

Personal assistants are ushering in the age of AI at home.

How the use of the voice is making machines smarter and humans more dependent.

Mona Lalwani, October 5, 2016

Google Home is the latest embodiment of a virtual assistant. The voice-activated speaker can help you make a dinner reservation, remind you to catch your flight, fire up your favorite playlist and even translate words for you on the fly. While the voice interface is expected to make quotidian tasks easier, it also gives the company unprecedented access to human patterns and preferences that are crucial to the next phase of artificial intelligence.

Comparing an AI agent to a personal assistant, as most companies have been doing of late, makes for a powerful metaphor. It is one that is indicative of the human capabilities that most major technology companies want their disembodied helpers to adopt. Over the last couple of years, with improvements in speech-recognition technology, Siri, Cortana and Google Now have slowly learned to move beyond the basics of weather updates to take on more complex responsibilities like managing your calendar or answering your queries. But products that invade our personal spaces -- like Amazon's Echo and Google Home -- point to a larger shift in human-device interaction that is currently underway.

Onstage demos of Google Home, which has the company's Assistant built into it, suggest a conversational capability that requires an

advanced understanding of human intent and context. The device relies almost entirely on the company's speech-recognition technology that has been in the making for almost a decade, since the early days of GOOG 411. But over the years, the basic telephone-based directory search has grown into the much more complex Google Now.

The drastic jump in the Android assistant's capabilities has come from neural net training and deep learning techniques that have allowed scientists to boost speech-recognition technology to a point where it is now starting to learn the nuances of human behavior through the medium of voice.

Using the voice to communicate with an outside entity makes for an intimate and innately human experience. "Speech is the most dominant way that humanity has been communicating with each other," David Nahamoo, speech CTO at IBM Research, said over the phone. "When we communicate with the outside, we speak. But from outside to inside, we absorb information a lot better visually. It's because of our heritage and the evolution that we have gone through. From the standpoint of efficiency, speech is quickest way to get a point across."

"Voice changes the way people interact with their systems." - Françoise Beaufays, Google

Devices like Echo and Google Home, for instance, are built on speech recognition that can help you stay heads-up and hands-free while you multitask around the house. So instead of spending time swiping and typing, you can tell the personal assistant what you need or what you're looking for. It's that kind of ease and productivity that companies dangle in front of the users to have them adopt chatbots and personal assistants in their daily communications, but talking to devices also opens the door to a new kind of relationship.

"I think voice changes the way people interact with their systems," says Françoise Beaufays, a research scientist who works on speech recognition at Google. "For a long time when people were typing in their browsers for information, they would write something cryptic like 'Eiffel Tower height,' for example." The string of seemingly random words would instantly pull up search results on google.com with pictures, details and dimensions of the iconic French structure. But when speech recognition started to take shape with smartphone assistants, Beaufays says there was a clear change in communication.

"As people started feeling comfortable with speech, instead of being cryptic they started saying: 'Hey, what is the height of the Eiffel Tower?' or 'How tall is the Eiffel Tower?'," she says. "We saw that switch in the way people were addressing their devices in speech first and typing next. Using your voice is bringing in

more discursive type of interaction, and even though you know very well it's a machine you behave a little more human with it."

While a verbal exchange with a virtual assistant can make it easier to get things done, it also makes it easier for the companies to gain invaluable insight into the human world that's filled with vocal clues to feelings and preferences. "We're going from computing to understanding," says James Barrat, author of *Our Final Invention: Artificial Intelligence and the End of the Human Era*. "It's not just us chatting. These machines are listening to what we like and don't like, how we speak and what we speak about. It's greater access to how we think."

In the world of AI, data is the currency that will set one company apart from the other. Through voice searches, millions of vocal samples become available to the companies that are fine-tuning personal assistants. The stream of information is fed back into the system to improve the accuracy of the algorithms, but it also gives the companies access to the complexities of human intent. In effect, using the voice to communicate with an AI helper only makes it smarter.

A lot can be gleaned from the vocal communication. Words and intonations start to give away user patterns, preferences and even emotions over time. That kind of insight into the mindset of the user is critical to the next wave of personalized AI that is already taking shape at companies like Google, Amazon and Facebook.

Smart, talking AIs at home will fire up the ecosystem of the Internet of Things, taking it from novelty machines to necessities. With companies aspiring to make their assistants omnipresent and their machines more interconnectable, they need capable speech recognition to get the job done.

"There's a parallel thrust," says Vlad Sejnoha, CTO at Nuance Communications, one of the leaders in voice-recognition technologies.

"You'll interact with your smart fridge or printer in a more natural way but also see a portable personal assistant that lives in a cloud and follows you around to help you navigate a complex world." Google Home, much like Amazon's Echo, already comes with partnerships that are useful around the house. You can use the speaker to control your Chromecast, Nest and Philips Hue lights.

In addition to navigating the immediate physical world, an omnipresent assistant could potentially become a gateway to unfamiliar settings or foreign languages too. In the spot aired during the Google event this week, the company demonstrated that Home has the ability to tap Google Translate to respond with accurate translations from English to Spanish. But whether the machine can comprehend foreign accents and translate the reverse, remains to be seen.

"These machines are listening to what we like and don't like, how we speak and what we speak about. It's greater access to how we think." - James Barrat

Failing to comprehend different accents has been one of the biggest downfalls of most digital assistants on smartphones today. Scientists building these systems often talk about the lack of data as one of the biggest obstacles to understanding new accents and languages. The copious amounts of information required to make that possible calls for massive investments from the companies. Taking the technology straight to people's homes opens up a steady stream of data that can be used for tests back in the research labs.

A lot of the building blocks are starting to fall into place for devices like Google Home to become efficient personal assistants. And even though, there's a need to be more vigilant of the ways human-device interactions are starting to shift; most voice interface developers believe it's a necessary change that will extend human capabilities.

"Having an AI that is your agent and helps you exist in the world better, gets you better information and services is hugely exciting," says Sejnoha. "As with anything there are uses that can be negative, we're all familiar with privacy and mining data. That's something we have to be thoughtful about, but the benefits far outweigh those scenarios."