

How Industries Are Meeting Consumer Expectations With Speech AI

Artificial intelligence is transforming customer experiences and helping companies provide on-demand service, fast delivery and personalized interactions.

Author: Cliff Edwards

Thanks to rapid technological advances, consumers have become accustomed to an unprecedented level of convenience and efficiency.

Smartphones make it easier than ever to search for a product and have it delivered right to the front door. Video chat technology lets friends and family on different continents connect with ease. With voice command tools, AI assistants can play songs, initiate phone calls or recommend the best Italian food in a 10-mile radius. AI algorithms can even predict which show users may want to watch next or suggest an article they may want to read before making a purchase.

It's no surprise, then, that customers expect fast and personalized interactions with companies. According to a Salesforce research report, 83% of consumers expect immediate engagement when they contact a company, while 73% expect companies to understand their unique needs and expectations. Nearly 60% of all customers want to avoid customer service altogether, preferring to resolve issues with self-service features.

Meeting such high consumer expectations places a massive burden on companies in every industry, including on their staff and technological needs — but speech AI can help.

Speech AI can understand and converse in natural language, creating opportunities for seamless, multilingual customer interactions while supplementing employee capabilities. It can power self-serve banking in the financial services industry, enable food kiosk avatars in restaurants, transcribe clinical notes in healthcare facilities or streamline bill payments for utility companies — helping businesses across industries deliver personalized customer experiences.

Most people now use both digital and traditional channels to access banking services, creating a demand for omnichannel, personalized customer support. However, higher demand for support coupled with a high agent churn rate has left many financial institutions struggling to keep up with the service and support needs of their customers.

Common consumer frustrations include difficulty with complex digital processes, a lack of helpful and readily available information, insufficient self-service options, long call wait times and communication difficulties with support agents.

According to a recent NVIDIA survey, the top AI use cases for financial service institutions are natural language processing (NLP) and large language models (LLMs). These models automate customer service interactions and process large bodies of unstructured financial data to provide AI-driven insights that support all lines of business across financial institutions — from risk management and fraud detection to algorithmic trading and customer service.

By providing speech-equipped self-service options and supporting customer service agents with AI-powered virtual assistants, banks can improve customer experiences while controlling costs. AI voice assistants can be trained on finance-specific vocabulary and rephrasing techniques to confirm understanding of a user's request before offering answers.

Kore.ai, a conversational AI software company, trained its BankAssist solution on 400-plus retail banking use cases for interactive voice response, web, mobile, SMS and social media channels. Customers can use a voice assistant to transfer funds, pay bills, report lost cards, dispute charges,

reset passwords and more.

Kore.ai's agent voice assistant has also helps live agents provide personalized suggestions so they can resolve issues faster. The solution has been shown to improve live agent efficiency by cutting customer handling time by 40% with a return on investment of \$2.30 per voice session .

With such trends, expect financial institutions to accelerate the deployment of speech AI to streamline customer support and reduce wait times, offer more self-service options, transcribe calls to speed loan processing and automate compliance, extract insights from spoken content and boost the overall productivity and speed of operations.

Heavy investments in 5G infrastructure and cut-throat competition to monetize and achieve profitable returns on new networks mean that maintaining customer satisfaction and brand loyalty is paramount in the telco industry.

According to an NVIDIA survey of 400-plus industry professionals, the top AI use cases in the telecom industry involve optimizing network operations and improving customer experiences. Seventy-three percent of respondents reported increased revenue from AI.

By using speech AI technologies to power chatbots, call-routing, self-service features and recommender systems, telcos can enhance and personalize customer engagements.

KT, a South Korean mobile operator with over 22 million users, has built GiGa Genie, an intelligent voice assistant that's been trained to understand and use the Korean language using LLMs. It has already conversed with over 8 million users.

By understanding voice commands, the GiGa Genie AI speaker can support people with tasks like turning on smart TVs or lights, sending text messages or providing real-time traffic updates.

KT has also strengthened its AI-powered Customer Contact Center with transformer-based speech AI models that can independently handle over 100,000 calls per day. A generative AI component of the system autonomously responds to customers with suggested resolutions or transfers them to human agents for more nuanced questions and solutions.

Telecommunications companies are expected to lean into speech AI to build more customer self-service capabilities, optimize network performance and enhance overall customer satisfaction.

The food service industry is expected to reach \$997 billion in sales in 2023 , and its workforce is projected to grow by 500,000 openings. Meanwhile, elevated demand for drive-thru, curbside pickup and home delivery suggests a permanent shift in consumer dining preferences. This shift creates the challenge of hiring, training and retaining staff in an industry with notoriously high turnover rates — all while meeting consumer expectations for fast and fresh service.

Drive-thru order assistants and in-store food kiosks equipped with speech AI can help ease the burden. For example, speech-equipped avatars can help automate the ordering process by offering menu recommendations, suggesting promotions, customizing options or passing food orders directly to the kitchen for preparation.

HuEx , a Toronto-based startup and member of NVIDIA Inception , has designed a multilingual automated order assistant to enhance drive-thru operations. Known as AIDA, the AI assistant receives and responds to orders at the drive-thru speaker box while simultaneously transcribing voice orders into text for food-prep staff.

AIDA understands 300,000-plus product combinations with 90% accuracy , from common requests such as "coffee with milk" to less common requests such as "coffee with butter." It can even understand different accents and dialects to ensure a seamless ordering experience for a diverse population of consumers.

Speech AI streamlines the order process by speeding fulfillment, reducing miscommunication and minimizing customer wait times. Early movers will also begin to use speech AI to extract customer

insights from voice interactions to inform menu options, make upsell recommendations and improve overall operational efficiency while reducing costs.

In the post-pandemic era, the digitization of healthcare is continuing to accelerate. Telemedicine and computer vision support remote patient monitoring, voice-activated clinical systems help patients check in and receive zero-touch care and speech recognition technology supports clinical documentation responsibilities. Per IDC, 36% of survey respondents indicated that they had deployed digital assistants for patient healthcare.

Automated speech recognition and NLP models can now capture, recognize, understand and summarize key details in medical settings. At the Conference for Machine Intelligence in Medical Imaging, NVIDIA researchers showcased a state-of-the-art pretrained architecture with speech-to-text functionality to extract clinical entities from doctor-patient conversations. The model identifies clinical words — including symptoms, medication names, diagnoses and recommended treatments — and automatically updates medical records.

This technology can ease the burden of manual note-taking and has the potential to accelerate insurance and billing processes while also creating consultation recaps for caregivers. Relieved of administrative tasks, physicians can focus on patient care to deliver superior experiences.

Artisight, an AI platform for healthcare, uses speech recognition to power zero-touch check-ins and speech synthesis to notify patients in the waiting room when the doctor is available. Over 1,200 patients per day use Artisight kiosks, which help streamline registration processes, improve patient experiences, eliminate data entry errors with automation and boost staff productivity.

As healthcare moves toward a smart hospital model, expect to see speech AI play a bigger role in supporting medical professionals and powering low-touch experiences for patients. This may include risk factor prediction and diagnosis through clinical note analysis, translation services for multilingual care centers, medical dictation and transcription and automation of other administrative tasks.

Faced with increasing demand for clean energy, high operating costs and a workforce retiring in greater numbers, energy and utility companies are looking for ways to do more with less.

To drive new efficiencies, prepare for the future of energy and meet ever-rising customer expectations, utilities can use speech AI. Voice-based customer service can enable customers to report outages, inquire about billing and receive support on other issues without agent intervention. Speech AI can streamline meter reading, support field technicians with voice notes and voice commands to access work orders and enable utilities to analyze customer preferences with NLP.

Minerva CQ, an AI assistant designed specifically for retail energy use cases, supports customer service agents by transcribing conversations into text in real time. Text is fed into Minerva CQ's AI models, which analyze customer sentiment, intent, propensity and more.

By dynamically listening, the AI assistant populates an agent's screen with dialogue suggestions, behavioral cues, personalized offers and sentiment analysis. A knowledge-surfacing feature pulls up a customer's energy usage history and suggests decarbonization options — arming agents with the information needed to help customers make informed decisions about their energy consumption.

With the AI assistant providing consistent, simple explanations on energy sources, tariff plans, billing changes and optimal spending, customer service agents can effortlessly guide customers to the most ideal energy plan. After deploying Minerva CQ, one utility provider reported a 44% reduction in call handling time, a 12.5% increase in first-contact resolution and average savings of \$2.67 per call.

Speech AI is expected to continue to help utility providers reduce training costs, remove friction from customer service interactions and equip field technicians with voice-activated tools to boost productivity and improve safety — all while enhancing customer satisfaction.

Because public service programs are often underfunded and understaffed, citizens seeking vital services and information are at times left waiting and frustrated. To address this challenge, some

federal- and state-level agencies are turning to speech AI to achieve more timely service delivery.

The Federal Emergency Management Agency uses automated speech recognition systems to manage emergency hotlines, analyze distress signals and direct resources efficiently. The U.S. Social Security Administration uses an interactive voice response system and virtual assistants to respond to inquiries about social security benefits and application processes and to provide general information.

The Department of Veterans Affairs has appointed a director of AI to oversee the integration of the technology into its healthcare systems. The VA uses speech recognition technology to power note-taking during telehealth appointments. It has also developed an advanced automated speech transcription engine to help score neuropsychological tests for analysis of cognitive decline in older patients.

Additional opportunities for speech AI in the public sector include real-time language translation services for citizen interactions, public events or visiting diplomats. Public agencies that handle a large volume of calls can benefit from multilingual voice-based interfaces to allow citizens to access information, make inquiries or request services in different languages.

Speech and translation AI can also automate document processing by converting multilingual audio recordings or spoken content into translated text to streamline compliance processes, improve data accuracy and enhance administrative task efficiency. Speech AI additionally has the potential to expand access to services for people with visual or mobility impairments.

From vehicle sales to service scheduling, speech AI can bring numerous benefits to automakers, dealerships, drivers and passengers alike.

Before visiting a dealership in person, more than half of vehicle shoppers begin their search online, then make the first contact with a phone call to collect information. Speech AI chatbots trained on vehicle manuals can answer questions on technological capabilities, navigation, safety, warranty, maintenance costs and more. AI chatbots can also schedule test drives, answer pricing questions and inform shoppers of which models are in stock. This enables automotive manufacturers to differentiate their dealership networks through intelligent and automated engagements with customers.

Manufacturers are building advanced speech AI into vehicles and apps to improve driving experiences, safety and service. Onboard AI assistants can execute natural language voice commands for navigation, infotainment, general vehicle diagnostics and querying user manuals. Without the need to operate physical controls or touch screens, drivers can keep their hands on the wheel and eyes on the road.

Speech AI can help maximize vehicle up-time for commercial fleets. AI trained on technical service bulletins and software update cadences lets technicians provide more accurate quotes for repairs, identify key information before putting the car on a lift and swiftly supply vehicle repair updates to commercial and small business customers.

With insights from driver voice commands and bug reports, manufacturers can also improve vehicle design and operating software. As self-driving cars become more advanced, expect speech AI to play a critical role in how drivers operate vehicles, troubleshoot issues, call for assistance and schedule maintenance.

Speech AI has the potential to impact nearly every industry.

In Smart Cities , speech AI can be used to handle distress calls and provide emergency responders with crucial information. In Mexico City, the United Nations Office on Drugs and Crime is developing a speech AI program to analyze 911 calls to prevent gender violence. By analyzing distress calls, AI can identify keywords, signals and patterns to help prevent domestic violence against women. Speech AI can also be used to deliver multilingual services in public spaces and improve access to transit for people who are visually impaired.

In higher education and research , speech AI can automatically transcribe lectures and research interviews, providing students with detailed notes and saving researchers the time spent compiling qualitative data. Speech AI also facilitates the translation of educational content to various languages, increasing its accessibility.

AI translation powered by LLMs is making it easier to consume entertainment and streaming content online in any language. Netflix , for example, is using AI to automatically translate subtitles into multiple languages. Meanwhile, startup Papercup is using AI to automate video content dubbing to reach global audiences in their local languages.

In the modern consumer landscape, it's imperative that companies provide convenient, personalized customer experiences. Businesses can use NLP and the translation capabilities of speech AI to transform the way they operate and interact with customers in real time on a global scale.

Companies across industries are using speech AI to deliver rapid, multilingual customer service responses, self-service features and information and automation tools to empower employees to provide higher-value experiences.

To help enterprises in every industry realize the benefits of speech, translation and conversational AI, NVIDIA offers a suite of technologies.

NVIDIA Riva , a GPU-accelerated multilingual speech and translation AI software development kit, powers fully customizable real-time conversational AI pipelines for automatic speech recognition, text-to-speech and neural machine translation applications.

NVIDIA Tokkio , built on the NVIDIA Omniverse Avatar Cloud Engine , offers cloud-native services to create virtual assistants and digital humans that can serve as AI customer service agents.

These tools enable developers to quickly deploy high-accuracy applications with the real-time response speed needed for superior employee and customer experiences.

Join the free Speech AI Day on Sept. 20 to hear from renowned speech and translation AI leaders about groundbreaking research, real-world applications and open-source contributions.

Original URL: <https://blogs.nvidia.com/blog/2023/09/07/speech-ai-for-industries/>