What AI Can Really Do Right Now

Voice Recognition

The Acoustic Model: Sounds Into Data

Your voice is measured by frequency, the wavelengths of sound at a specific moment. When you speak to Alexa, the software breaks down your command into 25-millisecond slivers, then converts each wavelength measurement into digestible numbers. The software compares those sonic signatures to its catalog of sounds until its confidence scores are high enough that it can assume that you said, "Order more dog food."

The Language Model: Getting Meaning

Watch your phone screen while using Siri or Google Assistant, and you'll see the transcription swap out words as you speak. That's the software comparing the words it thinks you've said to its stores of example sentences, which inform how it understands syntax and vocabulary. "The language model is trained on billions and billions of words of text," says Rohit Prasad, the head scientist behind Amazon's Alexa division. "The web, catalogs, our assets, all of that goes into the language model that judges how likely one word is to follow in a sequence." His example: "Play music by Sting" is far more likely than "Play Sting music by." Also, the software considers context to be more accurate. Ask "Who is in the cast of *Dark Knight*?" and it will read the result for the massively popular *The Dark Knight* Batman movie, rather than the lesser known 2017 art film *Dark Night*.

Self-Improvement

Acoustic and language models constantly adjust to how people use them. That's where A.I., specifically machine learning, comes in. "When you barge in on Alexa," Prasad says, referring to when users have to rephrase or specify a command or ask Alexa to stop, "we know we must've done something wrong. That's a cue to learn." For example, if you say, "Play The Martian," the device will consider whether you want

the audiobook or the Matt Damon movie. If the device needs to ask which one to play, and the number of users requesting the movie far outweighs the audiobook, the device might later default to playing the movie.

What's Coming Next

"Somebody who worked for one of these big companies told me that one of the most fascinating things you can do with all this data is to find out what people are asking about that the system doesn't do," says professor Alan Black of Carnegie Mellon University's Language Technologies Institute. "When the thing says, 'I can't do that,' you find the words recognized, and discover that it's something really interesting." Responding to unpredictable requests is part of the goal for student researchers competing for the $2.5 million Alexa Prize. The challenge: produce a chat bot that will converse with a human—ask intelligent follow-up questions, add new information—for 20 minutes. That's not difficult for a customer-service chat bot that's only concerned with khakis, but in this situation, as with conversations between humans, there are digressions and spontaneity. A program capable of that will mean a huge leap for A.I.

Advertisement - Continue Reading Below

Where A.I Has Already Infiltrated Everyday Life

No More Spam: Thanks to machine learning and years of email data, Google says that Gmail is now accurate to 99.9 percent when identifying (and quarantining) spam or phishing emails. The next application for email is context-specific automatic replies. Gmail analyzes the message you've received ("Meet on Tuesday or Wednesday?") and automatically generates responses to choose from ("Let's do Tuesday").

Faster Netflix: Ever stream a movie over bad Wi-Fi? The video gets pixelated as the app tries to use less data and keep playing. To prevent this, Netflix built software that identifies how visual a scene is, then decides how to divvy up bandwidth. So when you're watching a climactic Marvel movie battle, Netflix will use all available signal strength for the scene. Family Guy reruns? It won't strain your LTE to show Stewie in 4K.

Longer Battery Life: Your iPhone will analyze app use and motion-sensor data to predict what you're doing throughout the day. Say it needs to pull lots of data for a software update. The phone will use that data to figure out when you're likely to be driving to work. So rather than tax your phone's battery to search for cell towers to pull all that data, iOS waits until you're likely asleep and the phone is plugged in and near decent Wi-Fi.

Facial Recognition

How It Was Developed

Dr. Joseph Atick, a physicist and mathematician, helped create the first facial-recognition systems in the early 1990s. "Back then, we used statistical analysis to classify facial features as a set of patterns," he says. "It's all pattern recognition. You detect faces by detecting the coincidence of a set of local features—eyes, nose, the edges of your mouth. The chances of these things coming together are very low." Twenty-five years later, computers can do this on their own through deep learning, a multilayered system of pattern-recognition software. With the billions of pictures we upload every day to Google, Snapchat, and Facebook, those computers have an endless supply of data.

How It Works

After analyzing a still or video image, the computer converts it into a digital file, called a face print. This is a searchable measure of the size of facial features and the relationship between them. The software collects as many as 40 data points, but they aren't always used: "The more distinguishing the face—a Jay Leno chin, a monstrous nose," Atick says, "the fewer features the computer needs to identify it."

If you grow a beard or get a haircut, the software adapts. The computer knows hair can change, so it blocks it out. What's truly important to identification is the center of your face: a triangle between your temples, down to a point just below your mouth. That is the essence of an ID.

Advertisement - Continue Reading Below

Atick compares these digital representations to a city map. "If I start telling you about landmarks—Grand Central Terminal, Union Square, the Statue of Liberty—and where they are relative to each other, you'd realize I was talking about New York, not

London." There are other statues in the world, but they aren't the same size as the Statue of Liberty, or they're farther from a church. We all have noses and eyebrows, but they're not all identical, nor is their location.

Is Alexa Eavesdropping?

We've asked Google, Amazon, and Apple whether their devices are listening before you turn them on with a wake word. The answer is always no. That said, we've also asked an Air Force general if he has an Echo. Also, no. Conclusion? An internet-connected microphone is always a risk for certain professions. For us civilians, using an Echo isn't much different from using a spam email filter. Software, not humans, is what's primarily analyzing your speech.

Where to Find A.I. in Your Life

Facebook: Tag your friend David?

The DMV: New license pictures automatically run through the DMV database to check for duplicates and aliases.

Phones: No need for a fingerprint when you look into the camera to unlock your Samsung S8.

Airbnb: You can confirm your identity in the app by taking a selfie, which is checked against the driver's license you scanned when you created your account.

Driving: In late 2016, New York governor Andrew Cuomo proposed installing cameras that will capture faces at bridges and tunnels in New York City and in Pennsylvania Station, the city's commuter train hub.

Minneapolis: Delta is currently testing a system that scans your face and compares it to your passport photo to automate the bag-check process.

Germany: A billboard playing a beer ad only runs when it senses a woman walking by.

Whatever Happened to Watson?

Getty

Since the mid-2000s, Watson has been the name for IBM's artificial intelligence system, most famous for beating Jeopardy champ Ken Jennings in 2011. In the years

since, IBM has built up partnerships with medical organizations like Memorial Sloan-Kettering and medical-test giant Quest Diagnostic to develop software capable of quickly and accurately diagnosing patients. But as one doctor we spoke to who worked on Watson put it, "It's way further behind than I thought it would be." Other people we interviewed had similar suspicions of Watson's relevance in the face of progress made at places like Google and Facebook. To its credit, IBM has access to loads of medical data that could turn into useful systems in medicine. But we're still waiting on convincing evidence of Watson's usefulness.

Advertisement - Continue Reading Below

The Silicon Powering It All

To understand why Nvidia made so much money in the past few years, we need one bit of jargon: GPU. The graphics processing unit is a bit of silicon and transistors found in every laptop and smartphone. Unlike the central processing unit (CPU), which balances resources to run browser tabs and email apps simultaneously, the GPU is focused, made for splitting up and processing masses of similar data quickly. It's a deli slicer to the CPU's Leatherman.

For years, GPUs like the kind Nvidia made were used for rendering pretty graphics, as in Call of Duty or an Audi infotainment system. But in the mid-2000s, researchers realized that the GPU's design was useful for processes like deep learning. Later, advances in training voice- and photo-recognition software showed even more of a need for hardware that could quickly process a ton of unvarying information, the same stuff used to make 3D graphics. So Nvidia started building GPUs specifically for A.I.

"It's not like they got blindsided by it, but it just so happened that computing moved in their direction," says Alex Kolicich, a partner at 8VC, a venture-capital firm where he specializes in A.I. investments. "If you look over the last year or year and a half, Nvidia's stock has exploded for this exact reason." Now, Google and Apple are building their own GPUs, but Nvidia is far ahead. Pick almost anywhere that's doing important work with A.I.—Facebook, Microsoft, Amazon, Tesla. Nvidia is the hardware running it.

The All-Knowing A.I.

Advertisement - Continue Reading Below

The protest is escalating, so a witness starts a Facebook Live video. It shows a shouting mass standing in the middle of a major street. Banjo's software sees it, then checks local news sites to see if there are any parades scheduled for today. No. This is unusual. One of the protesters tweets a photo, geotagged to the same neighborhood as the Facebook Live video. Banjo analyzes the photo and recognizes that the glossy rectangle is a phalanx of police with riot shields. Then, someone sets a car on fire. No one has tweeted or Instagrammed that, but since Banjo suspects that something might be happening here, it checks the feed from nearby traffic cams. The orange and red slivers are real flames, Banjo's image-recognition system concludes, much bigger and more lethal than you'd see in a video of a stovetop or a beach bonfire. The software starts listening to police radio for codes—and so on as Banjo pulls more and more information until it can confidently send a push notification to its subscriber: We have a riot.

"We call it multi-source detection," says CEO Damien Patton. "Using different data sources to corroborate that something is real and happening right now." Banjo's humble job: Take the same type of image and voice recognition you'd find in Google Photos or an Amazon Echo, and apply it to the incomprehensible volumes of text, audio, photos, and video flowing through Facebook, Instagram, Twitter, traffic cameras, 911 call centers, police radios, and weather reports.

A few years ago, Banjo software informed a local news station that an Amazon data center had caught fire. The broadcast went out, and Amazon stock fell slightly.

But Banjo knew it first, which is why this particular use of A.I. is of interest for any industry in which being first means money. "Our biggest focus is the media industry and broadcasters," Patton says. The company also just opened an office in Park City, Utah, dedicated to investments in futures.

Sounds ominous, but Banjo says it strips users' personal information from social-media data. And Banjo does not work with any government or government entities. Why? "We don't want this technology to be used for Minority Report–type surveillance and monitoring," he says. "We're an A.I. company. And A.I. can be used for good, or evil."

How to Be a Part of the A.I. Future

Learn a Language

Building A.I. systems means learning to code. "Python is the place to start," says Alex Teichman, CEO of A.I. startup Lighthouse—he also participated in the 2007 DARPA Challenge while getting his Ph.D. from Stanford. "It's super easy to get into and it's concise." Download the Python app from Python.org. If you've already got Python down, it's on to C++, which gives better control over large amounts of information, essential when working with large data sets. Ideally, you'll learn the basics of both.

Learn a Language

Sebastian Thrun's A.I. research is a big part of why the term "self-driving car" exists. But he's also the cofounder of an online education company called Udacity that has (often free) video lessons on machine learning or the aforementioned Python.

Get the Software

TensorFlow is free machine learning software made by Google. With a few lines of code and some example data (handwriting samples, cat pictures), TensorFlow will do all the hard work. You'll end up with an app that can recognize handwriting or identify cats in photos. "Everybody uses TensorFlow," says Alex Kolicich, the venture capitalist specializing in A.I.—he also helped build Google Street View. "If you can code," Kolicich says, "you can learn it in a week or two."

A World of Caution, from Elon Musk

Getty

"If I were to guess at what our biggest existential threat is," the man said, referring to artificial intelligence, "it's probably that." He continued: "We're summoning the demon." It's a quote you'd expect from a TV news guest or a science-fiction author. But that was Elon Musk, the guy running Tesla while testing rockets and planning Mars colonization. He has A.I. hysteria, too?So, in typical Musk fashion, he did something about it. He gave $10 million to the Future of Life Institute, where board members like Stephen Hawking and Morgan Freeman focus on ensuring that new artificial-intelligence systems are designed to benefit humanity, not narrow interests. Then, Musk, along with eight other backers, put $1 billion into OpenAI. Its mission: to design A.I. that prioritizes humans, and ensure that everyone has access to it. "It's not self-evident that you want to share the recipe for building these very powerful

systems," says Ilya Sutskever, a cofounder of OpenAI and head of research. At 31 years old, he's also among the youngest of the A.I. scientists responsible for the deep-learning breakthroughs of the early 2010s. And, as with Musk, his concerns would sound alarmist coming from anybody else. "We are focused on building A.I. that will be not only capable, but act in our best interests." Today, that means building software for tasks like identifying false news stories. But the far future is on his mind. "There are disagreements among experts about the speed of progress," he says. "But we are talking about systems that are going to be dramatically smarter than humans in every respect—better artists, better scientists. And when you have systems that are clever, creative, and knowledgeable, they will have extreme power."

This story appears in the November 2017 issue.