm also well aware that many of you may look at that list and see big, flashing "Danger" warnings. So let me explain why AI isn't the beginning of SkyNet, the Singularity, or the end of the world. The AI programs that you've probably heard of, like Google's AlphaGo, are great examples of what we call "Narrow AI." These systems are capable of looking at vast amounts of data, extracting patterns, making predictions, and acting based on those predictions. With a game like Go, the computer can study every game ever recorded and model every likely outcome. But if you expanded the board size, the system wouldn't know what to do. It also wouldn't know how to play chess, or poker, or even Go Fish. That's really the difference between "Narrow AI" and "General AI." General AI refers to a system that demonstrates intelligent behavior as advanced--or more advanced--as a human across a broad range of cognitive tasks. General AI is what everyone is worried about. But the truth is that we have no real idea how to go from Narrow AI to General AI, which is why the general consensus of experts is that we're decades away from even having to worry about it.

Even with the major advances being made today, artificial intelligence is nowhere close to people in breadth and depth of human perception, reasoning, communication, creativity. It can't undertake creative intelligence tasks or put together unfamiliar combinations of ideas. It lacks social intelligence. These restrictions will hold AI back from outthinking humanity—and from displacing people in many of the jobs they hold today. In addition, robots will struggle with physical tasks that require unpredictable movement or fine motor skills. They may ultimately replace repetitive tasks such as welding and packaging, but have limited utility for construction or forestry.

While I don't believe machines will take over the world, I do understand the very real concern about the impact of machines on our workforce. We have already seen robots and computer programs displace workers in a variety of professions. These changes aren't limited to one industry, or even low-income workers. Research by McKinsey has shown that even CEOs will have the nature of their jobs change as technology improves. But those jobs won't go away; they'll change, as automation replaces some tasks and creates time for others. Instead of asking which jobs will exist in five, ten, or twenty years, we should ask which activities will be automated, and what kinds of high-value activities we'll be able to do in their place. Let me give you a few examples.

First, consider teachers. In the near future, computers will be able to remove the burden of grading and administrative tasks. But that won't replace teachers; rather, it will free up time so they can focus on student interaction, customize materials, and make sure that they're challenging high-performing students and assisting those who otherwise might have fallen behind.

Or take lawyers. While new software has replaced some tasks related to document review, drafting, and research, computers can't counsel clients, write a compelling brief, or persuade a judge. They can't be lawyers--but they can change the type of work that lawyers are doing.

That's just as true for a lot of professions related to office or administrative work. Imagine the time that you'd get back if AI programs could extract all the information needed from email rather than forcing you to sift through it. The time saved could be reallocated to the types of creative, high-value tasks that computers are really bad at.

When you look at the positives of what computing can do, you start to see the potential that technology offers across industries: robots that will let us build vertical farms, and data to better target our use of fertilizers and pesticides; online marketplaces where you will find personalized products, try them on in a virtual fitting room, and then connect with a tailoring service that can customize them at an affordable price; democratized access to marketplaces for buyers and sellers of any kind of service; administrative programs that can help you figure out everything you need to know for your next meeting and help you get there on time.

Technology will replace all the things that we don't need humans to do, and free us up to do all the stuff that we always wished we had time to dream up and make happen.

I wasn't joking earlier when I mentioned an Iron Man suit.

In the movies, Iron Man is the hero of the stories. But to me, the more interesting story is the engineering that it will take to bring him to life. Those are the kinds of jobs that we need to get our kids ready for. Instead of worrying about jobs, let's make coding a universal feature of our education system, alongside literacy and math, and teach kids that programming gives them the power to create anything that they can imagine. Let's institute training programs that enable people displaced by automated programs into the new jobs created by these advances.

Changes won't be zero-sum. The automating of certain tasks doesn't need to mean that there is less to do. In fact, it may mean the opposite. Not that long ago, our economy was primarily agricultural. People worked on farms. With improvements to technology, people worried about what would happen to the farm laborers whose jobs no longer existed. What did they do? They got new jobs, created as a result of the new things we could do because of those same advances in technology.

Our workforce changed, it didn't disappear. It was a messy, sometimes painful process, but it ultimately led to significant growth and many of the advances that we rely on today. I foresee a similar change in our near future, and the good news is that we can take steps to reduce the pain that this change might cause.

We're already seeing positive results from programs that have taken this path. For example, a web and software design company based out of Kentucky made a conscious choice to hire former coal miners and train them to write code. Now, this is only one small program—but imagine if we put in place policies that supported and incentivized this kind of retraining in a much broader way, and created a safety net for everyone dealing with this kind of transition.

We can't stop technology from improving or changing what our world looks like. But we can proactively deal with the consequence to ensure that change is for the better. Computers and people don't need to be competitors, especially when we can do so much more when we work together.

We've already seen what this can look like in medicine. For example, a new AI-based approach was tasked with reviewing images of lymph node cells to diagnose cancer. On its own, it had an error rate of 7.5%, worse than the 3.5% rate of human pathologists. But when both the AI system and the pathologist reviewed the data, the error rate went down to only 0.5%. Today, these systems may be

deployed in the world's most advanced cancer treatment centers. But imagine a future where every practitioner—even those working in small practices, in rural settings—had access to these systems. An overworked doctor may not have the time to review every new study and clinical trial. But working in tandem with these systems, they can offer their patients the most cutting-edge diagnosis and treatment options.

Or think of manufacturing. We assume that robots will lead to fewer jobs. But what if they really bring better jobs? We're seeing rising demand for customization and personalization in almost everything we buy. Today, we don't have the ability to keep up with this demand at a price point most people can afford. But by teaming robots and people together, we can re-shape the world of manufacturing to feature smaller factories, located closer to population centers, making fewer generic standardized goods and more customized products. Even better, people won't have to understand how the robots work in order to use them. The job market will be democratized, as will the ability to purchase the products they produce.

In one important sense, machines are like people: they have the potential to both help us and hurt us depending on the choices that we make. But in closing, I want to emphasize that it is up to us to decide what impact this technology will have.

So let's take a pragmatic look at all our fears about artificial intelligence, and establish some concrete benchmarks that tell us when we actually need to worry. The worst-case scenario isn't something that can happen overnight; if we know what warning signs to look for, we have the ability to hit the brakes before any trouble occurs.

Automated cars and robotic assembly lines shouldn't scare us; they should get us excited about what problems computers can help us solve. At CSAIL, we're excited about how we can work together—with other people, and with machines—to make all of our lives better in the little ways and the big ones.