

tonyinosaka
2 Followers

lhnatko recommended: [John Updike's Archive](#)

Recommend

- HOME PAGE
- TODAY'S PAPER
- VIDEO
- MOST POPULAR
- TIMES TOPICS

Try Times Reader 2.0 | Welcome, | [Log Out](#) | [Help](#)

The New York Times | International Herald Tribune

GLOBAL EDITION
Science

Search All NYTimes.com

- WORLD
- U.S.
- N.Y. / REGION
- BUSINESS
- TECHNOLOGY
- SCIENCE
- HEALTH
- SPORTS
- OPINION
- ARTS
- STYLE
- TRAVEL
- JOBS
- REAL ESTATE
- AUTOS

ENVIRONMENT SPACE & COSMOS



SMARTER THAN YOU THINK

Computers Learn to Listen, and Some Talk Back



Stuart Isett for The New York Times

Eric Horvitz's receptionist at Microsoft is a computer that can interact with visitors outside his office in Redmond, Wash.

By **STEVE LOHR** and **JOHN MARKOFF**
Published: June 24, 2010

“Hi, thanks for coming,” the medical assistant says, greeting a mother with her 5-year-old son. “Are you here for your child or yourself?”

Smarter Than You Think

Communicating With Computers

Articles in this series are examining the recent advances in artificial intelligence and robotics and their potential impact on society.

[Go to Series Page »](#)

Multimedia



Interactive Feature

[Building Smarter Machines](#)

The boy, the mother replies. He has diarrhea.

“Oh no, sorry to hear that,” she says, looking down at the boy.

The assistant asks the mother about other symptoms, including fever (“slight”) and abdominal pain (“He hasn’t been complaining”).

She turns again to the boy. “Has your tummy been hurting?” Yes, he replies.

After a few more questions, the assistant declares herself “not that concerned at this point.” She schedules an appointment with a doctor in a couple of days. The mother leads her son from the room, holding his hand.

- FACEBOOK
- TWITTER
- RECOMMEND
- COMMENTS (114)
- E-MAIL
- SEND TO PHONE
- PRINT
- REPRINTS
- SHARE



Travel Dispatch E-Mail



Sign up for the latest travel features, sent every Saturday. See [Sample](#)
tonysilva@mac.com [Sign Up](#)
[Change E-mail Address](#) | [Privacy Policy](#)

Ads by Google

[what's this?](#)

[Nokia Siemens Networks](#)

Long Term Evolution Improves Mobile Broadband Speed. See Examples Here!
www.NokiaSiemensNetworks.com

[Tokyo Apartments](#)

Short to Long Stay apartments. Now Up to 30% OFF!
www.tokyoapartments.jp

[Free Text To Speech](#)

Latest Technology. Let your PC Read Aloud with Natural Voices Now!
www.naturalreaders.com

[Text to Speech](#)

Amazing lifelike voices! Hear it. Try it free. [NextUp.com](#)

[Induction Lighting SOLARA](#)

The future of lighting is extremely long lasting and eco-friendly
www.amko.com.tw/SOLARA

[Contact Center & IP PBX](#)

Call Center, Voice Logger, IVR, ACD Integrated Solution
Contact Center
www.intelx.co.id

[Advertise on NYTimes.com](#)

MOST POPULAR

- E-MAILED
- BLOGGED
- SEARCHED
- VIEWED

- David Brooks: The Culture of Exposure
- Success and Scrutiny at Hebrew Charter School
- Smarter Than You Think: Computers Learn to Listen, and Some Talk Back
- What Broke My Father's Heart
- On New iPhone, a Mystery of Dropped Calls
- On Location: In the Catskills, Comfort in a Gingerbread House
- Roger Cohen: Feeling Bleu
- Choice Tables: In Oakland, Innovation Is on the Menu
- The Vanishing Mind: Promise Seen for Detection of Alzheimer's
- Paul Krugman: The Renminbi Runaround

[Go to Complete List »](#)



Graphic

A Doctor's Office, Revisited



Video

Medical Bayesian Kiosk

Related

[Jobs Created and Displaced](#) (June 25, 2010)

[Speech Recognition's Early Days](#) (June 25, 2010)

[Smarter Than You Think: What Is I.B.M.'s Watson?](#) (June 20, 2010)

RSS Feed

[Get Science News From The New York Times »](#)

[Enlarge This Image](#)



Jodi Hilton for The New York Times

Dr. Michael Lee with Chantelia Garrity and her daughter this month in Norwood, Mass. Dr. Lee uses speech software to transcribe his recorded notes.

Readers' Comments

Readers shared their thoughts on this article.

[Read All Comments \(114\) »](#)

But he keeps looking back at the assistant, fascinated, as if reluctant to leave.

Maybe that is because the assistant is the disembodied likeness of a woman's face on a computer screen — a no-frills avatar. Her words of sympathy are jerky, flat and mechanical. But she has the right stuff — the ability to understand speech, recognize pediatric conditions and reason according to simple rules — to make an initial diagnosis of a childhood ailment and its seriousness. And to win the trust of a little boy.

“Our young children and grandchildren will think it is completely natural to talk to machines that look at them and understand them,” said Eric Horvitz, a computer scientist at [Microsoft's](#) research laboratory who led the medical avatar project, one of several intended to show how people and computers may communicate before long.

For decades, computer scientists have been pursuing artificial intelligence — the use of computers to simulate human thinking. But in recent years, rapid progress has been made in machines that can listen, speak, see, reason and learn, in their way. The prospect, according to scientists and economists, is not only that artificial intelligence will transform the way humans and machines communicate and collaborate, but will also eliminate millions of jobs, create many others and change the nature of work and daily routines.

The artificial intelligence technology that has moved furthest into the mainstream is computer understanding of what humans are saying. People increasingly talk to their cellphones to find things, instead of typing. Both [Google's](#) and [Microsoft's](#) search services now respond to voice commands. More drivers are asking their cars to do things like find directions or play music.

The number of American doctors using speech software to record and transcribe accounts of patient visits and treatments has more than tripled in the past three years to 150,000. The progress is striking. A few years ago, supraspinatus (a rotator cuff muscle) got translated as “fish banana.” Today, the software transcribes all kinds of medical terminology letter perfect, doctors say. It has more trouble with other words and grammar, requiring wording changes in about one of every four sentences,

doctors say.

“It's unbelievably better than it was five years ago,” said Dr. Michael A. Lee, a pediatrician in Norwood, Mass., who now routinely uses transcription software. “But it struggles with ‘she’ and ‘he,’ for some reason. When I say ‘she,’ it writes ‘he.’ The technology is sexist. It likes to write ‘he.’ ”

Meanwhile, translation software being tested by the [Defense Advanced Research Projects Agency](#) is fast enough to keep up with some simple conversations. With some troops in Iraq, English is translated to Arabic and Arabic to English. But there is still a long way to go. When a soldier asked a civilian, “What are you transporting in your truck?” the Arabic reply was that the truck was “carrying tomatoes.” But the English translation became “pregnant tomatoes.” The speech software understood “carrying,” but not the

Stay connected to the news. Without being connected to the Web.

Times Reader 2.0.
The portable, versatile news reader.

[LEARN MORE ►](#)

The New York Times
Times Reader 2.0
Powered by ADOBE AIR

context.

Yet if far from perfect, speech recognition software is good enough to be useful in more ways all the time. Take call centers. Today, voice software enables many calls to be automated entirely. And more advanced systems can understand even a perplexed, rambling customer with a misbehaving product well enough to route the caller to someone trained in that product, saving time and frustration for the customer. They can detect anger in a caller's voice and respond accordingly — usually by routing the call to a manager.

So the outlook is uncertain for many of the estimated four million workers in American call centers or the nation's 100,000 medical transcriptionists, whose jobs were already threatened by outsourcing abroad. "Basic work that can be automated is in the bull's-eye of both technology and globalization, and the rise of artificial intelligence just magnifies that reality," said [Erik Brynjolfsson](#), an economist at the Sloan School of Management at the [Massachusetts Institute of Technology](#).

Still, Mr. Brynjolfsson says artificial intelligence will also spur innovation and create opportunities, both for individuals and entrepreneurial companies, just as the Internet has led to new businesses like Google and new forms of communication like blogs and social networking. Smart machines, experts predict, will someday tutor students, assist surgeons and safely drive cars.

The Digital Assistant

"Hi, are you looking for Eric?" asks the receptionist outside the office of Eric Horvitz at Microsoft.

This assistant is an avatar, a time manager for office workers. Behind the female face on the screen is an arsenal of computing technology including speech understanding, image recognition and machine learning. The digital assistant taps databases that include the boss's calendar of meetings and appointments going back years, and his work patterns. Its software monitors his phone calls by length, person spoken to, time of day and day of the week. It also tracks his location and computer use by applications used — e-mail, writing documents, browsing the Web — for how long and time of day.

When a colleague asks when Mr. Horvitz's meeting or phone call may be over, the avatar reviews that data looking for patterns — for example, how long have calls to this person typically lasted, at similar times of day and days of the week, when Mr. Horvitz was also browsing the Web while talking? "He should be free in five or six minutes," the avatar decides.

The avatar has a database of all the boss's colleagues at work and relationships, from research team members to senior management, and it can schedule meetings. Mr. Horvitz has given the avatar rules for the kinds of meetings that are more and less interruptible. A session with a research peer, requiring deep concentration, may be scored as less interruptible than a meeting with a senior executive. "It's O.K. to interrupt him," the assistant tells a visitor. "Just go in."

As part of the project, the researchers plan to program the avatar to engage in "work-related chitchat" with colleagues who are waiting.

The conversation could be about the boss's day: "Eric's been in back-to-back meetings this afternoon. But he's looking forward to seeing you." Or work done with the boss: "Yes, you were in the big quarterly review with Eric last month." Or even a local team: "How about that Mariners game last night?"

Mr. Horvitz shares a human administrative assistant with other senior scientists. The avatar's face is modeled after her. At Microsoft, workers typically handle their own calendars. So the main benefit of the personal assistant, Mr. Horvitz says, is to manage his time better and coordinate his work with colleagues'. "I think of it as an extension of

me,” he said. “The result is a broader, more effective Eric.”

Computers with artificial intelligence can be thought of as the machine equivalent of idiot savants. They can be extremely good at skills that challenge the smartest humans, playing chess like a grandmaster or answering “Jeopardy!” questions like a champion. Yet those skills are in narrow domains of knowledge. What is far harder for a computer is common-sense skills like understanding the context of language and social situations when talking — taking turns in conversation, for example.

The scheduling assistant can plumb vast data vaults in a fraction of a second to find a pattern, but a few unfamiliar words leave it baffled. Jokes, irony and sarcasm do not compute.

That brittleness can lead to mistakes. In the case of the office assistant, it might be a meeting missed or a scheduling mix-up. But the medical assistant could make more serious mistakes, like an incorrect diagnosis or a seriously ill child sent home.

The Microsoft projects are only research initiatives, but they suggest where things are headed. And as speech recognition and other artificial intelligence technologies take on more tasks, there are concerns about the social impact of the technology and too little attention paid to its limitations.

Smart machines, some warn, could be used as tools to isolate corporations, government and the affluent from the rest of society. Instead of people listening to restive customers and citizens, they say, it will be machines.

“Robot voices could be the perfect wall to protect institutions that don’t want to deal with complaints,” said Jaron Lanier, a computer scientist and author of “You Are Not a Gadget” (Knopf, 2010).

Smarter Devices

“I’m looking for a reservation for two people tomorrow night at 8 at a romantic restaurant within walking distance.”

That spoken request seems simple enough, but for a computer to respond intelligently requires a ballet of more than a dozen technologies.

A host of companies — [AT&T](#), Microsoft, Google and startups — are investing in services that hint at the concept of machines that can act on spoken commands. They go well beyond voice-enabled Internet search.

Perhaps the furthest along is Siri, a Silicon Valley company offering a “virtual personal assistant,” a collection of software programs that can listen to a request, find information and take action.

In this case, Siri, presented as an [iPhone](#) application, sends the spoken request for a romantic restaurant as an audio file to computers operated by [Nuance Communications](#), the largest speech-recognition company, which convert it to text. The text is then returned to Siri’s computers, which make educated guesses about the meaning.

“It’s a bit like the task faced by a waiter for whom English is a second language in a noisy restaurant,” said [Tom Gruber](#), an artificial intelligence researcher and co-founder of Siri. “It isn’t perfect, but in context the waiter can usually figure out what you want.”

The Siri system taps more data to decide if it is seeking a romantic restaurant or romantic comedy. It knows the location of the phone and has rules for the meaning of phrases like “within walking distance.” It scans online restaurant review services like [Yelp](#) and Gayot for “romantic.”

Siri takes the winnowed list of restaurants, contacts the online reservation service Open Table and gets matches for those with tables available at 8 the next day. Those restaurants are then displayed on the user’s phone, and the reservation can be

completed by tapping a button on the screen. The elaborate digital dance can be completed in a few seconds — when it works.

[Apple](#) is so impressed that it bought Siri in April in a private transaction estimated at more than \$200 million.

Nelson Walters, an [MTV](#) television producer in New York, is a Siri fan. It saves him time and impresses his girlfriend. “I will no longer get lost in searching Yelp for restaurant recommendations,” he said. But occasionally, Mr. Walters said, Siri stumbles. Recently, he asked Siri for the location of a sushi restaurant he knew. Siri replied with directions to an Asian escort service. “I swear that’s not what I was looking for,” he said.

Mr. Gruber said Siri had heard an unfamiliar Japanese word, but did not know the context and guessed wrong.

In cars, too, speech recognition systems have vastly improved. In just three years, the [Ford Motor Company](#), using Nuance software, has increased the number of speech commands its vehicles recognize from 100 words to 10,000 words and phrases.

Systems like Ford’s Sync are becoming popular options in new cars. They are also seen by some safety specialists as a defense, if imperfect, against the distracting array of small screens for GPS devices, smartphones and the like.

Later this summer, a new model of the Ford Edge will recognize complete addresses, including city and state spoken in a single phrase, and respond by offering turn-by-turn directions.

To the Customer’s Rescue

“Please select one of the following products from our menu,” the electronics giant [Panasonic](#) used to tell callers seeking help with products from power tools to plasma televisions.

It was not working. Callers took an average of 2 1/2 minutes merely to wade through the menu, and 40 percent hung up in frustration. “We were drowning in calls,” recalled Donald Szczepaniak, vice president of customer service. Panasonic reached out to AT&T Labs in 2005 for help.

The AT&T researchers worked with thousands of hours of recorded calls to the Panasonic center, in Chesapeake, Va., to build statistical models of words and phrases that callers used to describe products and problems, and to create a database that is constantly updated. “It’s a baby, and the more data you give it, the smarter it becomes,” said Mazin Gilbert, a speech technology expert at AT&T Labs.

The goal of the system is to identify key words — among a person’s spoken phrases and sentences — so an automated assistant can intelligently reply.

“How may I help you?” asked the automated female voice in one recording.

“I was watching ‘[American Idol](#)’ with my dog on Channel 5,” a distraught woman on the line said recently, “and suddenly my TV was stuck in Spanish.”

“What kind of TV?” the automated assistant asked, suggesting choices that include plasma, LCD and others.

“LCD,” replied the woman, and her call was sent to an agent trained in solving problems with LCD models.

Simple problems — like product registration or where to take a product for repairs — can be resolved in the automated system alone. That technology has improved, but callers have also become more comfortable speaking to the system. A surprising number sign off by saying, “Thank you.”

Some callers, especially younger ones, also make things easier for the computer by uttering a key phrase like “plasma help,” Mr. Szczepaniak said. “I call it the Google-ization of the customer,” he said.

Over all, half of the calls to Panasonic are handled in the automated system, up from 10 percent five years ago, estimated Lorraine Robbins, a manager.

But the other half of calls are more complex problems — like connecting a digital television to a cable box. In those cases, the speech recognition system quickly routes a call to an agent trained on the product, so far more problems are resolved with a single call. Today, Panasonic resolves one million more customer problems a year with 1.6 million fewer total calls than five years ago. The cost of resolving a customer issue has declined by 50 percent.

The speech technology’s automated problem sorting has enabled Panasonic to globalize its customer service, with inquiries about older and simpler products routed to its call centers in the Philippines and Jamaica. The Virginia center now focuses on high-end Panasonic products like plasma TVs and home theater equipment. And while the center’s head count at 200 is the same as five years ago, the workers are more skilled these days. Those who have stayed have often been retrained.

Antoine Andujar, a call center agent for more than five years, attended electronics courses taught at the call center by instructors from a local community college. He used to handle many products, but now specializes in issues with plasma and LCD televisions.

Mr. Andujar completed his electronics certification program last year, and continues to study. “You have to move up in skills,” he said. “At this point, you have to be certified in electronics to get in the door here as a Panasonic employee.”

The Efficient Listener

“This call may be recorded for quality assurance purposes.”

But at a growing number of consumer call centers, technical support desks and company hot lines, the listener is a computer. One that can recognize not only words but also emotions — and listen for trends in customer complaints.

In the telephone industry, for example, companies use speech recognition software to provide an early warning about changes in a competitor’s calling plans. By detecting the frequent use of names like AT&T and other carriers, the software can alert the company to a rival that lowered prices, for example, far faster than would hundreds of customer service agents. The companies then have their customer agents make counteroffers to callers thinking of canceling service.

Similar software, used by [Aetna](#), began to notice the phrase “[cash for clunkers](#)” in hundreds of calls to its call center one weekend last year. It turned out that tens of thousands of car shoppers responding to the government incentive were calling for insurance quotes. Aetna created insurance offers for those particular callers and added workers to handle the volume.

And as Apple’s new smartphone surged in popularity several years ago, GoDaddy, an Internet services company, learned from its call-monitoring software that callers did not know how to use GoDaddy on their iPhones. The company rushed to retrain its agents to respond to the calls and pushed out an application allowing its users to control its service directly from the iPhone.

Certain emotions are now routinely detected at many call centers, by recognizing specific words or phrases, or by detecting other attributes in conversations. Voicesense, an Israeli developer of speech analysis software, has algorithms that measure a dozen indicators, including breathing, conversation pace and tone, to warn agents and supervisors that callers have become upset or volatile.

The real issue with artificial intelligence, as with any technology, is how it will be used. Automation is a remarkable tool of efficiency and convenience. Using an A.T.M. to make cash deposits and withdrawals beats standing in line to wait for a teller. If an automated voice system in a call center can answer a question, the machine is a better solution than lingering on hold for a customer service agent.

Indeed, the increasing usefulness of artificial intelligence — answering questions, completing simple tasks and assisting professionals — means the technology will spread, despite the risks. It will be up to people to guide how it is used.

“It’s not human intelligence, but it’s getting to be very good machine intelligence,” said Andries van Dam, a professor of computer science at [Brown University](#). “There are going to be all sorts of errors and problems, and you need human checks and balances, but having artificial intelligence is way better than not having it.”

A version of this article appeared in print on June 25, 2010, on page A1 of the New York edition.

Times Reader 2.0: Daily delivery of The Times - straight to your computer.

Subscribe for just \$4.62 a week.

Related Searches

- * Artificial Intelligence
- * Computers and the Internet
- * Science and Technology
- * Software

- [+ Get E-Mail Alerts](#)

COMMENTS (114)

E-MAIL

PRINT

REPRINTS

INSIDE NYTIMES.COM

MOVIES »



In 'Restrepo,' Loving Life and Making War

OPINION »

Can Obama Create More Jobs Soon?

A Room for Debate forum on what it would take to put more people to work.

MUSIC »



Solemn, Strange and Demanding Sounds

U.S. »



Coffee with Lobbyists

OPINION »



Op-Ed: The Forgotten War, Remembered

WORLD »



Dressed Up and Ready, Canada Town Is Stood Up

[Home](#) | [World](#) | [U.S.](#) | [N.Y. / Region](#) | [Business](#) | [Technology](#) | [Science](#) | [Health](#) | [Sports](#) | [Opinion](#) | [Arts](#) | [Style](#) | [Travel](#) | [Jobs](#) | [Real Estate](#) | [Autos](#) | [Back to Top](#)

Copyright 2010 The New York Times Company | [Privacy](#) | [Terms of Service](#) | [Search](#) | [Corrections](#) | [RSS](#) | [First Look](#) | [Help](#) | [Contact Us](#) | [Work for Us](#) | [Advertise](#) | [Site Map](#)

MORE IN SCIENCE (3 OF 42 ARTICLES)

Advances in Oil Spill Cleanup L Valdez

[Read More »](#)