BDO®

BDO KNOWS:
TECHNOLOGY

VOICE RECOGNITION TURNS UP THE VOLUME

By Hank Galligan and Mark Giamo

The voice recognition market will be worth $18 billion by 2023, according to Markets and Markets.

What's driving its explosive growth? There are many reasons, but among the most prominent is a growing desire for the personalized, convenient services that voice recognition products can offer. The devices and applications using voice recognition software for businesses and consumers can run the gamut, from security verification tools for mobile personal banking apps to voice-controlled food delivery applications.

This machine learning-enabled software is on its way to becoming the primary input and output system for computers and other tech devices. New developments in speech processing and machine learning, lower costs of implementation, and the technology's greater processing power have all contributed to businesses' and consumers' ability to adopt the software in their everyday personal and professional lives. People are now more eager than ever to interact with machines that can not only respond to humans, but engage with people by processing people's individual selves within a specific context. Voice recognition can now tell the difference between a doctor's instructions and those of a lawyer's, allowing it to be adopted by a variety of industry professionals. The tech industry has been anticipating its success for years; its rise is eagerly awaited, and many industries are ripe for its benefits.

FROM BABY TALK TO TOTAL FLUENCY

Voice recognition harnesses the power of machine learning to allow humans to interact with technology in ways never seen before. At its foundation, voice recognition technology encompasses a statistical method called the Hidden Markov Model (HMM) that measures the probability of speech patterns. A feedback loop is where machine learning enters the picture. With the loop, guided by algorithms structured by machine learning, the voice recognition software can teach itself how to better understand human speech patterns and respond accordingly.

HOW DO I GET MORE INFORMATION?

BRIAN BERNING, Cincinnati
513-592-2420 / bberning@bdo.com

TIM CLACKETT, Los Angeles
310-557-8201 / tclackett@bdo.com

SLADE FESTER, Silicon Valley
408-352-1951 / sfester@bdo.com

DEMETRIOS FRANGISKATOS, New York
212-885-7397 / dfrangiskatos@bdo.com

HANK GALLIGAN, Boston
617-422-7521 / hgalligan@bdo.com

AFTAB JAMIL, Silicon Valley
408-352-1999 / ajamil@bdo.com

BRYAN LORELLO, Austin
512 391-3515 / blorello@bdo.com

ANTHONY REH, Atlanta
404-979-7148 / areh@bdo.com

DAVID YASUKOCHI, Orange County
714-913-2597 / dyasukochi@bdo.com

HOW DOES SIRI WORK?

Automatic Speech recognition (ASR)

NLP annotation
(tokenizzation, POS tagging, NER, ...)

Pattern-action mappings

Service manager

Internal & external APIS

email

SMS

maps

weather

stocks

etc.

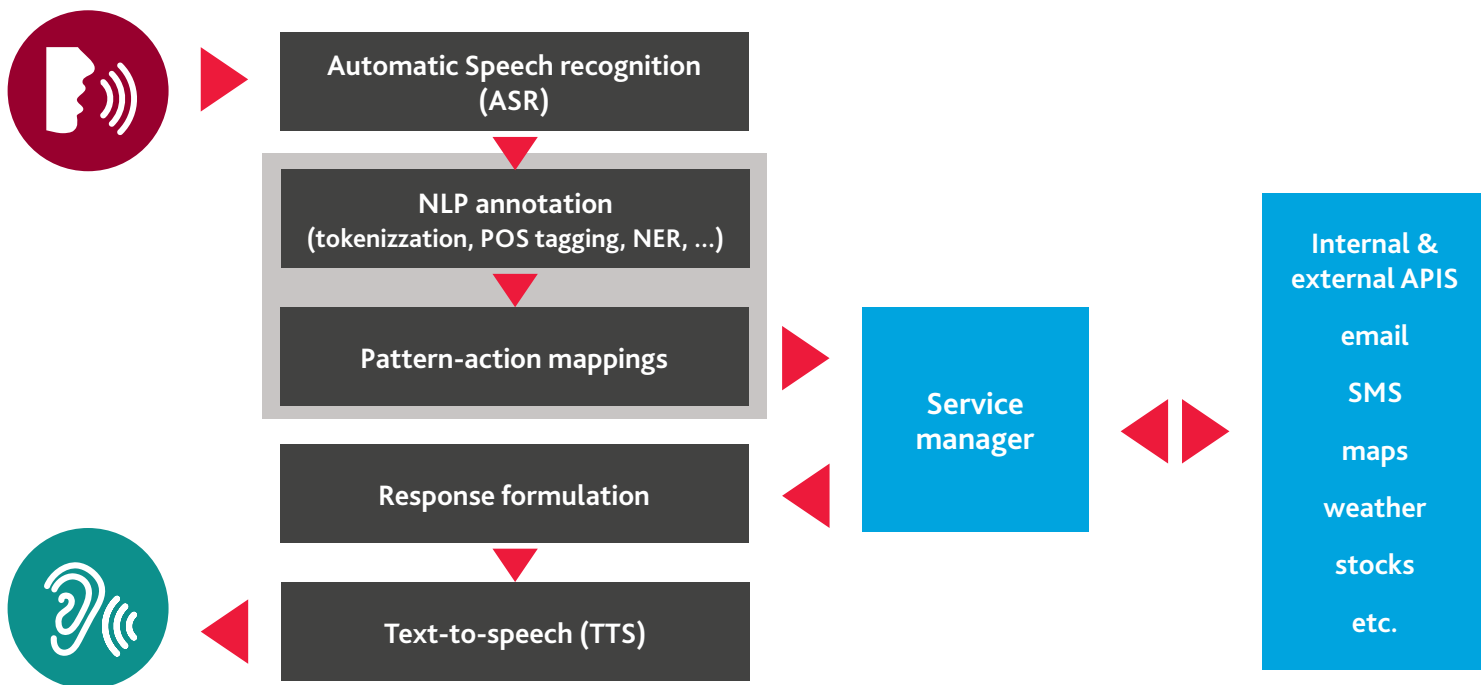**Response formulation**

Text-to-speech (TTS)

Image originally developed by Dianna McDougall; Source: NLU course, Stanford University

In the 1950s, speech computation originated with computers called the "Audrey" that understood numbers. Then IBM's "shoebox" technology developed the ability to comprehend 16 words. The technology quickly advanced during the 1970s, when the U.S. Department of Defense heavily invested in research and development projects. Gradually, researchers and scientists strengthened machine learning's capabilities through research in automatic speech recognition, natural language processing, pattern-action mappings, response formulation, text-to-speech processes, and other machine learning subsectors. Now, speech recognition uses these various machine learning functions so that a device can listen to a user's request and respond, or independently propose an action to its users.

Further developments, such as MIT's **power-efficient chip** specialized for automatic speech recognition, ensure bright prospects for voice recognition's continued adoption.

BEYOND CONSUMER PRODUCTS

Today, many consumers know and recognize personal tech products that incorporate voice recognition, such as Apple's Siri application or Amazon's Alexa. Using these products for customer service, translation, and personal assistance are just a few of the current consumer-facing services.

Nevertheless, while voice recognition software's adoption has exploded, its full potential has yet to be realized, with current applications only beginning to scratch the surface — especially when it comes to B2B disruption. Eventually, voice recognition could become the primary input and output process for personal computers. It is expected to play a huge role in developing customer support software, translation, and dictation applications, to name a few possibilities.

Banking and payment applications are particularly prime for speech recognition's capabilities. Already, banks have begun to use speech recognition to strengthen their security authentication processes, while payment applications allow users to complete payments simply by speaking into their devices. Thirty-one percent of consumers are expected to use speech recognition for payments by 2022, **according** to Business Insider Intelligence. Search engine's optimization levels could also improve as more internet users begin using speech recognition, which would spill over into virtually every B2B industry imaginable. If speech recognition becomes the most popular computer interface, every industry will have to rethink how customers interact with their digital offerings.

The less it costs to develop voice recognition software and the more energy-efficient the hardware becomes, the easier it will be to integrate voice recognition into products and services. Both tech and non-tech companies will need to figure out how to do so to keep up with competition and demand.

RAISE YOUR VOICE: BUSINESS OPPORTUNITIES FOR TECH

Voice and speech recognition applications demand new components, products and services and provide new business opportunities for tech companies. For midmarket and high-growth tech companies, some of the areas for growth and innovation include:

▶ **Integrate with Existing Systems:** Companies will need to understand how interactions with voice recognition could best benefit their customers' needs based on both their typical input and output devices. If a device fails to communicate or integrate with key technologies, its adoption rate has a good chance of quickly spiraling downward. Solution providers will need to think about how consumers are interacting with the software across various platforms, from desktops to tablets to cell phones.

▶ **Expand Software Options:** Once companies have considered how to introduce voice recognition to existing and new customers, they'll need to decide what type of software will be most compatible with the intended service or product. Since every application of voice recognition may have a different intention, companies will need to decide whether to buy a licensed software that's already developed or build its own software from scratch.

This provides a huge business opportunity for tech companies to create a range of voice recognition software offerings with features for different applications. The U.S. Bureau of Labor Statistics **estimates** that the software industry's output will grow to $205.6 billion in 2022, with voice recognition innovation poised to help drive its expansion.

Enhancing the accuracy of voice recognition is another opportunity for software developers. Since voice recognition's machine learning algorithms learn via listening to real humans speak, the technology will also have to learn different dialects, styles of speaking and languages. As voice recognition is integrated into more products and services, or even replaces certain types of work, its accuracy and dependability will be key to its success.

▶ **Give Shape to the Voice:** Voice recognition also creates enormous market opportunities for hardware companies that manufacture the physical components used to give the software a tangible device or body. Chips enabled with swift processing capabilities and robotics parts will be needed to create personal assistant devices like Amazon's Alexa, or to fashion robotics voiced by the software.

The International Data Corporation (IDC) **forecasts** global spending on robotics, including purchases of robotic systems, system hardware, software, robotics-related services, after-market robotics hardware, commercial purchase of drones and

after-market drone hardware, to more than double by 2020, growing from $91.5 billion in 2016 to more than $188 billion in 2020. For tech companies developing voice recognition services, the growing robotics industry presents a sizable market opportunity to integrate voice recognition into the hardware of the future.

▶ **Ensure Private Conversations:** Data and information privacy concerns will shape many developments in emerging tech, and companies have the opportunity to develop products and services that meet these evolving cybersecurity needs. When devices eventually sit inside of consumers' homes and offices, significant privacy questions need to be addressed. If a device has the capability to be listening to its surroundings all the time, what kind of implications does this technology have for confidential information sharing?

In China, for example, security researchers recently **discovered** a way to silently activate voice recognition systems, meaning that people nearby are not aware they are being listened to and potentially recorded. Although concrete regulations around voice recognition have yet to be finalized in the U.S., cyber breaches are top of mind in the tech industry, especially with technology like voice recognition that can capture word-for-word conversations.

In the broader scheme of emerging technology, voice recognition represents only a small piece of modern innovation. Once companies learn to master software such as voice recognition, partnering these new applications with the power of automaton, virtual reality, or differentiated artificial reality can exponentially grow a business' tech power.

Tech companies and professionals who embrace voice recognition stand to seize big opportunities in the year ahead.

Hank Galligan is an assurance director in BDO's Technology practice with a focus on software. He can be reached at hgalligan@bdo.com.

Mark Giamo is an assurance managing partner with a focus on the technology industry. He can be reached at mgiamo@bdo.com.

People who know Technology, know BDO.

BDO TECHNOLOGY PRACTICE

BDO is a national professional services firm providing assurance, tax, and advisory services to a wide range of publicly traded and privately held companies. Guided by core values including competence, honesty and integrity, professionalism, dedication, responsibility and accountability for 100 years, we have provided quality service and leadership through the active involvement of our most experienced and committed professionals.

BDO works with a wide variety of technology clients, ranging from multinational Fortune 500 corporations to more entrepreneurial businesses, on myriad accounting, tax and other financial issues.

ABOUT BDO

BDO is the brand name for BDO USA, LLP, a U.S. professional services firm providing assurance, tax, and advisory services to a wide range of publicly traded and privately held companies. For more than 100 years, BDO has provided quality service through the active involvement of experienced and committed professionals. The firm serves clients through more than 60 offices and over 550 independent alliance firm locations nationwide. As an independent Member Firm of BDO International Limited, BDO serves multi-national clients through a global network of 67,700 people working out of 1,400 offices across 158 countries.

BDO USA, LLP, a Delaware limited liability partnership, is the U.S. member of BDO International Limited, a UK company limited by guarantee, and forms part of the international BDO network of independent member firms. BDO is the brand name for the BDO network and for each of the BDO Member Firms. For more information please visit: www.bdo.com.

Material discussed is meant to provide general information and should not be acted on without professional advice tailored to your firm's individual needs.