Stanford Team Aims at Alexa and Siri With a Privacy-Minded Alternative

By John Markoff
9-11 minutes

Credit: Dalbert B. Vilarino
Palo Alto, Calif. — It has been almost two decades since Google started to dominate internet search the way Microsoft dominated software for personal computers a generation earlier.

Now computer scientists at Stanford University are warning about the consequences of a race to control what they believe will be the next key consumer technology market — virtual assistants like Amazon’s Alexa and Google Assistant.

The group at Stanford, led by Monica Lam, a computer systems designer, last month received a $3 million grant from the National Science Foundation. The grant is for an internet service they hope will serve as a Switzerland of sorts for systems that use human language to control computers, smartphones and internet devices in homes and offices.

The researchers’ biggest concern is that virtual assistants, as they are designed today, could have a far greater impact on consumer information than today’s websites and apps. Putting that information in the hands of one big company or a tiny clique, they say, could erase what is left of online privacy.

“A monopoly assistant platform has access to data in all our different accounts. They will have more knowledge than Amazon, Facebook and Google combined,” Dr. Lam said in an interview.

From left, Dan Boneh, Monica Lam, Jen King, Chris Manning,
James Landay and David Mazieres, seated, are computer science researchers at Stanford University. Credit Brian Flaherty for The New York Times.

Image
Virtual assistants have access to a broader and more personal range of data than, say, a search engine. A virtual assistant can be like a personal secretary, with access to many of the most intimate details of your life.

Dr. Lam is collaborating with a group of Stanford faculty members and students to build a virtual assistant that would allow individuals and corporations to avoid surrendering personal information as well as retain a degree of independence from giant technology companies.

The system from Dr. Lam’s group is called Almond. In a recent paper, they argued for an approach in which virtual assistant software is decentralized and connected by programming standards that will make it possible for consumers to choose where their information is stored and how it is shared.

A first version of the service was released last year, and the Stanford researchers are now trying to build an alliance with larger technology and consumer companies.

The market for virtual assistants, which do tasks as varied as selecting music or turning on and off thermostats and lights, is booming. Earlier this year Google said that, including Android-based phones, it was nearing installing its Assistant service on one billion devices. Amazon said it had sold more than 100 million Echo and related devices.

Virtual assistants have not attracted significant scrutiny from
government regulators because the market is still small. But a handful of companies — Amazon, Google, Apple and Microsoft — are already dominating it.

The Google Assistant and Alexa from Amazon have captured a huge chunk of the digital assistant market. Credit Joe Buglewicz for The New York Times

Image

The Google Assistant and Alexa from Amazon have captured a huge chunk of the digital assistant market. Credit Joe Buglewicz for The New York Times

“This is the same situation as with operating systems and browsers,” said Tim Wu, a Columbia University law professor who has written extensively on the concentration of economic power in the internet era. “It’s competitively dangerous.”

The Stanford researchers are hoping to gain support by making their software freely available to users of smartphones, computers and consumer appliances.

They are encouraging makers of consumer products to connect their devices to the Almond virtual assistant through a Wikipedia-style service they call Thingpedia. It is a shared database in which any manufacturer or internet service could specify how its product or service would interact with the Almond virtual assistant.

They also hope Almond can leapfrog existing virtual assistant systems in its ability to understand complex language. Virtual assistants are doing a better job of understanding what humans say, but they have made much less progress in understanding what those words mean. Context and nuance are difficult for a machine
to understand.

While simple phrases like “What is the temperature?” or “Play a Beatles song” are now routinely handled by computer assistants, routine human interactions that require understanding of context or rely on something that was previously spoken are much more challenging.

The services now do best in specific domains, like all the questions you might get about controlling your Spotify account.

The Stanford researchers are hoping to gain support by making Almond freely available to users of smartphones, computers and consumer appliances.

It has taken years of work to get to this point. Three decades ago, Apple commissioned a group of employees that included Bud Colligan, Hugh Dubberly, Mike Liebhold and Doris Mitsch to create a video showing how in the future humans might interact with computers using spoken language.

The video, known as “Knowledge Navigator,” featured an absent-minded professor who talked with a computing system to perform everyday tasks and academic research.

The demonstration inspired a number of developers, including the artificial intelligence researchers Adam Cheyer and Tom Gruber, who began research on virtual assistants while they were still at SRI International, an independent research laboratory in Menlo Park,
In 2010, Apple acquired the start-up and then released its technology for the iPhone the following year.

Since then, Siri has faced stiff competition. Last year, Amazon said it had 10,000 employees working on its Alexa service, many of them focusing on improving the ability to understand complex commands.

The Stanford researchers argue that Alexa’s approach, even with thousands of employees, will never be able to adequately deal with the complexity and variability of human language because it is incredibly labor-intensive and may not extend to more complex conversations.

An Amazon Echo Dot. Virtual assistants are doing a better job of understanding what humans say, but they have made much less progress in understanding what those words mean. Credit Monica Jorge for The New York Times

Amazon researchers, on the other hand, have said that having access to vast amounts of data will give them a meaningful advantage in developing more sophisticated conversational software.

The Stanford researchers have developed a system named Genie that simplifies the task of training a so-called neural network. They
are improving the accuracy of their service by creating test data, some of it generated by humans and the rest by sentences created by special test programs.

While machine accuracy in understanding spoken words is now routinely above 90 percent, accuracy in understanding complex natural language is substantially lower. A recent paper by the Stanford researchers, which described a significant advance in language understanding, still reached only 62 percent accuracy on “realistic user inputs,” actual statements produced by human test subjects in written form.

[Get the Bits newsletter for the latest from Silicon Valley and the technology industry.]

Dr. Gruber, who recently left Apple after heading advanced development there, remains skeptical of any near-term technical breakthrough that will make it possible for virtual assistants to have humanlike understanding.

“When you get a question and you don’t know what domain it’s in, then you have this very complicated problem of massive ambiguity,” he said.

Dr. Lam said the threat to privacy cannot be overstated. For example, she noted that Wynn Resorts in Las Vegas last year installed Amazon Echo devices in rooms.

“Once they said that what happens in Las Vegas stays there,” she said. “Now that’s no longer necessarily true. Now it might end up in Seattle.”

Correction: July 11, 2019

An earlier version of this article misstated who was involved in the
creation of the “Knowledge Navigator” video. It was a group of Apple employees that included Bud Colligan, Hugh Dubberly, Mike Liebhold and Doris Mitsch. The computer scientist Alan Kay did not have a role in the video’s development.

Follow John Markoff on Twitter: @markoff.

Interested in All Things Tech? Get the Bits newsletter for the latest from Silicon Valley. And sign up for the personal tech newsletter for advice and tips on the technology changing how you live.

A version of this article appears in print on June 14, 2019, Section B, Page 1 of the New York edition with the headline: Digital Help That Puts User Privacy First? What?. Order Reprints | Today’s Paper | Subscribe