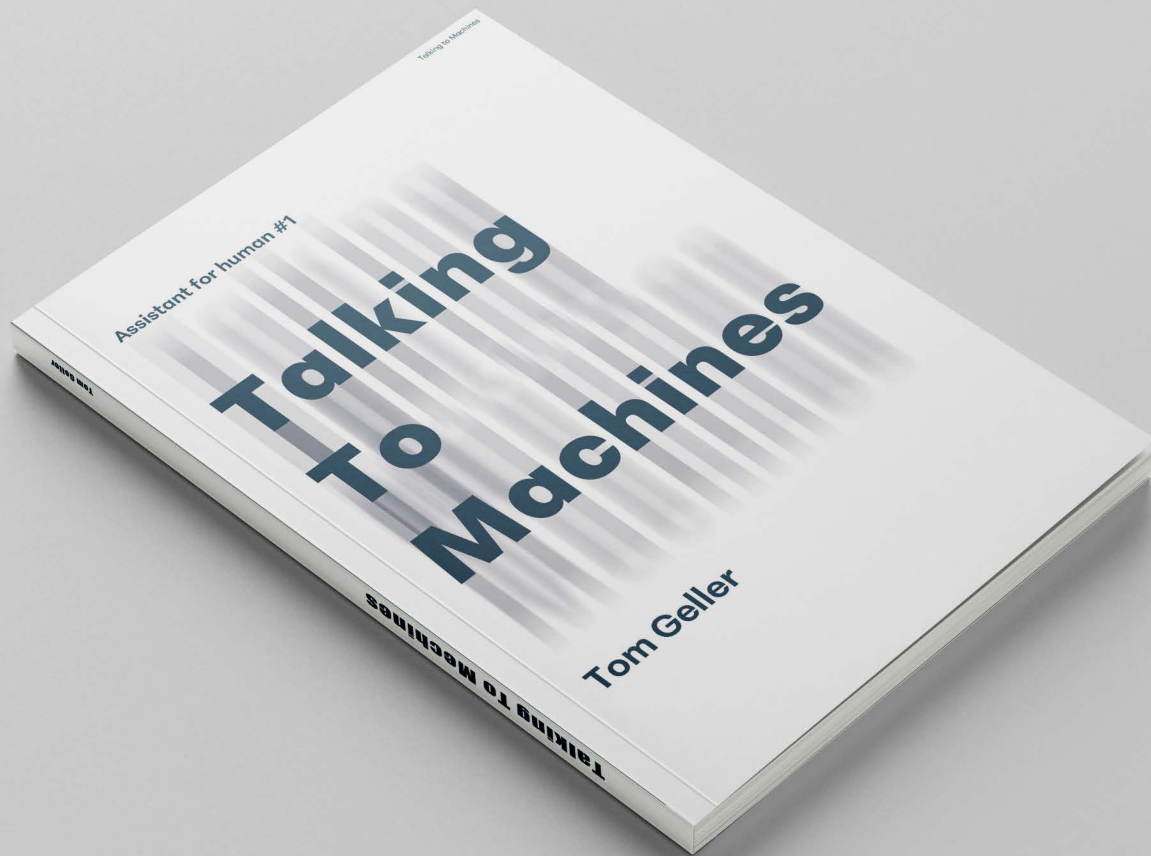




Talking To Machines

Assistant for human #1

Talking To Machines

Tom Geller

Talking To Machines

Tom Geller



Voice recognition programs like Siri are now capable of understanding spoken commands, recognizing a conversation's context, and answering questions in a conversational manner. From scheduling one's appointments to how digital assistants like Siri emerge at the intersection of speech recognition, AI, cloud computing, and interface design, discover how technologies once conceived to solve have transformed the way humans interact with machines—and what that means for our future.

SIRI
Talking To Machines
Tom Geller

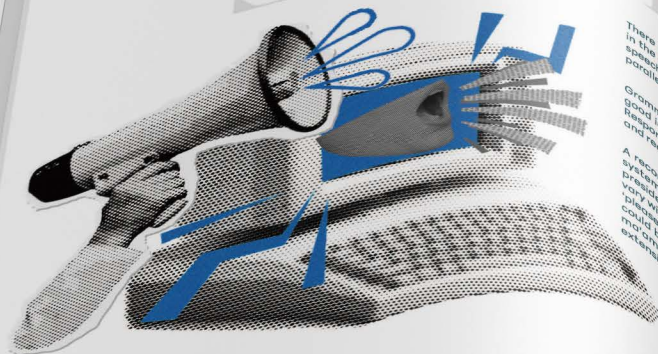


Assistant for human #1

Tom Geller

**Talking
To
Machines**

1 Executing Commands



There is a clear need between the spoken command and its fulfillment, though. The first step in the process is to convert the audio of speech into meaning. The two main applications of speech recognition—dictation and command recognition—have forced researchers to pursue parallel methods that balance vocabulary, accent, and context needs.

Grammar-based voice recognition is optimized for situations where the program has a very good idea about what the speaker will say. Its most common application is in Interactive Voice Response (IVR) systems, such as those that some airlines use to interpret spoken reservations and requests for information. These are often conversational.

A grammar-based voice system asks the speaker a question, then listens for the response. As a result, the system must understand only a limited vocabulary. But according to Dan Faulkner, vice president of Nuance, a company that makes software for the Enterprise Business Unit at Nuance, responses can vary widely. "When a phone prompt system tells you to say 'yes' or 'no,' you can say 'yes' or 'no,' but in the Southern states some people will say 'yes' or 'yup,' and in the Northern states some people will say 'yes' or 'yeah.' So even for something as simple as 'yes' or 'no,' you need quite an extensive list of phrases."

TALKING TO MACHINES

More than four million iPhone 4S's featuring Siri were sold during its first six weeks, and although use might seem to be as simple as speaking, it's not. It's a whole new world of possibilities, and it's going to be beyond simple description.

SIRI COULD HELP

Siri represents an important moment when voice recognition, machine learning, and artificial intelligence meet in a way that's not just useful, but transformative. As a product manager, I've seen the potential of this technology, and I've seen the potential of Siri. As a product manager, I've seen the potential of this technology, and I've seen the potential of Siri.

The news about Siri is that it works. People have tried to get computers to answer questions conversationally for at least 15 years, but only now has the technology reached a threshold where people overall like it.

The iPhone's popularity also gives intelligent software assistants wider exposure than they would get otherwise. Roger K. Moore, editor in chief of the Journal of Computer Speech and Language, points out that "the field of research hasn't changed dramatically. What's new is that Siri's brought together several complementary technologies together. Our business has been going for many years. Only now, with Siri, everybody knows about it."

OUR DAILY LIFE COULD BE CHANGE

CONTENTS

- 01** Talking to Machines
An introduction to how voice control's Siri enables
machines to understand human speech.
- 02** Executing Commands
Examining the process of converting spoken
commands into machine-executable actions.
- 03** Learning and Organizing
How a system organizes information and learns from
interactions to improve performance.
- 04** Measures of Success
Evaluating the effectiveness of conversational
systems through a variety of user experience
research measures, focusing on user expectations.

Talking to Machines

Voice recognition programs like Siri are new
examples of understanding spoken commands,
recognizing a conversation's context and answering
questions in a personable manner.



When Apple integrated Siri into the iOS operating
system for iPhone, it spurred iPhone owners to start
talking to their phones as well as through them.

The program, which converts spoken commands
such as "Schedule dinner with Lisa at 6 tonight"
into calendar appointments, Web searches, and
other tasks, is the most widely distributed example of a
cognitive assistant to date.



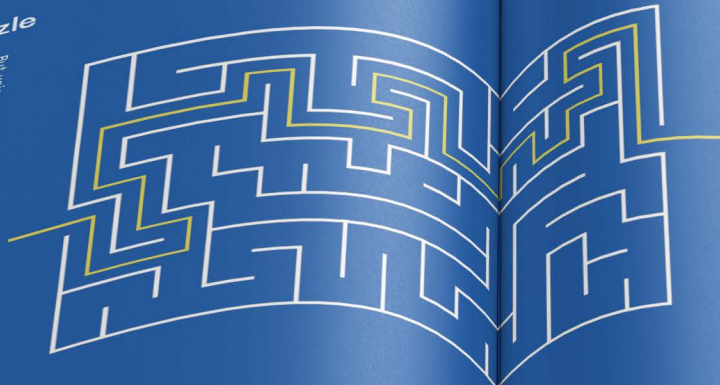


Learning and Organizing

© 2019 Apple Inc. All rights reserved. Learn more at <https://www.apple.com/education/>

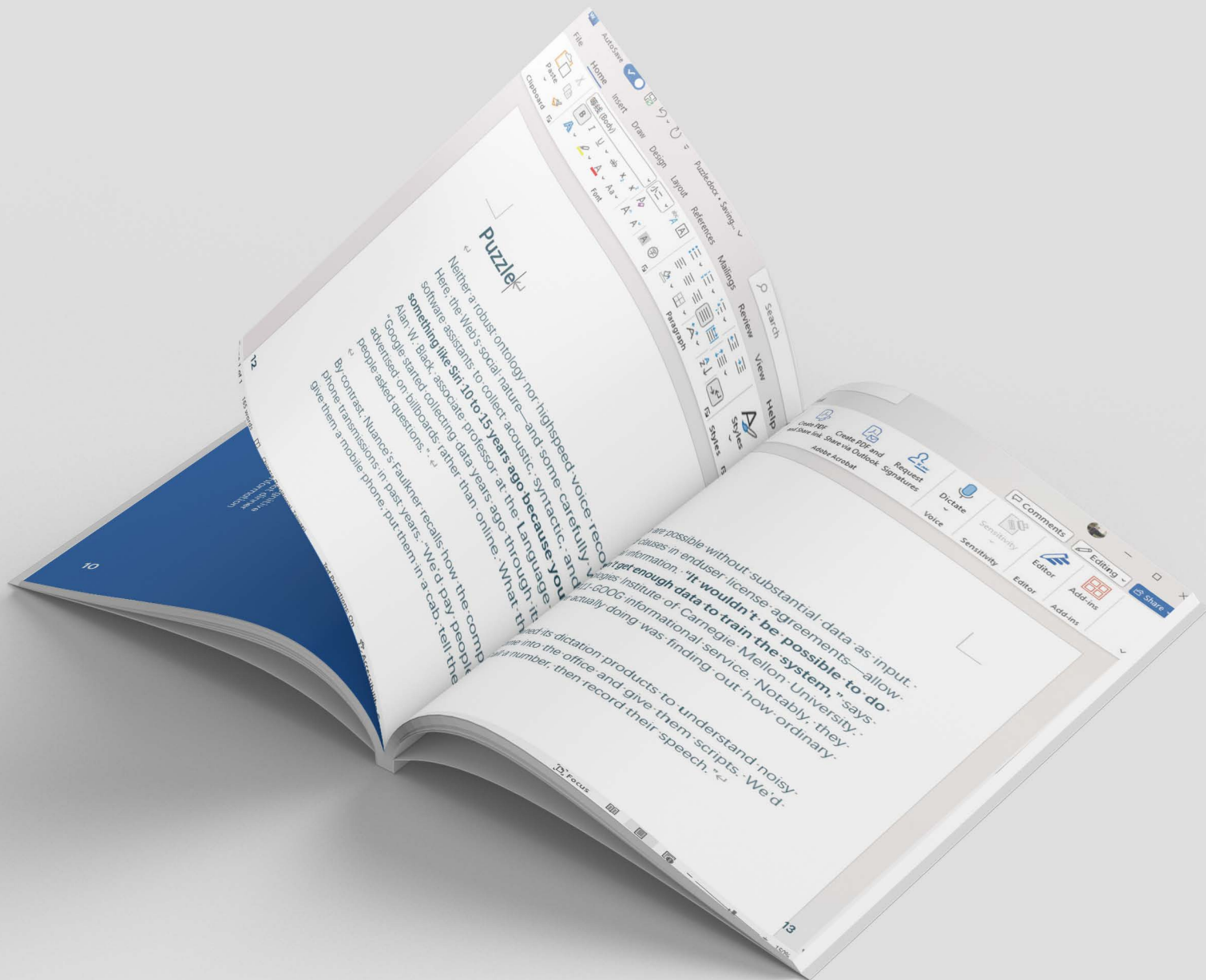
puzzle

But voice recognition is only a small part of the puzzle. Before a cognitive assistant can schedule that dinner with Lisa, it needs to understand that dinner is an event of limited duration that Lisa is a person whose contact information is found in the server's address book, and so forth.



Some of this understanding came from a Defense Advanced Research Projects Agency (DARPA)-funded project, Cognitive Assistant that Learns and Organizes (CALO), which was part of DARPA's PAL (Personalized Assistant that Learns) program. CALO's focus was not on voice recognition per se, or on natural language understanding. CALO's focus was computer interaction in general. Rather, it was about making computer systems learn in everyday settings, such as learning to recognize concepts, it then had to relate them through an underlying ontology, and trigger desktop applications and Web services.

investigator C. Raymond Perrault describes the challenge of solving for ambiguities in natural language. "Movie titles are typically just phrases in the language," Perrault points out, "so you could say 'Cat me two tickets to The Artist,' and the system would recognize that phrase as a movie title. On CALO we tried to solve for such ambiguities robustly, and would make it easy to build systems that do as well." Humans to build systems that do using context, while teaching a computer to learn them, requires addressing very long lists of such things as building a computer movies, and organizations, along with a solid categorization system to manage them. Ideas from CALO suggested a new approach to the development of a spoken interface to a set of Web services, eventually realized in Siri, which was built by a company that spun out of SRI International.





Measures of Success

Measures of Success

Ultimately, these steps serve the single goal of delivering a relevant, true, and useful response with appropriate speed. But at Carnegie Mellon's Black points out, a software engineer's job is not done if it only leaves a "One-shots" measure of employee dialogue systems is "one-shots". He notes, "and if that's the only goal, you can have an interaction that's successful and takes the time, yet is unpleasant. So satisfaction is another goal."

Black believes Siri delivers that satisfaction partly through its ability to "talk back". When Siri gives a response, he says, "it has a character. It wants to make you know who you are. You can tell it to call you 'Mister' or 'Don't bother' or 'You're late' and it will do it. It's a character. It also gives you a lot of information about the user's state. For instance, if

