

SUMMER 2020
WWW.SPEECHTECHMAG.COM

Speech TECHNOLOGY

>> As Speech Translation
Advances, Barriers
Continue to Fall

>> Companies Turn to IVAs to
Deal with COVID Surges

>> A New Approach to
Conversational AI

Teaching with Speech

Voice technology
has infiltrated
the learning
environment,
improving schools
for students and
educators alike



Join Us for These Learning Opportunities

Visit speechtechmag.com/webinars and start your education.

KEEP UP-TO-DATE on the latest technology innovations

- › Pick a topic that you want to explore.
- › Spend an hour listening to industry experts.
- › Get your questions answered in real time.
- › Sign up now and you will be sent a reminder prior to your session.

THAT'S ALL IT TAKES TO STAY CURRENT.

ATTEND A WEBINAR AND BE ENTERED TO WIN A \$100 AMAZON GIFT CARD.



Our 2020 Roundtable Topics Include:

- Apr 14: *The Future of Voice Technologies—An Executive Briefing***
Find out what's coming down the road from the people developing tomorrow's tools— and start planning for the voice-based future.
- Jul 14: *How AI Behavioral Modeling Improves Customer and Employee Satisfaction***
The move to work-from-home because of COVID-19 resulted in a rapid acceleration of the digital transformation needs of businesses around the world. The next step was figuring out how to manage employees from disparate locations to ensure continued customer satisfaction.
- Sep 22: *Merging Voice With Visual Displays***
Find out how you can use voice to transform everything from retail kiosks to interactions with IoT devices.
- Dec 08: *AI and the Power of Speech Analytics***
Take a deep dive into how AI is changing what speech analytics can tell us and how it can improve your relationship with your customers.

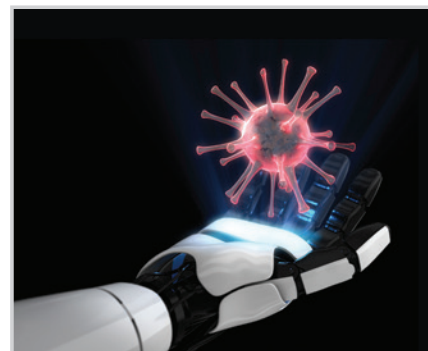
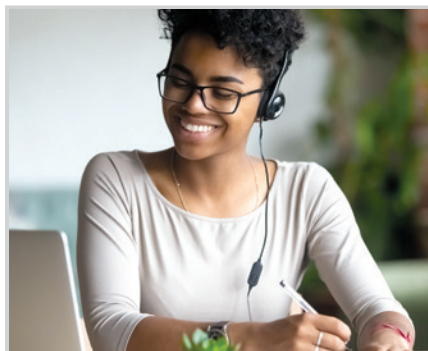
All events are available on-demand after the live dates.

For information on becoming a Roundtable Webcast sponsor, contact:

LaShawn Fugate

859-278-2223 X 104

lashawn@infotoday.com



COVER STORY

14 Teaching with Speech

Voice technology has infiltrated the learning environment, improving schools for students and educators alike.

BY ERIK J. MARTIN

FEATURES

20 As Speech Translation Advances, Barriers Continue to Fall

Technology improvements have pushed speech translation out to a smaller world.

BY PHILLIP BRITT

24 Companies Turn to IVAs to Deal with COVID Surges

Automated voice solutions helped contact centers handle rising contact volume during the pandemic.

BY LEONARD KLIE

FYI

8 A New Approach to Conversational AI Uses Symbolic Neural Networks

A joint venture between IBM and MIT relies on machine vision in combination with a broad AI that can multitask and cover multiple domains, but which also can read data from a variety of sources, whether the data is structured or unstructured.

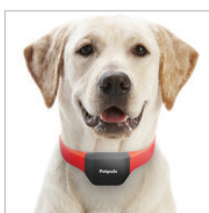
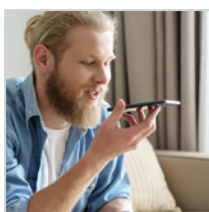
9 Speech Aids in COVID-19 Recovery

A voice-enabled app can help companies detect and monitor employees with the coronavirus; meanwhile, voice and speech samples are being collected from COVID-19 patients at various stages of the disease to help develop an algorithm for identifying coronavirus carriers.

10 Overheard/Underheard:

A New Collar Gives Dogs a Voice

An AI-powered dog collar uses voice recognition technology to detect and analyze five emotional states—happy, anxious, angry, sad, or relaxed—in dogs and tracks their physical activity.



COLUMNS

4 Editor's Letter

Thanks, COVID, for the Speech Boom
BY LEONARD KLIE

5 Business Case

IVAs: Using AI to Serve Customers and Contact Centers
BY DONNA FLUSS

6 View from AVIOS

Speech Technology Is Not Yet Conversational Technology
BY EMMETT COIN

12 Interact

Mitigating Bias in Speech Recognition Systems
BY KASHYAP KOMPPELLA

13 Voice Value

Automated Solutions Have Become Work-from-Home Agents, Too
BY NANCY JAMISON

29 Inside Speech

Amid the Pandemic, Speech Technologies Expand the Telehealth Options
BY KEVIN BROWN

30 Forward Thinking

How Can Speech Technology Help?
BY MOSHE YUDKOWSKY

Thanks, COVID, for the Speech Boom

I would never say that the coronavirus pandemic was a good thing, but if there is one positive that we can draw from the outbreak, it would be the very widespread adoption of speech technologies that would likely not have happened without it. Throughout this issue, we highlight many ways in which speech technologies have helped companies and their customers and employees, teachers and their students, healthcare providers and their patients, and countless others interact more effectively, more safely, and more cost-efficiently. Speech technologies are even being used to identify people who are likely infected with COVID-19 so companies and employees can go back to work with confidence. And once employees are in the office, speech is also helping to stop the transmission of the disease from shared surfaces.

“The COVID-19 pandemic sped up the adoption rate of speech technologies among the population more broadly,” observes Matt Muldoon, North American president at text-to-speech systems provider ReadSpeaker, in this issue’s cover story, “Teaching with Speech” (page 14).

In the same article, Sara Marie Hasbun, founder and managing director of Meridian Linguistics, echoes that sentiment: “The coronavirus has sped up the adoption of speech technology, mostly just because people are more willing to try something new,” she states.

Among its uses in educational settings, speech technology is helping with transcription and captioning of lectures; translations; monitoring the well-being of students; as a reading and writing aid; in distance learning (something which became the normal amid coronavirus school closures); and as an assistive technology for disabled students.

Speech technology’s use as a translation aid also gained solid ground in the COVID-19 world as face-to-face business and social meetings have given way to digital communications. Our second feature, “As Speech Translation Advances, Barriers Continue to Fall” (page 20), highlights how contact centers were able to address new kinds of queries from customers all over the world. “The pandemic has provided us an opportunity to build our knowledge rapidly,” says Jen Snell, vice president of product marketing for Verint Intelligent Self-Service, in the article.

For more on how speech technologies helped contact centers handle rising interaction volumes, read our third feature, “Companies Turn to IVAs to Deal with COVID Surges” (page 24). Speech technologies enabled companies to quickly pivot their contact centers to work-from-home models and then to deflect huge amounts of queries away from smaller agent pools that were challenged in an untold number of ways. Amid the pandemic, intelligent virtual assistants became “mission critical” for companies, the article points out.

And while the coronavirus hastened their adoption, none of these use cases would have been possible had it not been for the many technology advances that were already under way long before the spring, when the coronavirus really took hold. Each of these technologies—and many more not discussed in these pages—have been getting better over the years, thanks to more computing power that enables applications to be housed on smaller devices. Thanks to advances in artificial intelligence, machine learning, and natural language processing, speech applications are more accurate, better able to isolate speech in noisy environments, better able to derive insights from the words spoken and the manner in which they are delivered, and better able to surface information that can improve customer interactions.

It’s unfortunate that it took a worldwide crisis for organizations to change, but we are glad that they did. And while we hope that the world never has to weather a storm of this magnitude again, we can take comfort in the fact that if it does, organizations will be much better prepared thanks to speech technologies.

As this issue was going to press, the United States was in the throes of several regional outbreaks, but all experts agree that once the country settles into our new normal, speech technology adoption will increase even further.

While no one can forecast the future, it is certainly not a stretch by any means to say that speech tech is here to stay, and the coronavirus had a large role in that.

LEONARD KLIE, Editor
lklie@infotoday.com

Editor Leonard Klie
lklie@infotoday.com

Managing Editor Chris Cronis
ccronis@infotoday.com

Senior Designer Laura Hegyi

Contributors Phillip Britt, Kevin Brown, Emmett Coin, Donna Fluss, Nancy Jamison, Kashyap Kompella, Erik J. Martin, Moshe Yudkowsky

Editorial Advisory Board The Editorial Advisory Board of *Speech Technology* magazine is composed of prominent figures in research, development, and applications of speech technology. The members will assist the magazine’s editorial staff by recommending articles and column topics or prospective authors, and offering advice on technical matters and industry trends. Additional responsibilities will include critiquing the magazine’s content and design. If you are interested in participating, contact Editor Leonard Klie (lklie@infotoday.com), *Speech Technology* magazine, 143 Old Marlton Pike, Medford, NJ 08055.

PUBLISHER / ADVERTISING SALES / PRINT & ONLINE
Publisher/Advertising Sales, *Speech Technology Magazine*
 LaShawn Fugate lshawn@infotoday.com (609) 654-6266

MARKETING
Marketing Manager, Events & Circulation Sheila Willison
 sheila@infotoday.com • **Director of Web Events** DawnEl Harris
 dawnel@infotoday.com

CORPORATE HEADQUARTERS
 Information Today, Inc., 143 Old Marlton Pike, Medford, NJ 08055

EXECUTIVE MANAGEMENT
President and CEO Thomas H. Hogan • **Chairman** Roger R. Bilboul •
CFO Mike Flaherty • **Group Publisher** Bob Fernekees

INFORMATION TECHNOLOGY
Vice President, Information Technology Bill Spence

PRODUCTION
Production Manager Tiffany J. Chamenko •
Ad Trafficking Coordinator Jacqueline Crawford

Speech Technology (ISSN: 1088-5803; USPS 23156) is published four times a year, January (Winter), April (Spring), July (Summer), and October (Fall), by Information Today, Inc., 143 Old Marlton Pike, Medford, NJ 08055 USA; Phone: (609) 654-6266; Fax: (609) 654-4309; Internet: www.infotoday.com. Registered in U.S. Patent & Trademark Office.

Periodicals postage paid at Vincentown, NJ, and additional mailing offices.

© Copyright 2020, Information Today, Inc. All rights reserved. No part of this publication may be reproduced in whole or in part in any medium without the express permission of the publisher. PRINTED IN USA

POSTMASTER: Send address changes to *Speech Technology*, P.O. Box 3006, Northbrook, IL 60065-3006

Rights and Permissions
 Permission to photocopy items is granted by Information Today, Inc. provided that a base fee of \$3.50 plus \$0.50 per page is paid directly to Copyright Clearance Center (CCC), or provided that your organization maintains an appropriate license with CCC.

Visit copyright.com to obtain permission to use these materials in academic coursepacks or for library reserves, interlibrary loans, document delivery services, or as classroom handouts; for permission to send copies via email or post copies on a corporate intranet or extranet; or for permission to republish materials in books, textbooks, and newsletters.

Contact CCC at 222 Rosewood Drive, Danvers, MA 01923; (978) 750-8400; Fax: (978) 646-8600; www.copyright.com. If you live outside the USA, request permission from your local Reproduction Rights Organization. (For a list of international agencies, consult www.ifrr.org.)

For all other requests, including making copies for use as commercial reprints or for other sales, marketing, promotional and publicity uses, contact the publisher in advance of using the material. For a copy of our Rights and Permissions Request form, contact Lauree Padgett, lpadgett@infotoday.com.

Online Access Visit our Web site at www.speechtechmag.com

Contents also available online under direct licensing arrangements with EBSCO, NewsBank, ProQuest, and Gale, and through redistribution arrangements with information service providers, including Dow Jones Factiva, LexisNexis, OCLC, STN International, and Westlaw.

Subscription Information
SUBSCRIPTIONS: Print version free to qualified recipients within the U.S. Digital version free to qualified recipients worldwide. Subscription rates for non-qualified subscribers: U.S. subscription rate—\$79.95; Canada and Mexico—\$89.95; overseas delivery—\$145. All rates to be prepaid in U.S. funds. Subscribe online: www.speechtechmag.com/subscribe.

BACK ISSUES: \$8 per copy, prepaid only.

CHANGE OF ADDRESS: Mail requests, including a copy of the current address label from a recent issue and indicating the new address, to *Speech Technology Magazine*, P.O. Box 3006, Northbrook, IL 60065-3006 or call (847) 559-7301.

Reprints For quality reprints of 500 copies or more, contact Sheila Willison at (877) 993-9767, sheila@infotoday.com.

Disclaimers Acceptance of an advertisement does not imply an endorsement by the publisher. Views expressed by authors and other contributors are entirely their own and do not necessarily reflect the views of the publisher. While best efforts to ensure editorial accuracy of the content are exercised, publisher assumes no liability for any information contained in this publication. The publisher can accept no responsibility for the return of unsolicited manuscripts or the loss of photos.

Privacy Policy Occasionally we make a portion of our mailing list available to organizations whose products or services we think might be of interest to our customers. If you do not wish to receive such mailings, please send a copy of your mailing label with a request to be removed from the third-party mailing list to *Speech Technology* magazine, Customer Service, P.O. Box 3599, Northbrook, IL 60065-9736 or call (800) 248-0588 or (847) 559-7301.

Editorial Office 143 Old Marlton Pike, Medford, NJ 08055; www.speechtechmag.com





IVAs: Using AI to Serve Customers and Contact Centers

True omnichannel customer service is becoming closer to reality

Intelligent virtual agents (IVAs) represent the future of omnichannel self-service, a new standard of voice and digital self-service in a channel-optimized format. They can also provide contact centers with context-aware guided support and relevant information for each individual customer interaction, for agent-assisted or escalated self-service interactions. And with the COVID-19 pandemic putting increased stress on businesses, customer service departments, and remote workers in need of support, these are welcome developments (see related story on page 24).

DMG Consulting defines an IVA this way:

A system that utilizes artificial intelligence, machine learning, advanced speech technologies (including natural language understanding/natural language processing/natural language generation [NLU/NLP/NLG]) to simulate live and unstructured cognitive conversations for voice, text, and digital interactions via a digital persona.

IVAs—which are powered by sophisticated underlying technologies that enable them to simulate cognitive reasoning and respond to human beings in a conversational manner in phone and digital interactions—support omnichannel environments so customers can start in one channel and move seamlessly to another. IVAs leverage machine-readable, context-aware knowledge bases (or other data sources and repositories) to store and retrieve the data needed to respond in a personalized and contextually relevant manner to human questions or input. They can mimic human cognitive functions such as understanding and sentiment, and they use machine learning to continuously improve their accuracy and effectiveness over time; their intelligence is continually evolving based on knowledge gained from each interaction, which is assimilated and leveraged in future conversations and transactions. IVAs may include visual representations—i.e., avatars—on websites or within mobile apps, or they can just be speech-enabled.

There are dozens of artificial intelligence (AI) technologies available today, but the three that are core for IVAs are NLP/NLU/NLG, real-time analytics, and machine learning (ML).

The NLP grouping of technologies—which include transcription, speech-to-text, and text-to-speech—allow organizations to understand what customers are saying; they're designed to find meaning and insights in conversations, whether spoken or written. NLP enables computers to understand the meaning without a predefined syntax for the content. It also allows computers to respond to people in their own language. Practical applications of these technologies

in contact centers include speech analytics, IVAs, robotics, and other capabilities.

Real-time analytics encompasses a highly diverse group of technologies and applications. Real-time analytics frequently takes and acts upon the input from an NLP solution. It may also draw upon historical data, a customer relationship management (CRM) solution, a sales system, marketing databases, inventories, etc., to determine the most appropriate action to take. The challenge is to do this fast enough to enable a live or automated “agent” to act in near real time. Billions of dollars have been invested in a large variety of real-time analytics solutions over the past 20 years. All of the older applications were rules-based, and most were too slow to be helpful in a real-time servicing scenario, which is what is needed in a contact center.

Machine learning, an essential component of AI, is a highly complex set of algorithms that can automatically “learn” and identify trends, patterns, opportunities, etc., in a dataset. (A dataset may include recordings or transcribed texts, such as tweets and emails.) ML can operate in three modes: supervised, semi-supervised, and unsupervised. It can be used to provide the most current information about the customer journey, which should be a priority for all service organizations. Machine learning is already being used in a growing number of contact center solutions, including speech analytics and workforce management.

Additionally, in the era of AI and real-time analytics, data repositories are a must-have. AI-enabled solutions enhance self-service and agent-assisted interactions by accessing knowledge bases, other data repositories, and third-party sources to locate and retrieve the most appropriate resource material to provide to a prospect or customer, or to answer questions.

Final Thoughts

IVAs are the future of omnichannel self-service and for helping employees, not just contact center agents, perform their jobs. In the past couple of years, billions of dollars in research and development have been invested in AI-related products, and IVAs have been the beneficiaries of a lot of this investment. There's still a lot more work to be done before IVAs can come close to replicating an interaction between a live agent and a customer, but some of the IVAs are already significantly more effective, responsive, flexible, and accurate than IVRs. ☒

Donna Fluss (donna.fluss@dmgconsult.com) is president of DMG Consulting, a provider of contact center, analytics, and back-office market research and consulting.



VIEW FROM AVIOS

By Emmett Coin

Speech Technology Is Not Yet Conversational Technology

The industry still lacks tools to support true conversation

Speech technology has become ubiquitous; everyone is suddenly speaking to machines. Voice is now a thing. Devices like Amazon Echo and Google Home are existing technology, not gadgets of the future.

Automatic speech recognition (ASR) is good—even astonishingly good if you’re in the sweet spot of the training space. Text-to-speech (TTS) is not unpleasant. Yet something is missing.

Where is the conversation? Why can’t we converse with machines?

Conversational technology remains at the same level as interactive voice response (IVR) design from two decades ago. Granted, some narrow improvements, like “intent recognition,” have emerged to ease the task of understanding what an utterance means. But the overarching conversational design is still a flowchart labyrinth: IF heard X, THEN say Y, and JUMP to a new position (state) in the “conversation” flowchart.

For the most part, the voice interactions we have today involve a single human utterance: a command or query, after which the human gets a simple response—lights turn on, Rachmaninoff is played, or a question is answered (“there are 3.26156 light years in a parsec”). If you need to do more or ask a follow-up, you start with a fresh slate. You must reformulate a self-contained, complete, unambiguous utterance. (Note: Some skills/actions execute multiple turns; I’ll address that later.)

My daughter, a professional psychologist in D.C., related to me a suggestion from one of her industry newsletters describing how to get into artificial intelligence (AI). She knew I’d chuckle. It quoted Mark Cuban explaining that anybody can get into AI these days: just write a script for Alexa. Amusing, but it made me wonder: If people with their finger on the pulse of high-tech venture capital think

that an IVR script is tantamount to AI, then what does the rest of the public think it is?

Those who have created skills/actions with multiple turns know it is just coding a flowchart with all of the potential twists and turns that could possibly happen. It is tedious and fraught with a combinatorial explosion of details. At best you will anticipate a subset of those details (experience helps here). It must be tested with a range of people that do surprising (human) things that the coder never dreamed of.

This is why out of thousands of multi-turn skills/actions, most of them (90 percent?) are awful to the point of being unusable. And of the remainder, most of those are bad enough to be abandoned after a few encounters. The

remaining few fit into a category we could characterize as usable. The point is that simple command or query applications are relatively easy to create. But conversations (multi-turn applications) remain difficult and expensive and require deep expertise.

I and the AVIOS community believe that the industry lacks tools to support true conversation. In my talks I often ask the audience to imagine the conversation between Albert Einstein and a 3-year-old. No one doubts what would happen. AI and the kid would

find a topic where their experience overlapped. Their conversation would wander around the intersection of their expertise (e.g., Einstein might say “Venn I was a child...”). Occasionally the conversation would step outside their common experience (curved spacetime, or bubblegum-flavored yogurt), and one of the conversational partners would offer to expand the envelope of that common knowledge (“like a ball on a stretched sheet” or “it is fun to squeeze out of a tube”).

Even though their worldviews are vastly different in scope, they can both effortlessly use their generic (meta-)

The voice interactions we have today involve a single human utterance, after which the human gets a simple response.

conversation engine. Neither had to study or make complex plans for how to respond. It followed a meta-level pattern. Yet their “specific” conversation is unique in the universe. How does that happen?

Often in science fiction people converse with machines. Once one excludes the extraordinary machines (such as *2001: A Space Odyssey*’s HAL, *Ex Machina*’s Ava, *Star Wars*’ CP30, Isaac Asimov’s R. Daneel Olivaw, *Lost in Space*’s M3-B9, etc.), I like to imagine the sorts of human-machine exchanges we could have using today’s technology.

I think the fictional robot that comes closest to approximating such exchanges is Robbie the Robot from the classic 1956 film *Forbidden Planet*. I recommend the film in general (look up *Forbidden Planet* on IMDb for Robbie’s dialogue), but there are five or six interactions with Robbie that are noteworthy, and in particular a couple involving extended conversations to complete a task. Both of the tasks involve manufacturing something. Both of the tasks are in very different domains. Both have missing constraints as well as presumable constraints. There is a structure to his conversation at a meta-level.

In the world of skills/actions of today, these tasks would be crafted in completely isolated, elaborate silos. Today’s tools do not scale for building conversations. A pragmatic example is a sales job. Two positions are open: in the mattress department and in the shoe department. Other than the details of the products (pillow top, leather soles, queen-size, slip-on, etc.) how much do you have to be trained to have a conversation with the customer?

As a salesperson all you really need is a spreadsheet with columns for product names, parameter names,

price, and so on. All those things can be slotted into a standard, meta-customer interaction. The goal is the same: match the features to the customer desires and convince them to buy.

In an object-oriented programming paradigm, the developer would select an existing “salesperson” base class. Then the developer would create a new “shoe salesperson” class that would inherit all of the common behaviors. This new class would add nuances required for selling shoes.

In fact, in this thought experiment we could imagine that an “apparel salesperson” base class existed (of course inheriting behaviors from the “salesperson” class), and that the developer would choose that one on which to build their “shoe salesperson.”

A scheme like this fixes another major problem: multiple personality disorder. Alexa sounds like Alexa no matter what she talks about. But every developer imbues a different personality/manner to the application (unique presumptions and responses hard-coded into the flowchart). The human is forced to be a bit of a psychiatrist and recognize which personality they are speaking with.

Some say that neural nets will learn how to converse from transcribed conversations. This is not the place to detail the very tall hurdles that underlie this approach. But at AVIOS conferences we do dig into the details like this when we talk about practical, working technology with the leaders at the cutting-edge. Visit www.botsandassistantsconf.com and join us! [✉](#)

Alexa sounds like Alexa
no matter what she talks
about, but every developer
imbues a different
personality/manner to
the application.

Emmett Coin is the founder of ejTalk, where he researches and creates engines for human-computer conversation systems. Coin has been working in speech technology since the 1960s at MIT with Dennis Klatt, where he investigated early ASR and TTS.

A New Approach to Conversational AI Uses Symbolic Neural Networks

A new approach to artificial intelligence could provide customer experience and digital transformation specialists with more accurate intent recognition, according to a new report from Opus Research.

In the report, “Neuro-Symbolic Artificial Intelligence and Potential Impact on Conversational Commerce,” Opus founder and lead analyst Dan Miller details a joint venture between IBM and MIT that relies on machine vision in combination with a broad AI that can multitask and cover multiple domains, but which also can read data from a variety of sources (text, video, audio, etc), regardless of whether the data is structured or unstructured.

This approach could enable users to “do more with less” and provide for greater transparency and privacy, according to Miller.

Employing the neuro-symbolic approach to conversational AI could enable companies to “add common sense” to their chatbots, intelligent virtual agents, interactive voice response systems, and the prompts provided to live agents.

“AI has been subjected to boom-then-bust cyclicity since the 1970s, but this time feels very different,” Miller says. “The MIT-IBM Watson AI Lab, the joint research effort forged in 2017, has put significant resources into a new approach that combines the probabilistic pattern recognition capabilities of today’s deep neural networks and deep understanding with the once-prevalent ‘symbolic’ approach to AI that is based on representations of problems, logic, and search that are considered more human-readable. Their approach offers the possibility of levels of accuracy that approach 100 percent for functions like image recognition or natural language understanding.”

With the new approach, companies will be able to improve their automated systems’ ability to understand the actions and utterances of those with whom they are trying to communicate, according to Miller. “The technology can enable companies to step up their level of customer care. It offers

improvement over being tossed around between live agents and chatbots.”

Another advantage of incorporating neuro-symbolic AI into customer conversations is architectural. Resources can be deployed in multiple clouds, as well as on processors embedded in smart endpoints, including smartphones, smart speakers, and automotive information/entertainment consoles. Thus, the ability to support ubiquitous, omnichannel, device-agnostic self-service strategies, including contact center, is simplified.

The technology requires far fewer resources than the technology used today to understand customer actions, Miller adds. In fact, neuro-symbolic systems could achieve better results in natural language processing and other tasks with as little as 1 percent of the data required by current alternatives, according to David Cox, IBM Research’s director of the MIT-IBM Watson AI Lab.

In 2018, the lab published research regarding the Neuro-Symbolic Concept Learner, a resource that eases the machine training. “As

conversations become the prevailing engagement model between brands and their customers, it will become increasingly apparent that successful task completion ultimately relies on solving a succession of non-related puzzles,” Miller explains.

Still a Little Futuristic

While the neuro-symbolic AI technology is in its embryonic stages, the speed of technology advancements today means it could enable very practical conversational commerce in a few years.

Less than a decade ago, machine vision was relatively new—and extremely expensive—in the robotics realm. But now several companies offer the technology with continuing refinements for better color definition, shape recognition, etc., enabling vision-capable robots to perform tasks that previously could only be done by humans.

Miller doesn’t think it will take that long for the IBM/MIT



“AI has been subjected to boom-then-bust cyclicity since the 1970s, but this time feels very different.”

technology to become commercially viable in conversational AI applications.

Concerns over consumer privacy could provide a catalyst to advance the adoption of the technology, according to Miller. “The neuro-symbolic AI requires a small fraction of the data that [deep neural network]-based approaches have relied on. This is vitally important because of the heightened concern surrounding privacy and protection of personal information. Enterprises had grown comfortable collecting or purchasing voluminous amounts of personal data about their best customers in order to provide highly personalized service.”

Miller adds: “Ideally, companies need to work exclusively with first-person data, meaning the information that an individual provides directly and voluntarily in the course of spoken or text conversations. This has no clear connection with the use of symbolic processing, per se, but in an era of heightened sensitivity to personal privacy, any approach that minimizes reliance on third-party data on individuals is a positive.

“In the years ahead, as it is productized and commercialized, we expect the technology to be a source of measurable improvement in customer experience and agent productivity” —*Phillip Britt*

Speech Aids in COVID-19 Recovery

As America slowly started returning to work after months of COVID-19-related shutdowns, speech technologies have been playing a huge role in the recovery efforts.

In early July, Sonde Health launched Sonde One, a voice-enabled app to help companies detect and monitor employees with the coronavirus.

Sonde One leverages the company’s advanced vocal biomarker platform and machine learning technology. With a single six-second voice sample, the app can detect respiratory symptoms such as coughing, shortness of breath, or chest tightness or pain. The app combines this voice analysis with a COVID-19 questionnaire and the subject’s body temperature to identify which employees have COVID-19 within one minute. Employers can deploy the solution to any number of staff anywhere.

“At Sonde, we believe that voice is a vital sign and a meaningful predictor of health,” said David Liu, CEO of Sonde Health, in a statement. “By analyzing a few seconds of speech, we can detect subtle changes in a person’s voice caused by common symptoms of respiratory disease. We built this capability into the Sonde One app so organizations can provide their employees with simple and fast COVID-19 monitoring in their pocket.”

To bring Sonde One to market, Sonde partnered with corporate wellness solutions provider Wellworks for You to make the tool available to its clients and their 1.4 million participating members.

“After months of remote work and distributed workforces, many employers want to get their employees back in the office; they’re just not sure of the best way to do it,” said

Thomas Tegler, president of Wellworks for You, in a statement. “Sonde One is an easy, non-invasive way for organizations to manage employee health risks during the pandemic and get workers back to the office safely.”

SHI International, a provider of technology products and services, is the first company to deploy Sonde One. The company will begin implementing the Sonde One app in August as it gradually begins bringing employees back to the workplace, starting with its Somerset, N.J., headquarters.

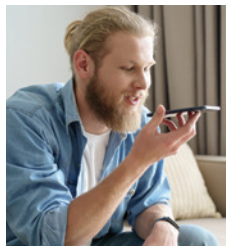
“As we navigate this pandemic, our employees’ health and wellness is our top priority,” said

Michael Haluska, vice president of human resources at SHI International, in a statement. “From the conference room to the warehouse floor, the Sonde One app will help us standardize our return-to-work policy to eliminate any confusion and keep all employees feeling safe.”

Similar research to detect potential carriers of the coronavirus using voice and speech processing was being conducted at Israel’s Afeka Tel-Aviv Academic College of Engineering with the Israeli Ministry of Defense.

“The target of the project is to develop an innovative algorithm that will be a breakthrough in the identification of potential carriers of the COVID-19 virus,” says Ami Moyal, a speech processing expert and president of Afeka College of Engineering.

Voice and speech samples (recordings) are being collected from COVID-19 patients at various stages of the disease, as



“Many employers want to get their employees back in the office; they’re just not sure of the best way to do it.”

well as from healthy individuals. Another control group of COVID-19-negative patients suffering from respiratory symptoms (such as flu symptoms) are also being recorded to differentiate them from COVID-19-positive patients. The goal is to create a representative database to be used in the research.

The data will be processed using a variety of tools, advanced speech processing techniques, and machine learning algorithms.

And once employees are in the office, speech is also helping to stop the transmission of the disease from shared surfaces. Instead of having to touch light switches, temperature

controls, doorknobs, keypads, and similar items around the office, companies can deploy voice interfaces similar to those found in smart home systems.

“Voice has already made significant inroads into the smart home space and certain enterprise domains,” says Lian Jye Su, a principal analyst at ABI Research.

A global emphasis on working from home, combined with advice to minimize COVID-19 transmission from shared surfaces, helped cement the benefits of smart home voice control for millions of consumers and enterprises, Su says.

—Leonard Klie

Overheard/Underheard

A New Collar Gives Dogs a Voice

There is a lot of speech technology innovation around health monitoring and using voice to detect certain medical conditions, but the same technology is now being used to improve the relationship between dogs and their pet parents.

Petpuls is an artificial intelligence-powered dog collar that uses voice recognition technology to detect and analyze five emotional states—happy, anxious, angry, sad, or relaxed—in dogs and tracks their physical activity.

The collar pairs via Wi-Fi with an accompanying iOS or Android app and uses a proprietary algorithm along with a database of more than 10,000 bark samples from 50 breeds of dogs to identify the five emotional states.

Petpuls claims to have an accuracy rate of higher than 80 percent, and it grows even more accurate over time as it records and accumulates the dog’s “voice” data.

This, company executives say, will help dog owners better understand their dog’s emotions and what triggers them.

The collar also functions as an activity tracker, with a built-in accelerometer sensor that automatically calculates the number of calories a dog burns from exercise by the hour, day, week, or month.

Petpuls also comes with a “My Pet Registration” feature for dog owners to create unique profiles of their dogs within the Petpuls Community and a “Pet Favorites” feature that allows dog owners to keep track of their dogs’ food, toys, and other dog friends who are also registered as part of the Petpuls Community

“Petpuls was born from my desire to use AI and IoT

technology to help dog lovers better understand their canine companions,” says Petpuls founder and CEO, Vincent Kim. “The AI-driven data accumulated by the Petpuls collar can create more meaningful relationships between dogs and the humans they regularly interact with, [including] dog owners, walkers, caretakers, and veterinarians. We envision the Petpuls platform as a large global community of dog lovers, a community that is safe, fun, playful, and informative, where dog owners not only track their dogs’ emotions and exercise, but also meet other dog aficionados, set up dog play dates, and talk about the latest in dog food, fashion, healthcare, and more.”

Petpuls will retail for \$99.

And speaking of non-humans talking to us, if another tech startup has its way, product packaging will be soon be able to tell us all about what’s inside the box. Third Aurora is reportedly working on a packaging prototype that will enable consumers to talk to the products they’re about to purchase.

If the prototype comes to fruition, frustrated shoppers with product-related questions will no longer have to wait idly for store assistants to help them. The product will tell them the answer instantly, all through a speech-enabled smartphone interface.

The concept was challenging for Third Aurora, which had to bring together a complicated mix of cutting-edge technologies, including artificial intelligence, language processing, and cloud data.

“It’s an exciting concept and an elegant solution,” says digital marketing expert Dave Chaffey, cofounder of Third Aurora. “And it fits the ethos that the best place to find the answer is from the source.” —Leonard Klie





Get social with us.



spechtechmag.com/social-media



INTERACT

By Kashyap Kompella

Mitigating Bias in Speech Recognition Systems

The fault, dear Siri, is not in our voices, but in our choices

Image recognition systems today are unrecognizable from those of the pre-deep-learning phase. And as their performance skyrocketed, so did their adoption. But with increased adoption, their limitations have also become quite visible. One particular area of concern is the accuracy of image recognition systems when it comes to minority groups. The systems make many errors, with grave consequences. When a facial recognition system fails to verify the identity of a user, it can mean shutting off access to resources or

inconveniencing them in other ways. Thankfully, there is growing awareness of the perils of algorithmic bias. Companies and researchers are figuring out the best ways to mitigate it.

A machine learning model is only as good as the data that is used to train it.

But enough about image recognition. What about speech recognition systems? Do they also suffer from bias?

Sadly, the answer is yes. Of course, there is no doubt that automatic speech recognition (ASR) software has improved significantly in recent years because of machine learning. ASR is used in our smart speakers and in our phones' virtual assistants. It drives use cases such as speech-to-text conversion, audio captioning, assistive technologies, medical transcription, and more. But the specter of bias haunts all machine learning applications, including speech recognition. Simply put, speech recognition software does not work well for certain demographic groups.

Consider this study published in the Proceedings of the National Academy of Sciences by researchers from Stanford and Georgetown universities. The researchers tested ASR technology from five leading vendors (Amazon, Apple, Google, IBM, and Microsoft) and found these troubling results:

- The word error rate for African-American speakers is about twice as high as it is for white speakers for the five big ASR systems.
- Overall, up to 23 percent of transcripts of audio snippets from African-American speakers were unusable, but the

corresponding number for white speakers is less than 2 percent.

This study did not look at the experience of other ethnic groups, but I'd expect to see similar results for those as well. Clearly, there is a problem. Let's understand what's causing it and how we can address it.

A machine learning model is only as good as the data that is used to train it. The training dataset may be large, but if that data is largely from—or is skewed toward—a particular demographic, then the models work well for that group, but not so well for other groups. Bias is simply a manifestation of a higher error rate for the group underrepresented in the training data. The aforementioned study demonstrates the results.

What's the fix? The training data should include different accents, spoken language variants, and pronunciations of non-native speakers. But the solution to reduce AI bias is not just technical in nature. The teams building the speech tech systems should represent a diversity of voices and experiences. No one sets out to build biased systems, but in the absence of diversity, teams may not be aware of their unconscious biases and blind spots.

The problem of AI bias in image recognition systems has received widespread attention, but so far speech recognition systems, though they have similar problems, have not been under the same level of scrutiny. But as their adoption increases and they are deployed for use cases such as recruitment (where the stakes are high, and the laws are strict), these issues will assume greater urgency.

When speech recognition works well, user experience improves and customer satisfaction increases. But when it misfires or doesn't work, it's a minor annoyance for some, but for others it can be a dehumanizing experience. Companies selling ASR systems and the organizations using them need to understand the issues, have a plan for mitigating bias, and help make AI in speech technology more inclusive. ☒

Kashyap Kompella is CEO of rpa2ai Research, a global AI industry analyst firm, and is the co-author of *Practical Artificial Intelligence: An Enterprise Playbook*.



Automated Solutions Have Become Work-from-Home Agents, Too

Adoption and interest in AI-infused capabilities got better because things got worse

Recently the contact center industry has seen a rapid acceleration in the development and adoption of artificial intelligence-infused capabilities across the spectrum. Included in this mix is an array of applications, including bots and intelligent virtual assistants (IVAs), predictive routing, process automation, voice biometrics, assisted and unassisted robotic process automation (RPA), and automated forecasting and quality assurance.

The brain trust of the speech technology world applauded. Perhaps we had entered the heyday of what we have been working toward. Improvements to the technology, many more deployments, and solid use cases are now being used to showcase the benefits that AI-infused solutions bring. Topmost are that RPA bots and IVAs can complete tasks or provide assistance to agents, improving performance, productivity, and customer satisfaction while reducing costs.

Everything Changed

This might sound strange, but suddenly adoption and interest in these solutions got even better because things got worse. I'm referring to the COVID-19 pandemic, which overnight sent the entire contact center industry scurrying to dust off business continuity and disaster recovery plans and move

Robotic process automation and a virtual workforce provide additional advantages in times of uncertainty and change.

tens of millions of contact center employees to a work-from-home (WFH) model (see related story on page 24). Across the board, contact center providers fashioned programs and offers to assist in the move, creating packages with temporary complementary remote agent solutions, along with websites of information, best practices, and other assistance.

And guess who went home with those agents? Yes, the growing digital workforce became WFH agents as well. For example, NICE, in addition to its CXone@Home program, introduced its NICE Employee Virtual Attendant (NEVA) solution in a NEVA@Home package. And in these virtual suitcases were additional AI-infused tools to help the IVAs and their human counterparts, including desktop automation and knowledge management. Another example came from Verint, which, along with its intelligent agent, provided a WFH support program that included prebuilt COVID-19 categories for Verint Speech Analytics to help identify customer and employee

business challenges and aid compliance, as well as a knowledge management (KM) starter package.

While such benefits have been touted frequently by vendors and the press, robotic process automation and a virtual workforce provide additional advantages in times of uncertainty and change. When supervisors suddenly can't directly interact with agents, desktop process automation, interaction and speech analytics, and emotion detection provide a window into how agents are handling the change. Are they distracted? Are they maintaining compliance? What are the new topics and issues arising from customers stuck in the same situation? What new information needs to be placed into knowledge management systems and training materials? How do you keep agents engaged when they can no longer seek help by looking up from their workstations?

Employing a virtual workforce at home with agents solves for these situations by enabling agents to work "side by side" with another agent. IVAs work tirelessly to relieve agents of tedious tasks. They can dip into knowledge databases and prepopulate agent screens, provide next-best-action advice, help maintain compliance, and perform post-interaction follow-up.

IVAs can proactively help agents in times of stress. For instance, using data gleaned from speech analytics, the IVA might change the agent script for debt collection in real time to be more empathetic to a customer who has just lost a job or loved one. Similarly, when combined with fraud detection packages, the IVA can take the burden off of the agent by acting as the gatekeeper between an agent and someone trying to fraudulently gain access to an account. The IVA can authenticate, ask more questions, send the caller to a fraud squad, or simply coach the agent in real time on how to handle the situation.

The top priorities of improving performance and reducing costs in customer contact have been with us since the inception of the call center. In the past decade, these have been joined by the need to improve the customer and worker experience as well. Fortunately, the wealth of technologies under the umbrella of AI have finally shown us a way forward in addressing all four top concerns in running a customer care organization. Now with the switch to WFH, which shows all signs of remaining a core work strategy, it's time for organizations to add new technologies to assist with the change. ☑

Nancy Jamison is an industry director in information and communications technologies at Frost & Sullivan. She can be reached at nancy.jamison@frost.com, or follow her on Twitter @NancyJami.

Teaching with Speech

Voice technology has infiltrated the learning environment, improving schools for students and educators alike • BY ERIK J. MARTIN





At a time when Chromebooks are replacing textbooks and digital assistants are often more of a go-to resource than human instructors, it's hardly a shock to learn that 95 percent of teachers employ technology in schools today, according to a recent Common Sense Media study.

But many are surprised at the speed and degree to which one branch of innovation has grown in the field of education: Speech technology now commands an increasingly larger voice in classrooms and homework zones. From smart speakers and voice dictation software to text-to-speech apps and cutting-edge devices used by the disabled, students and educators are turning to a growing array of impressive tools and resources that can significantly aid learning.

Ahmed Ali, principal engineer for the Arabic language technologies group at the Qatar Computing Research Institute, part of Hamad Bin Khalifa University, is one of many experts who notes that speech tech has made great strides in the education space in recent years.

"Today, there are three main speech technologies visible in the daily routines of many students: converting speech to text in the form of automatic speech recognition (ASR), converting text to speech, and extracting paralinguistic features from voice, such as emotion, age, and gender," Ali



says. “ASR, in particular, has opened up exciting possibilities across the age spectrum, enabling, for example, the digitizing of lectures to display spoken words as text in university classes.”

In classrooms with younger students, Google Home and Amazon Alexa have earned coveted spots as teaching aids.

“These smart speakers are increasingly utilized as educational resources to enhance the curriculum and provide relevant information like the news and weather,” says Mai Ling Chan, director of growth and achievement at Cognixion, a provider of artificial intelligence-powered assistive communication solutions. “Because they respond to all voice commands, students often have the opportunity to interact with their classroom smart speaker and are learning how to interact and control these new technologies.”

Matt Muldoon, North American president at ReadSpeaker, providers of the TextAid software for voice-enabled assignments, textbooks, and other tools for students, explains that accessibility and equal access in the classroom have been front-line issues in the education field for many years and a primary driver for speech in the classroom.

“Text-to-speech technology was traditionally used as an assistive technology tool for students with learning disabilities, dyslexia, and literacy challenges. But the pandemic has shown that students of all levels benefit from using the technology,” Muldoon says. “The COVID-19 pandemic sped up the adoption rate of speech technologies among the population more broadly. Though many educators were already using platforms like Zoom to teach students, speech technologies have been quickly integrated to enhance the online learning experience in recent months.”

Echoing those thoughts is Margaret Curley, a speech pathologist and manager of remote education and telese-
rvices at TherapyTravelers, a placement

firm that connects speech and language pathologists and other therapists with school districts and students.

“Speech technology is becoming pervasive across grade levels,” she says. “For instance, in public schools, Individualized Education Program (IEP) students often have accommodations for speech-to-text via apps that enable you to answer questions aloud on video that can be rerecorded until the student likes the results. Meanwhile, in colleges, there are entire departments dedicated to such accommodations.”

“Speech technology is becoming pervasive across grade levels.”

Tools of the Trade

Speech tech apps, devices, and resources are more plentiful in general today, so it is not surprising that they are more pervasive and easily adaptable for use by students in the classroom and at home.

Major players in this realm include Nuance Communications’ Dragon line of voice dictation tools for education; Amazon Alexa skills like Ask My Class and ClassAlexa; Google Docs and its voice typing feature; Echo360, which transcribes text from classroom videos and creates content that’s easily referenceable and searchable; Don Johnston’s Co:Writer, which aids with writing via speech translation and recognition; Talk Technologies’ Steno SR, a private speech-to-text microphone that empowers students to compose text with their voice privately in the classroom and read that text back to them; Voice4U, an interactive communication app designed for autistic and English language learner students; and Claro Software’s ClaroRead, a text-to-speech app that caters to kids with attention and visual deficiencies.

Another standout is Otter for Education, a web app that provides

speech-to-text transcription and is used on hundreds of university campuses around the world, including at UCLA, where the app provides live transcription and collaboration assistance to students with learning disabilities who need academic accommodations for taking notes.

Muldoon points out that many students today avail themselves of speech tech via free text-to-speech tools that can be installed as plug-ins on laptops.

“They listen to the text spoken aloud as they read it. And some of these

tools include the ability to download content that can be listened to by students offline and at their convenience; simultaneous highlight, which highlights words and sentences in different colors; and screen/

page masks, which present a horizontal bar that can be moved along with the reading to emphasize the lines that are being read while shading the rest of the screen to reduce distractions,” Muldoon says.

The benefits made possible by speech tech in the educational setting can be invaluable to teachers and pupils alike. Take the perks of speech recognition alone; according to Canada’s Edmonton Regional Learning Consortium and Alberta Education, it can help students do the following:

- meet grade-level expectations and aid them in writing, including composing and editing;
 - switch between speaking and typing as needed;
 - augment legibility and written output;
 - create written output that more aptly represents their actual oral language aptitudes;
 - boost endurance and decrease fatigue by reducing writing by hand or keyboard; and
 - improve pronunciation by providing a less stressful setting, especially for English language learners.
- Amazingly, first graders averaged more



than a 97 percent accuracy rate on post-study reading tests after using speech recognition tools, according to a 2018 study by the Missouri College of Education.

“Speech recognition technology is supportive of the learner because it allows them to use personally, culturally relevant grammar,” said Elizabeth Baker, professor of literacy studies in the Department of Learning, Teaching, and Curriculum at the Missouri College, in the study. “Children all have different backgrounds, and this technology allows them to learn to read while using their own frame of reference.”

Text-to-speech, meanwhile, can boost reading fluency, comprehension, and decoding and position children to independently perform better using appropriate grade-level tools and materials. And the ability to have their text spoken aloud on demand can generate crucial oral feedback that could help improve writing and composition.

“Leveraging text-to-speech tools helps students improve their grades and retain more information,” Muldoon says. “With 60 percent to 80 percent of students not disclosing their disability as they transition into post-secondary education, solutions like ReadSpeaker’s webReader and/or docReader help level the playing field for first-time online learners.”

Cindy Jiban, principal academic lead for NWEA, a research-based nonprofit that creates academic assessments for students, notes that speech tech is also improving early reading skill evaluation.

“The previous reading assessment practice in first- and second-grade classrooms was to have each student sit down with the teacher individually, one at a time, and read aloud,” she says, noting that these sessions typically took a week and rendered the teacher unavailable for actual teaching. “But now, particularly in the primary grades, oral reading assessment is rapidly shifting to

capitalize on automatic speech scoring capabilities. Students read out loud into boom mics on headsets, making it feasible to capture and score individual readings. Today, the assessment takes half an hour total each season. And thanks to speech tech scoring, three weeks of reading instruction are saved.”

Supporting Those with Special Needs

Some students, of course, stand to reap greater speech tech benefits than others. These include the reading-challenged, those with difficulty putting their thoughts into text, those with dyslexia and attention disorders, IEP students who encounter hardship communicating, and the physically disabled.

“Imagine a student with a cleft palate or oral cancer using augmentative and alternative communication (AAC) devices to answer questions in

“The coronavirus has sped up the adoption of speech technology, mostly just because people are more willing to try something new.”

class,” Curley says. “Picture a student with dexterity challenges using speech-to-text to answer homework or write essays. And ponder a student with auditory processing challenges who can listen to a story while following along with the text.”

Indeed, speech tech can bring a multitude of benefits to students with disabilities to ensure they have an equal education.

“Today’s text-to-speech tools read text out loud and may include features like page masking and highlighting tools to help students focus. Other features include a dictionary, diction and pronunciation guides, enlarged text, and dyslexia guides to help a wide variety

of diverse students access and engage with the content,” Muldoon adds.

Sara Maria Hasbun, founder and managing director of Meridian Linguistics, a Hong Kong business that supplies speech technology companies with training data, seconds that sentiment.

“If a student is struggling to read because of dyslexia, low vision, or even blindness, they can at least keep up with the rest of their studies through text-to-speech functionalities,” she says. “Students who are deaf or hard of hearing, on the other hand, can benefit from speech-to-text technology. Lip reading is incredibly taxing, and the reality is that sign language interpreters are expensive and not always available. Speech technology can help fill that gap.”

Speech tech advancements can even help the more severely disabled—such as students with cerebral palsy, who might not be able to coordinate muscle movements—produce effective speech at the conversational level.

“Having access to an iPad with an app with eye-tracking technology, like Cognixion’s Speakprose, can increase her ability to formulate and produce sentence-level phrases,” Chan says. “The student can ask and answer questions in class and participate in activities that would be otherwise extremely difficult for her to physically execute.”

Speech in Distance Learning

More than 90 percent of the world’s student population was impacted by school closures due to coronavirus. But if the pandemic taught us anything, it’s that distance learning can be effective with the right resources, including speech tech, in place.

“The coronavirus has sped up the adoption of speech technology, mostly just because people are more willing to try something new,” Hasbun says.



“Everyone has had to learn to integrate new technology into their lives, whether it’s Zoom or Google Meet for videoconferencing or Slack for office chitchat. So everyone is already in the right headspace to be accepting of technologies that have always been there but might have been more daunting before.”

However, with some level of distance learning likely to remain in place for the foreseeable future, “it’s imperative that access to content, web accessibility, and individual learning continue to be top-of-mind issues for educators across the country,” Muldoon advises.

Ali believes the continued COVID-19 lockdown provides a “great opportunity for us to capitalize on the usage of voice. We can transcribe speech, learn different paralinguistic features, and even detect anxiety levels from speech using the right software. Imagine how a teacher can use such a platform to monitor the well-being of all students to facilitate and personalize their learning.”

Best Practices

Yair Shapira, CEO and founder of AmplioSpeech, a provider of digital speech therapy to K-12 students, reminds educators and parents that the school schedule—as of this writing—remains uncertain for the fall and beyond.

“For the most vulnerable students, who struggle academically and often personally and socially at all times, the current times of turbulence might be disastrous,” Shapira says. “These children have already suffered compromised service since March, lack service during the summer, and might dramatically regress further if they do not receive intensive instruction tailored to their specific needs.”

Fortunately, technology can be a solution.

“We all need to be open to using tools like speech tech to minimize the disruptions and provide continuous, regular,

and consistent services that will become a safe haven for fragile students regardless of the location of the student and staff,” Shapira continues. “Such services can only be provided if the entire therapy cycle is managed by a single platform, including resources, automatic documentation, reimbursement, self-practice, etc. Harnessing the power of AI in systems such as AmplioSpeech can not only accelerate students’ progress, it can also enable oversight, reduce workloads, and empower clinicians.”

“Don’t forget to show staff how [the technology] can free up more time for those high-quality teacher-student interactions.”

But before investing in speech tech, schools and families should do their homework.

“Ideally, the speech technology solution they select will include features that position students to get the most out of using the technology,” Muldoon says.

School administrators also need to take a hard look at the costs and learning curve involved, too.

“If introducing the use of technology is a bit uncomfortable for some staff, don’t forget to show them how it can free up more time for those high-quality teacher-student interactions that really matter,” Jiban suggests.

It’s important as well to develop a school- or district-wide policy on appropriate technology use—one that addresses privacy concerns, decreases the risk of data breaches, and promotes open communication with and feedback from parents, experts agree.

While no one can forecast the future, it’s no stretch to say that speech tech is here to stay in the halls of academia as well as the home learning environment.

“To ensure that students have access to equal education, we’ll see more schools and universities adopting

text-to-speech technologies. When offered in e-learning content, especially assessments, text-to-speech allows for users who may have learning disabilities to remain in the classroom and participate with the assessment as all others do, ensuring that they feel included in the classroom,” Muldoon says.

Also, he adds, “as universities prepare to offer more online courses, they’ll need to ensure that they’re offering the latest technologies that give students, regardless of learning level, the best learning experience possible.”

Fred Singer, founder and CEO of Echo360, believes ASR will become an essential technology in the classroom for note taking and search. Students can, for example, link an analog lecture to digital textbooks on a range of devices; Singer

believes this will change the way students take notes and how they interact in class. “They won’t just get a recording of the class but all the keywords they need for sophisticated search,” he argues.

Hasbun is banking on speech technology continuing to increase accessibility for more students.

“Not just the disabled, but also people who speak other languages, or even just people whose eyes get tired staring at screens all day,” Hasbun says. “And I don’t think it will be a hard sell.”

But, for all that it can do in the classroom, speech technology will never replace the need for a flesh-and-blood teacher, Ali is convinced. “Instead, we’ll empower teachers to perform even better,” he says. ☒

Erik J. Martin is a Chicago area-based freelance writer and public relations expert whose articles have been featured in AARP The Magazine, Reader’s Digest, The Costco Connection, and other publications.

He often writes on topics related to real estate, business, technology, healthcare, insurance, and entertainment. He also publishes several blogs, including martin-spiration.com and cineversegroup.com.

2020 ANNUAL REFERENCE GUIDE

Your guide to today's trends and opportunities related to speech technology solutions and services.



Speech
TECHNOLOGY

<http://speechechmag.com/referenceguide>



As Speech Translation Advances, Barriers Continue to Fall

TECHNOLOGY IMPROVEMENTS HAVE PUSHED SPEECH TRANSLATION OUT TO A SMALLER WORLD • By Phillip Britt

Speech translation was already a rapidly growing market, thanks to evolving technologies, growing international travel, and business globalization. The COVID-19 pandemic slowed international travelers' need for translation capabilities, but that need is expected to rebound as the world re-emerges from the pandemic.

Market research firm Mordor Intelligence forecasts that the speech-to-speech translation market will grow at a compound annual rate of 9.4 percent between now and 2025 thanks to further globalization, which it says will pick up again.

Speech-to-speech translation helps break down language barriers, helping to foster international business and tourism and to equalize communication ability, the Mordor forecast says.

Since the growth of the tourism industry is fueling the demand for speech-to-speech translation solutions, vendors such as Raytheon, Lingmo, and LogBar are actively positioning their products along those lines.

Mordor Intelligence and others also point out that the technology behind speech-based translation has improved immensely in just the past year or two, sparking further demand and interest.

The most prolific advancements have been in the areas of artificial intelligence, neural processing, natural language understanding, speech recognition, and text-to-speech conversion, making speech translation much better than it was just a few years ago, says Sue Reager, chief technology officer at Language Preservation Technologies. "Translation has moved from good to wonderful. Translation is strong at a very high level."

Greater computing power and the ability to house it on smaller devices have also had a dramatic impact on speech translation technology.

Additionally, the technology has benefited from better audio quality—thanks largely to noise reduction and echo cancellation—that is possible on phones, headsets, intelligent virtual assistants, and other devices involved in translation, according to Reager.

"If the audio was crap, the speech translation was crap," Reager says, noting that those issues have mostly been resolved. "Now you have great technology and great translation."

Reager says speech translation has evolved from "an insane focus on English" a few years ago to excellent translation between most major languages. Mordor's research, for example, found a rising demand for Mandarin Chinese, which is not surprising given China's rising position in the world, but translation of less common languages is also possible today.

Reager's firm is working on translations to and from Cherokee and other Native American languages.

The ability to speed up and slow down outputs is also making translations better for use with animation. Reager notes that the popular children's television show *Sesame Street* is now available in Cherokee. The translation, she says, appears to be natural because the sound and puppet movements are in synch, which wasn't possible earlier.

Another major advancement is the ability to add emotion and colloquialisms in speech translation, Reager says, maintaining that these advances have taken speech translation far beyond the robotic-sounding monotony that plagued it in the past.

Several Market Participants

It's not just firms like Reager's that are leading the innovations in speech translation technology. Some of the top technology companies have recently advanced their language translation capabilities as well.

Google, for example, earlier this year launched Transcribe for Android, a feature of its Translate app that can convert live speech in one language into another language almost instantly. Google Transcribe rolled out with support for English, French, German, Hindi, Portuguese, Russian, Spanish, and Thai.

Reager also credits Google with offering translations in some less common languages where there is no return on investment. "That is heroism of the first magnitude," she says.

Fellow tech giant Amazon, meanwhile, has started offering several translation services through its AWS Lambda function, which connects with Amazon's three artificial intelligence-based language services, Amazon



Transcribe, Amazon Translate, and Amazon Polly. Moreover, Amazon Alexa can now provide voice translation from English to as many as 48 other languages.

“Machine translation works quite well with major language pairs,” says Evgeny Matusov, lead science architect for machine translation at AppTek, a provider of artificial intelligence, speech recognition, neural machine translation, and natural language understanding technologies.

AppTek in early July launched two consumer speech translation applications, AppTek Speech Translate and AppTek Speech Transcribe. With both applications, users can communicate in real time with other people in 17 languages (English, Arabic, Portuguese, Mandarin Chinese, Dutch, French, German, Greek, Hebrew, Italian, Japanese, Korean, Pashto, Persian, Russian, Spanish, and Turkish). The applications also support a variety of dialects, including 12 Arabic, two French, three Spanish, five English, two German, and two Portuguese dialects.

AppTek Speech Translate is a two-way speech communication application that offers conversational real-time streaming speech-to-speech translation using the latest in artificial intelligence and machine learning. With it, users speak directly into their mobile devices to transcribe and simultaneously translate spoken content in AppTek’s 17 languages; the app delivers the translation as a natural-sounding spoken output and even provides access to 100 preloaded offline phrases to help in emergency, medical, transportation, and dining situations.

AppTek Speech Transcribe can convert voice memos, lectures, meetings, interviews, and other spoken audio content into text in real time across AppTek’s 17 languages. It can also be used as an accessibility aid when speaking with deaf or hard-of-hearing individuals.

“These simple-to-use and highly accurate applications enable users to speak fluently in multiple languages across a wide range of domains, including travel,

healthcare, conversational, and more, greatly enhancing communications in a breadth of situations,” said AppTek CEO Mudar Yaghi in a statement at the time of the release.

Another tech leader, TransPerfect, is expanding the number of languages and dialects that its GlobalLink platform for transcription and subtitling supports through a partnership with AppTek.

Through the partnership, GlobalLink will be able to produce transcripts and subtitles among the 30 languages and dialects that AppTek currently supports. In addition, AppTek’s patented deep neural network-based intelligent line segmentation solution allows for the automated

“Machine translation works quite well with major language pairs.”

conversion of transcripts to subtitle files, with text broken up based on both syntax and semantics to reduce the amount of human effort needed to create subtitles.

More Important in Call Centers

While machine translation has many uses, it has become more important than ever in call centers since the onset of the COVID-19 pandemic, according to Jen Snell, vice president of product marketing for Verint Intelligent Self-Service. Call volume all over the world rose dramatically once government-mandated lockdowns began, particularly as people ramped up their online purchases and digital transactions. “The pandemic has provided us an opportunity to build our knowledge rapidly,” Snell relates.

Verint had already built its own speech recognition engine capable of supporting 70 languages for contact centers around the world. The engine offers flexibility with conversational AI, voicebot integrations, multiple speech recognizers, and semantic processors, as well as other Verint offerings.

In the contact center, Verint’s ASR helps improve operational efficiency with conversational AI to understand what callers are saying so that the system can provide the right self-service options or transfer calls to the right agents.

“We can offer faster out-of-the-box support,” Snell says of Verint’s machine translation capabilities. However, like other experts, she notes that machines can’t handle all translation tasks by themselves, particularly when the caller has a very heavy accent.

As translation technologies improve, Reager sees them moving into other verticals where they previously didn’t work well enough. One of the main ones is education.

Reager predicts that speech translation will become more embedded in the classroom to the point where students can take classes from several universities in different countries without language concerns. Online lectures, books, and other course materials could all be translated into several languages to enable students to assemble an education from anywhere.

Speech translation systems are also being used throughout the world in medical facilities, hotels, retail stores, factories, and police departments and are basically applicable anywhere that spoken language is used to communicate.

Speech translation can be used in so many places and for so many purposes now largely because of recent advancements in computing power that have enabled it to move to more devices and platforms, Reager and others contend.

Far from Perfect

However, despite all of the advancements, how much companies can rely strictly on machine translation and how much they need to augment machine capabilities with human interpreters depends on how precise the translation needs to be and the type of information being translated, according to AppTek’s Matusov.

“If you need perfect quality, there will be some editing involved,” he says.



The more technical or nuanced the information, the more a human will need to be involved in producing the final translation. Human translators likely will not be needed for mundane, simple work, but that is not likely to be the case for more complex materials, according to Matusov.

Another challenge with speech translation is the need to standardize interfaces and data formats to ensure that systems are compatible. Research in this area is being fostered by speech translation consortiums, such as C-STAR and A-STAR, which are working to design standards for bilingual corpora, interfaces, and data formats to connect speech translation modules internationally.

Other ongoing challenges include speaker-dependent variations in speaking styles and pronunciations and external factors, such as acoustic noise or speech by other speakers, in real-world situations. Current systems continue to work best in quiet environments with one person speaking at a time, which is not how most conversations occur.

For the best translations, context is also essential, experts agree. Not only will it help with gender-specific translations, but also with formal and informal translations, Matusov says.

“You need to know how to manipulate really good neural translation software to get the right result,” Reager points out.

Reager also notes that many different types of translation engines exist for different types of uses. Google, she said, is among the best in many cases, and many experts agree.

According to a study by online translations platform One Hour Translation, Google Assistant is one of the top-performing real-time voice translators, ahead of Apple’s Siri, Amazon’s Alexa, and Microsoft’s Cortana and Skype Translate.

One newcomer in the speech translation space, though, threatens Google’s leadership position, many experts content. That company, DeepL, is a

German firm that launched its Translator free machine translation service in August 2017. DeepL’s technology converts words between English and 10 other languages (Chinese, Dutch, French, German, Italian, Japanese, Polish, Portuguese, Russian, and Spanish) and proposes approximations of language equivalence among them all using a two-step process via an English pivot. That, many experts believe, could ultimately make DeepL more accurate and nuanced than Google Translate.

Reager says DeepL offers the best translation for Chinese speakers in the

“You need to know
how to manipulate
really good neural
translation software
to get the right result.”

United States, but for Chinese speakers in China, Sogou is a better option.

“Each speech translation engine has its own strengths and weaknesses,” Reager advises. “You should use the one that offers you the best return.”

When it comes to speech-based translation, many difficulties have already been overcome, but many others still remain. Most of them are related to the languages themselves. For one, some languages use different tenses and gender pronouns depending on the context, and translations have to take them into account.

The issue of gender misidentification is perhaps the hardest to address, and it goes beyond basic translation issues. To deal with it, researchers at the University of Trento in Italy in June proposed a benchmark—which it dubbed MuST-SHE—to evaluate whether speech translation systems fed textual data are constrained by the fact that sentences sometimes omit gender identity clues. The co-authors assert that these systems often exhibit gender bias, and that signals

beyond text (like audio) provide contextual clues that might reduce this bias.

In machine translation, gender bias is at least partially attributable to the differences in how languages refer to males and females. Romance languages, for example, incorporate gender agreement into their grammatical constructs, with verb endings and other sentence constructs dependent on the gender of the nouns involved.

The MuST-SHE standard is a multilingual test set designed to uncover gender bias in speech translation. It comprises roughly 1,000 audio recordings, transcripts, and translations in English-to-French and English-to-Italian pairs from the open source MuST-C corpus, annotated with qualitatively differentiated and balanced gender-related phenomena. MuST-C is the largest multilingual corpus available for the direct speech translation task.

“If, like human beings, ‘machine learning is what it eats,’ the different diet of machine translation and speech translation models can help them develop different skills,” writes the researchers. “By eating audio-text pairs, speech translation has a potential advantage: the possibility to infer speakers’ gender from input audio signals.”

The report comes after Google introduced gender-specific translations in Google Translate chiefly to address gender bias. Scientists have proposed a range of approaches to mitigate and measure it, most recently with a leaderboard challenge and set of metrics dubbed StereoSet.

There are other issues to be further resolved in speech translation, but experts agree that problems with mistranslations are becoming fewer and fewer as the underlying technologies continue to improve. And once that occurs, people will be able to resume international travel without any of the language and cultural barriers. ☒

Phillip Britt is a freelance writer based in the Chicago area. He can be reached at spenterprises@wowway.com.

Companies Turn to IVAs to Deal with COVID Surges

Automated voice solutions have helped contact centers handle rising interaction volume during the pandemic

BY LEONARD KLIE

As a response to the global COVID-19 pandemic, companies all over the world had to quickly pivot their contact center employees to work-from-home environments. On top of that, the pandemic led to dramatic increases in contact center volume, with the number of inbound conversations more than doubling for many support teams. Additionally, contact center agents had to field questions from customers about subjects they've previously never had to handle.

Companies didn't have to handle the pandemic alone, though. Conversational artificial intelligence (AI) in the form of intelligent virtual assistants (IVAs) became the go-to solution for companies to quickly and accurately provide information about how their operations were being affected by COVID-19.

The pandemic accelerated investments in digital touchpoints, including IVAs, as companies looked for ways to combat the deluge of inbound calls and unpredictable call volume, Forrester Research concludes in a recent report.

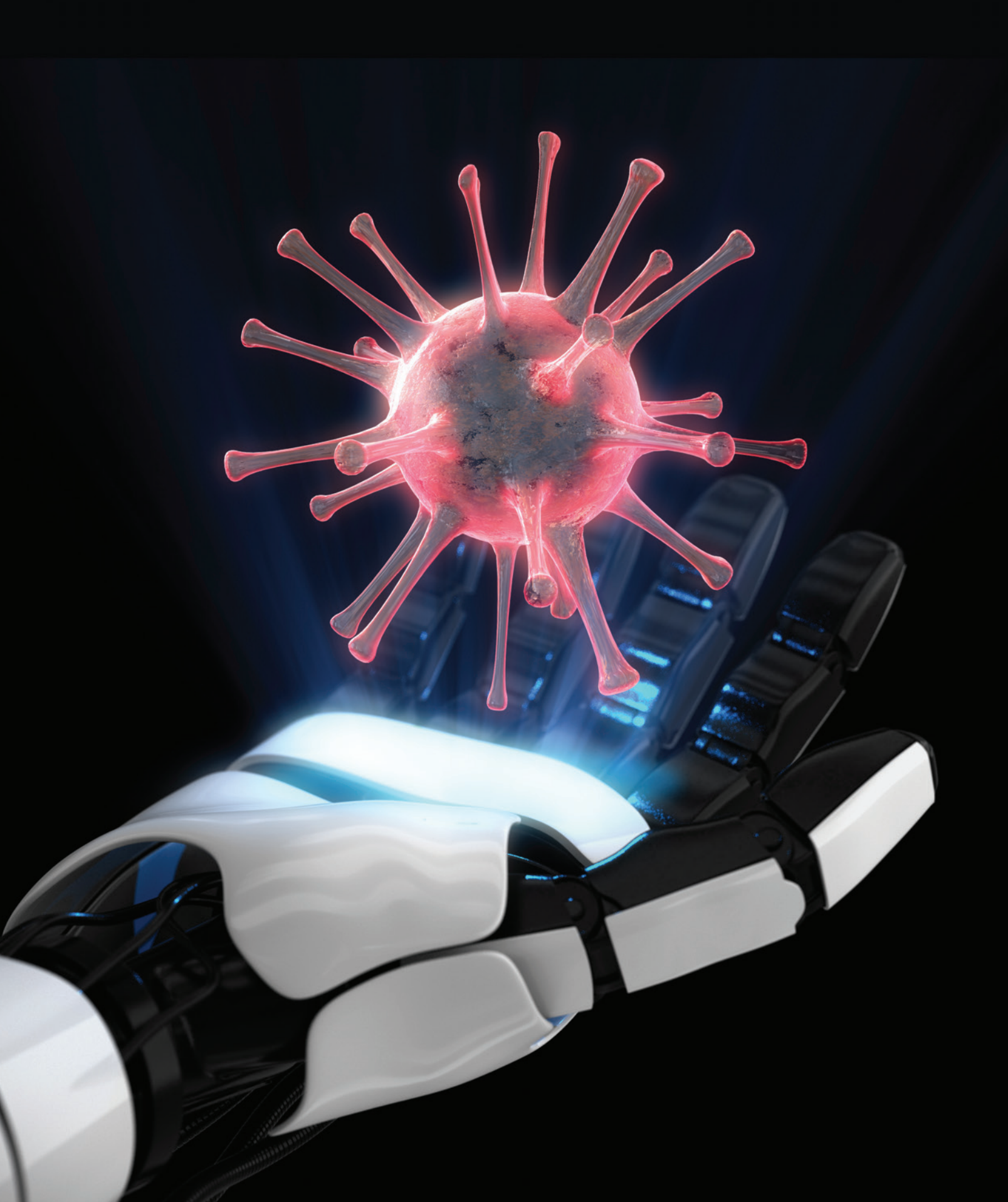
Others noticed the trend as well.

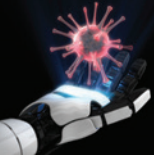
"We've seen how enterprises have quickly updated their current IVAs to handle pandemic-related inquiries. They've also put AI to work within their organizations to support employees adjusting to a new work environment and minimize the effort needed for agents to find the right information," says Jen Snell, vice president of product marketing for Verint Intelligent Self-Service.

"Too many businesses assume that operational efficiency cannot be achieved during a crisis. Yet, as we've seen in the past few months, when IVAs are answering thousands of queries at once, a contact center can scale both its response and its human touch. IVAs can prioritize which questions need the attention and empathy of a human agent, making the most of agents' time and talents. IVAs can also evolve to meet new needs, all without requiring a costly, complex, all-hands-on-deck sprint to make changes to critical software," Snell continues.

DMG Consulting, in a similar report, notes that IVAs serving on the front lines of self-service during the coronavirus outbreak eliminated any doubt about their efficacy.

"The pandemic has highlighted the need for organizations to develop and implement reliable disaster recovery (DR)/business continuity (BC) plans as well as comprehensive self-service solutions. In the world of customer service, IVAs are becoming mission-critical tools as organizations strive to maintain their service levels while responding to today's unprecedented service demands," the report concludes.





And because of that, nearly two-thirds of U.S. IT decision makers plan to increase their investments in automation in the coming year, noting that intelligent self-service has been key to remaining agile during the crisis, according to research by Opinion Matters and Inference Solutions.

A full 71 percent of IT decision makers agree that intelligent self-service automation helped their organizations remain agile, and 64 percent expect to increase automation investments over the coming year as a result of the crisis. More than a quarter (26 percent) will increase their investments by 10 percent or more.

Additionally, 69 percent expect customer service automation to play a significant role in their organizations' resilience over the next two years,

"Cloud-based, self-managed platforms that give businesses more control of their customer-and employee-facing self-service applications will be key to overcoming the barriers to adoption flagged in our research," said Callan Schebella, Inference Solutions' CEO, in a statement. "We are pleased to see that IT leaders plan to use automation not only to lower operating costs, but to better support their employees as they navigate the challenges of doing business in a post-pandemic world."

Industries Hardest Hit

The pandemic certainly disrupted business across just about every industry, but perhaps none were as hard hit as the travel and transportation industries, which saw event cancellations, closed borders, and other restrictions imposed almost overnight. Within a matter of days, travelers all over the world had to cancel or rebook their trips.

Companies in these two industries quickly updated their voice systems with information specific to the virus and the related travel restrictions. Their IVAs were reconfigured to direct customers to landing pages containing a wealth of information about service impacts, cancellation policies, and other essential information.

Corporate travel services provider

Gant Travel, for example, optimized its systems in this way. Working alongside NICE inContact's team of experts and partners, the Bloomington, Indiana-based organization eliminated silos at its contact center so agents could have the unified, omnichannel capabilities they needed to rapidly respond to customers in real time.

Since implementing NICE inContact CXOne, Gant has achieved the following outcomes:

- reduced telephone costs by 7 percent;
- reduced average handling time by 5 percent;
- reduced queue wait times by 46 percent by eliminating extraneous and redundant IVR functionality; and
- increased quality scores for voice and email to 93 percent.

"Being in the business of facilitating

"Cloud-based, self-managed platforms that give businesses more control of their self-service applications will be key to overcoming barriers to adoption."

travel, our success hinges on our ability to ensure that client trips are seamless," said Jim Conner, director of operations at Gant Travel, in a statement. "When an issue arises, every second we're not spending getting things back on track is a missed opportunity, which is why ensuring an optimally efficient contact center is so critical. NICE inContact helped us solve for our major problem areas, improving both the agent and customer experiences."

ZIM, a global container shipping industry powerhouse, saw a huge surge in customer contacts from shippers who were nervous about vulnerable supply chains, changing cargo routes, border shutdowns, and similar concerns.

ZIM worked with Orange Business Systems to deploy its Managed Contact Center (MCC), which employs automation and AI to route callers to the most

appropriate agents based on their latest profiles and interactions.

In-depth data analytics and reporting provided by the solution gives ZIM complete visibility into contact center activities and customer communications. This functionality allows the company to mine data and analyze customer needs. In addition, a dashboard and real-time monitoring system gives ZIM a 360-degree view of call center queues on a dynamic digital whiteboard located in each of ZIM's call centers. Thanks to such tools, ZIM can take advantage of an optimized response time and have visibility on relevant data for answering the calls.

"We succeeded in maintaining our high level of customer experience across our global branches, despite the COVID-19 pandemic implications, necessitating many of our customer service teams to work from home with very short notice. This was in large part thanks to Orange Business Services' technology and capabilities," explains Assaf Tiran, ZIM's global customer service vice president.

Other Industries Challenged

Not surprisingly, healthcare companies were also stretched like never before when the pandemic hit.

AccuRx, which provides a platform used by doctors to communicate with patients, experienced a massive surge in incoming customer support questions early in the outbreak. The number of queries jumped 10 times the normal amount almost immediately. Its customer support team had to quickly add automation and artificial intelligence to help meet the increased demand without further taxing its already-limited personnel.

With the help of Intercom, AccuRX scaled support with chatbots to instantly resolve new inquiries while collecting critical information that it could pass on to support representatives when needed.

AccuRX set up quite a few auto-responses using Intercom's Resolution Bot and deployed Intercom's Custom Bots to gather data about customers so agents could understand them and their issues before chatting.

Integrating Intelligent Assistants for a Competitive Edge

VERINT®

Removing Blind Spots from Intelligent Assistant Integration

Intelligent assistance adoption is on the rise, and we don't see this trend changing any time soon. This isn't only because of COVID-19 and the urgent need to scale to meet intense customer demand—it's also because companies are deploying pre-planned innovation roadmaps. As more companies look to integrate intelligent assistants into their IT environments and business processes, many are encountering new challenges, both technical and cultural, for which they don't have an established playbook.

With that in mind, here are three imperatives when integrating intelligent assistants into your existing processes.

1. REDEFINE AI

The AI hype cycle is unrelenting, and many have come to think of AI as a silver bullet for the enterprise. Yes, AI means *artificial intelligence*, but in practice, businesses would be better served by thinking of this technology as *augmented intelligence*. AI is a tool that helps people make the most of their time and effort. And, like any tool, it needs to be operated effectively and in the right situations by actual people.

By thinking of AI as augmented intelligence instead of artificial intelligence, you force yourself to consider the technology through the lens of your business. What exactly is worth augmenting in your organization? Phrased another way: What could stand to be strengthened through AI, and what is better left to your human workforce?

Asking these questions helps identify where AI can provide the most value. In fact, conversation analysis systems are available to help your business identify intelligent assistant use cases, allowing you to accurately identify automation opportunities and accelerate solutions. Understanding where and how you should deploy an intelligent virtual assistant (IVA) that emulates and augments intelligence is the first step toward ensuring you are effectively integrating this technology into your processes.

2. MOVE BEYOND DEFLECTION

Across industries, stakeholder expectations are converging. And intelligence across the enterprise to engage customers and support employees is required. The experience of doing business with a company that has yet to deploy AI will soon feel like doing business with a company that doesn't have a website. Still, many businesses set antiquated goals for AI deployments that anchor this powerful new technology to the business logic of the late 1990s.

One of the most frequent planning errors is when an organization integrates intelligent assistants into business processes that were designed with the singular goal of call deflection. In doing so, they limit the value they can derive from AI. What good is an artificially intelligent assistant that is designed to avoid interaction with customers? The whole premise of intelligent assistants is to provide scaled, personalized engagement. Yet too many businesses integrate them into processes that have for decades been built to cut costs, mitigate risk, and minimize contact.



As a tool, intelligent assistants not only help you address the symptom of high chat volume, but also help diagnose the root cause of customer challenges, and in turn create high-value solutions that go beyond simple cost-cutting goals. For example, one Verint® customer saw increased revenue generation from cross-selling and up-selling when our intelligent assistant was implemented. The reason was simple: intelligent assistants never forget to cross-sell and up-sell when talking to a customer, but humans do. That revenue would never have been realized if the AI was integrated into a process designed to achieve traditional contact center KPIs.

3. UPSKILL FOR UPSIDE

Intelligent assistants are a force multiplier, not a replacement, for your workforce. As you examine your existing processes to determine where an intelligent assistant could provide the most value, it's all too easy to overlook the opportunity to focus your people on higher order, higher value work. What are your employees empowered to do, if they aren't bogged down by high-volume, repetitive tasks?

AI works best when it is integrated not just into technical processes, but into real human workflows, as well. Human oversight plays an essential role in ensuring that a business gets what it needs out of an intelligent assistant. In fact, it's people who bring it all together, because your people remain your greatest asset. AI should bring out the best in your employees because it takes on the work that underutilizes them, allowing them to grow their strengths.

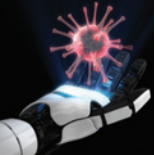
Verint has years of experience delivering innovative, AI-powered cloud solutions to simplify, modernize, and automate customer engagement. Contact us today to learn more. ■

1-800-4VERINT

www.verint.com/engagement

www.verintblog.com

© 2020 Verint Systems Inc. All Rights Reserved Worldwide.



In the telecommunications arena, Canadian wireless services provider Koodo Mobile faced the same challenges. It turned to Ada's automation-first, artificial intelligence-powered customer service platform to meet the demand.

Koodo Support, powered by Ada, saved customers thousands of hours of wait time by automating the customer service experience, freeing up agents to handle the most pressing customer inquiries. The technology has enabled Koodo's customers to manage 60 percent of their queries on their own through self-service without ever needing to speak to live agents.

Another industry that suffered greatly during the pandemic was the nonprofit sector. The Trevor Project, a 24/7 suicide prevention and crisis intervention organization for gay, lesbian, bisexual, and transgender youth, saw its call volume more than double. With a record number of calls, chats, and texts coming into its Lifeline program and counselors temporarily forced to work remotely, the Trevor Project needed to enable its staff to quickly transition to a virtual workplace without sacrificing the quality of connections to its clients or security to ensure their privacy.

The Trevor Project leveraged Vonage technology to transition to a full-time remote and 100 percent virtual contact center during the COVID-19 crisis within days. Vonage also enabled the Trevor Project to build custom automated tools that could surface critical information, connecting those in need of support to the appropriate crisis counselors.

Since switching to Vonage Contact Center, the Trevor Project has experienced close to 100 percent uptime and 24/7 on-call support.

Beyond the Contact Center

Customer-facing teams aren't the only ones using IVA technology to help streamline operations. Organizations are also using it to help other employees who found themselves working from home without the in-person guidance

and regular feedback that they would receive in office settings.

"We've seen businesses successfully deploy IVAs as an internal resource to support employees," Snell recalls. "Employee-facing IVAs have helped offload lower-level questions and provided access to knowledge in real time without relying on a manager or HR resource."

When used internally, IVAs can also provide employees assistance in setting up technology and accessing accounts and solve common challenges like videoconferencing and other IT needs.

The Inference/Opinion Matters research found that 54 percent of IT managers expect employee service automation to play a significant role in

"Employee-facing IVAs have helped offload lower-level questions and provided access to knowledge in real time without relying on a manager or HR resource."

their organizations' resilience; 39 percent expect supply chain automation; and 37 percent flagged robotic process automation. Increasing employee efficiency and productivity topped the list of priorities in respondents' automation strategies, followed by reducing costs and improving the ability to meet customer demands.

Better Than IVRs

Among the reasons for the increased interest in IVAs, DMG Consulting concludes in its report, is that IVAs have many advantages over what it calls "outdated and inflexible" interactive voice response (IVR) systems. IVAs, it says, can be programmed and implemented rapidly. Additionally, IVAs use artificial intelligence to respond to customer inquiries with concierge-style service;

they better support natural language conversations; and when they need to escalate a transaction to a live agent, they can pass on the transaction with context and stay on the line, using machine learning to listen and learn so that they can handle similar transactions more effectively on their own.

IVAs can improve internal business processes as they assist employees by automating tasks, increasing productivity, and improving accuracy.

As an added benefit, conversational intelligence can offer a real-time analysis of customer intents, queries, and even speech patterns during IVA interactions. This, many experts agree, allows businesses to marshal their resources to the greatest areas of customer need without engaging in huge, overly complicated data science projects.

"It's unfortunate that [the pandemic] had to be the catalyst for organizations to change, but for many, if they don't right now, they're going to have a tough time surviving and thriving, because their customers will go elsewhere," said Ben Smith, global vice president and general manager of customer experience solutions at Verint, during a session at Verint's Engage virtual conference in early June.

"While we hope the world doesn't see another crisis on the scale of this pandemic, most contact centers will be much better prepared next time, with an IVA at their side, to weather the storm," Snell maintains. "Most experts suggest we're still only in the early innings of this pandemic and that we'll go through several new phases of operation until we reach a new normal. This is precisely the environment that conversational intelligence helps businesses navigate."

DMG Consulting further expects that once the post-COVID-19 recovery process is under way and the new normal takes hold, adoption of IVAs will increase even further. ☒

Leonard Klie is the editor of Speech Technology. He can be reached at lklie@infoday.com.



Amid the Pandemic, Speech Technologies Expand the Telehealth Options

The telehealth approach is now providing treatments across the healthcare spectrum

Telehealth has rapidly grown over the past decade for both cost-management and convenience reasons, but the recent pandemic has injected a more critical motivation for providers and patients alike.

Prior to 2020, many healthcare businesses separated telehealth from their traditional care lines. But with the ongoing importance of physical distancing, providers are now using telehealth in their everyday practices.

In fact, the telehealth approach is now providing treatments across the entire healthcare spectrum. One example is VoiceSense's predictive speech analytics solutions, which can assist behavioral health professionals with monitoring depression and other mental health issues. VoiceSense is language-independent, tracking speech patterns over time and detecting when the patterns change, thereby alerting a mental health provider when the technology deems the patient's mental state is changing. For many patients, the alerts provide immediate and early feedback that can assist

with behavioral self-care, thereby empowering patients to improve their health.

Much like biometric trackers, VoiceSense's remote monitoring gives healthcare providers insights into patients' well-being through real-time analysis of voice patterns. These mental biomarkers act just like heart rate and blood pressure monitors to deliver early alerts that assist providers with intervention

Expanding telehealth with at-risk populations is critical to reopening the economy.

and more efficient in-person and telehealth visits.

Expanding telehealth with at-risk populations is critical to reopening the economy while protecting the elderly and people with preexisting conditions. LifePod offers a solution that enables both healthcare providers and families to remotely monitor and support their patients and loved ones.

LifePod allows health providers to create routines for patients, such as exercise and medication reminders, including refills. Patients respond to LifePod's questions

and reminders to verify that they are maintaining their healthcare regimen. Personal routines can include family members checking whether the patient requires groceries to be delivered, thereby eliminating the need for a trip to the grocery store.

LifePod's speech interface resembles that of Amazon's Alexa, Apple's Siri, and Google Assistant, so it is user-friendly and easy to learn even for technology-challenged people. Furthermore, it integrates with smart home devices so patients do not have to learn to use another personal assistant to turn on and off lights, change the thermostat, or lock doors. LifePod's combination of healthcare help and family inclusion in the daily lives of patients can prevent having to move independent but at-risk people to higher-risk care facilities while reducing personal trips to stores, pharmacies, and the like.

Finally, even "good old-fashioned" telehealth is quickly morphing due to the use of natural language processing for a greater number of patients across a broader treatment continuum. Many healthcare organizations discovered the value of applying speech analytics to a much larger patient base due to the explosive growth of telehealth during the COVID-19 pandemic. Prior to the pandemic, many if not most healthcare organizations used telehealth for specific cases such as ICU reviews by specialists, or distance care for small patient populations. So speech analytics stayed on the periphery. Now that many traditional office visits moved to telehealth, patient and provider conversations are recorded, and the power of speech analytics is being tested in care areas not previously considered for speech analytics.

COVID-19 quickly and by necessity changed the delivery of healthcare, and speech technologies are supporting these changes that are likely permanent. Healthcare will continue to drive new and important uses of speech technologies in the near future. ☒

Kevin Brown is customer experience application architect at Banner Health, where he specializes in voice and web customer experience solutions. He has more than 25 years of experience designing and delivering speech-enabled solutions. You can reach him at kevin.brown@voxpertus.com



FORWARD THINKING

By Moshe Yudkowsky

How Can Speech Technology Help?

Looking at all the ways speech recognition can step up amid the pandemic

COVID-19, unlike most other epidemics in our experience, has a place for amateurs and non-physicians to help. Remarkable improvements in technology made this possible.

Here's one piece of technology that has moved in the space of about three years from obscure and expensive to commonplace and relatively cheap. Many of my friends own 3-D printers, expert users who design and print parts for their various projects, both new designs and repairs. When the Great Mask Shortage began in March 2020, they stepped up and began churning out masks.

Then came a wave of higher-tech applications, but before I get there, I want to thank the lower-tech and personal contributions by the denizens of my Chicago neighborhood. A charitable organization found funding to distribute thousands of meals each week to people suddenly thrown out of work. The local volunteer fast-response paramedics found themselves with their hands full, and citywide, Chicago's Community Emergency Response Team provided 311 staffing to answer COVID-19-related calls. These

people donated time and sweat and money to help their neighbors, and we all owe them, and people like them across the world, a debt of gratitude.

The main issue that came to my mind about speech technology and the pandemic is related to medicine.

The high-tech projects of COVID-19 included, as one might expect, websites dedicated to various ventures: databases about the epidemic itself; dedicated communication channels between researchers; support sites for various new charitable groups. Other sites supported volunteer and commercial ventures that seem like obvious projects, but, in my opinion, only in retrospect. For example, I think we can all agree now that any project that offers grandparents (such as myself) a way to more easily read books to distant grandchildren is extremely welcome.

Still other projects have the usual ominous ring. "Contact tracing" via apps on phones is one example—whether voluntary or not, and in many places, not. Here in the United States, I'm not entirely sanguine about signing up Big Data companies to pore through their records to track tens of millions of citizens.

And then we come to speech technology. What can speech technology do to help?

If you've ever been to a call center, or for that matter any cubicle farm, you've no doubt noticed it's not set up

for social distancing ("social distancing" is a misnomer, by the way—it's asocial distancing). This means workers will work from home; this means a great deal of work for system administrators. While in theory this means the flexibility of text-to-speech (TTS) and automatic speech recognition (ASR) could help, in practice I've yet to run across a business using TTS for calls, or one that made a significant switch to ASR to listen to callers. The reason is quite simple: If your underlying call center technology does not currently support a remote-based workforce, ASR and TTS will not make this more fundamental problem vanish.

But the main issue that came to my mind about speech technology and the pandemic is related to medicine. I've seen any number of web pages that let you determine, based on your symptoms, if you should be tested for COVID-19. I know there are chatbots (experimental ones, as far as I can tell) to analyze symptoms, and I fully expect to see dedicated kiosks, complete with ASR, that will take over the tedious task of screening visitors; expect some to be deployed later this year to at least some businesses. I'll guess that airports will get them eventually as well. Having answered the COVID-19 screening questions any number of times, what I'd really like at this point is an ASR application for my phone that hears the questions from the kiosk and answers on my behalf. No healthcare system I've used over the past two months offers automated pre-visit screening over the phone.

But what really intrigues me is the possibility of phone-based applications that screen for COVID-19 based on biometrics. Set aside smart watches that monitor your temperature, heart rate, or even your EKG. What about a smartphone application that analyzes speech to determine if you're ill with COVID-19? Research groups have conjectured that COVID-19, which inflames the respiratory system, would produce measurable changes in speech, changes these groups say they've detected (see the related story on page 9).

Perhaps this work will pan out and perhaps not. Maybe vaccines will arrive before we know for sure. Regardless, even if this work does not help this time around, it lays the foundation for future work. Perhaps one day a rapid test for a new respiratory pathogen can be deployed, not by shipping throat swabs, but deploying phone-based applications—a worthy goal, one everyone in the speech technology field can be proud of. ☒

Moshe Yudkowsky, Ph.D., is the president of Disaggregate Consulting and author of *The Pebble and the Avalanche: How Taking Things Apart Creates Revolutions*. He can be reached at speech@pobox.com.

Speech Technology Research Center

Leverage RPA to Help Modernize Customer Engagement

SPONSORED BY: Verint

As consumer expectations for faster service increase, organizations need to look for additional ways they can help their contact center agents interact with customers more efficiently, delivering the right response the first time. RPA in its many forms can help organizations achieve these goals.

The logo for Verint, featuring the word "VERINT" in a bold, blue, sans-serif font with a registered trademark symbol.

Learn How to Attain Secure Customer Authentication

SPONSORED BY: LumenVox

There's no fraud strategy more foolproof than multifactor authentication. One multifactor modality that strengthens the authentication process dramatically is passive voice biometrics. Learn more about this key strategy to help meet increasing demand.

The logo for LumenVox, featuring a stylized icon of three vertical bars of varying heights in green and orange, followed by the text "LumenVox" in a bold, sans-serif font and "Speech + Authentication" in a smaller font below it.

What to Look For in a Voice AI Partner

SPONSORED BY: Observe.AI

By learning from these use cases, benefits, and best practices, you can transform your quality assurance program. Here's what to look for in a Voice AI partner.

The logo for Observe.AI, featuring a stylized icon of three vertical bars of varying heights in black and white, followed by the text "OBSERVE.AI" in a bold, sans-serif font.

Staying Connected With Your Customers

SPONSORED BY: Concentrix

Research conducted by Market Strategies International has found that consumers are looking for a company to meet them where they already are on mobile, social, and the web; proactively reach out to them with relevant products and services; and make it easy to do business with you. Learn how business messaging can help!

The logo for Concentrix, featuring a stylized icon of three curved lines in blue, green, and orange, followed by the text "CONCENTRIX" in a bold, sans-serif font.

A 5-Step Process for Improving Home-Based Agent Performance

SPONSORED BY: NICE Nexidia

The transition of contact center employees from brick-and-mortar to home-based offices is a reality that may become a long-term work arrangement. As such, it is an opportunity to redefine quality management and coaching for improved communication and collaboration within any work environment. Read this white paper to learn how a systematic five-step process can drive managed and sustained agent performance improvements.

The logo for NICE Nexidia, featuring the word "NICE" in a bold, blue, sans-serif font, followed by a small blue square, and then the word "nexidia" in a lowercase, blue, sans-serif font.

Download these white papers and more today!
www.speechtechmag.com/IndustryWhitepapers

Speech
TECHNOLOGY
magazine

DELIVERED TO YOUR INBOX EACH WEDNESDAY,
Speech Technology eWeekly provides comprehensive updates on what's
happening in speech technology and the world of artificial intelligence.

Featuring accurate news, in-depth feature articles, case studies, Industry Voices,
columns from luminaries in the field, and more, it's a newsletter you can look forward
to, as it delivers fresh and accurate news from across the industry.

spechtechmag.com/newsletters

THIS WEEK'S SPEECH INDUSTRY NEWS



SUBSCRIBE

Speech
TECHNOLOGY