Market Guide for Speech-to-Text Solutions

Published: 22 April 2020 **ID:** G00451355

Analyst(s): Anthony Mullen, Bern Elliot, Adrian Lee

STT solutions transcribe speech into text, enabling a range of applications that include text analytics, conversational assistants, speech data mining and speech translation. Application leaders must consider the specific capabilities they are looking for to properly evaluate vendor offerings.

Key Findings

- Speech-to-text (STT) providers have enlarged their offerings, beyond simple transcription, to offer a raft of voice-related services, from authentication to real-time alerts, compliance and emotion detection.
- Companies deploying STT applications typically underestimate the amount of customization in the form of dictionary and language assets needed for their use cases.
- STT language and acoustic assets, such as taxonomies and lexicons, are not yet very transferable or interoperable with other natural language technology (NLT) platforms.

Recommendations

For successful deployments of STT-powered applications, application leaders responsible for customer service and support technology should:

- Pick a solution that can meet the needs of their industry, one that supports customization and features that can serve multiple use cases.
- Budget for additional work to configure the solution for the language of the business by testing vendors in POCs with typical content to gauge any additional effort.
- Ensure that STT, NLT and analytic applications can leverage STT investments by validating that applications can share access to key semantic assets such as dictionaries, taxonomies, synonyms and lexicons.

Strategic Planning Assumption

By 2025, 40% of all inbound voice communications to call centers will use speech-to-text technology.

Market Definition

Gartner defines speech-to-text (STT) platforms as business applications that process speech content, either live or in batch to produce:

- A **transcript** of the conversation
- **Metadata** about the call, the callers, attributes of call, emotional context
- **Value-added services** (e.g., biometric, legal)
- **Workflow tools** to support downstream work (e.g., intent detection, CRM updates)

The capabilities of STT solutions vary. At a minimum, providers can offer a set of generic APIs with no tailored industry offering. More advanced solutions support complex deployments of edge technologies tailored to specific industries such as medical and legal. As natural language experiences are rapidly adopted by customers, users and employees, STT solutions must address a number of deployment configurations and be tailored for end-user domain knowledge to improve their accuracy.

Popular applications of the technology include:

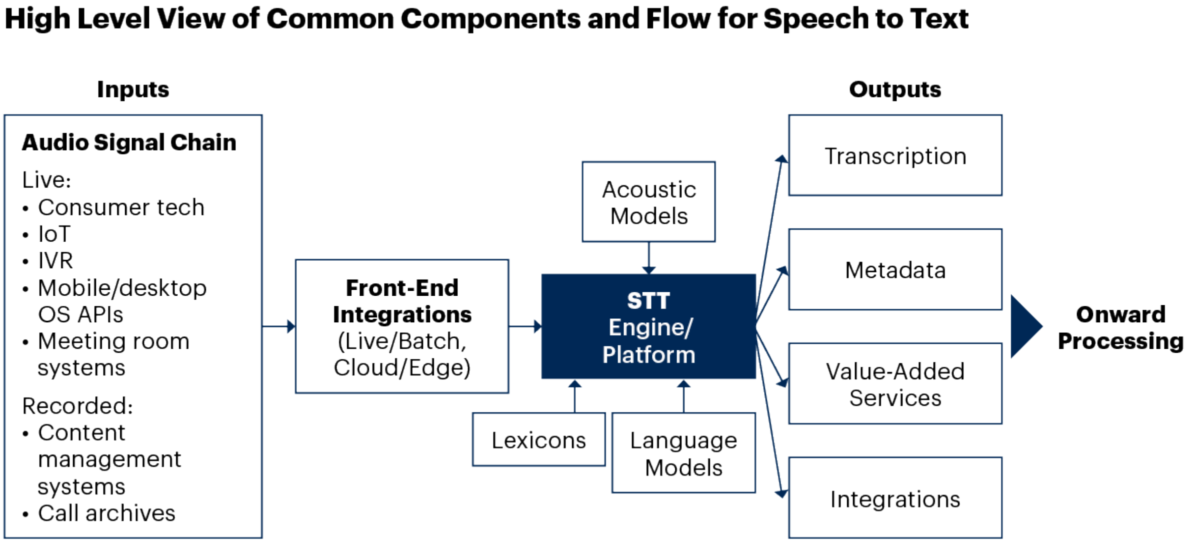
- **Customer-facing experiences** ranging from product, sales and customer service to commerce.
- **Employee use cases** such as meeting room solutions, virtual assistants and more generally as application voice-controlled front ends.
- **Business intelligence use cases** to enable text analytics systems.

Whereas before these platforms were considered as a stand-alone purchase, we now see more mature buyers consider them in light of other NLT capabilities. These buyers require virtual assistants, translation, analytics and natural language generation (NLG), and the ability to share language models, as part of a wider communications automation suite.

Market Description

While there are a number of much smaller steps in the mechanics to converting speech to text, the broad sweep of how these systems operate is shown in Figure 1. The platforms interface to an audio signal (live or in batch) and use neural networks in conjunction with semantic models to create a variety of outputs.

Figure 1. High-Level View of Common Components and Flow for Speech to Text



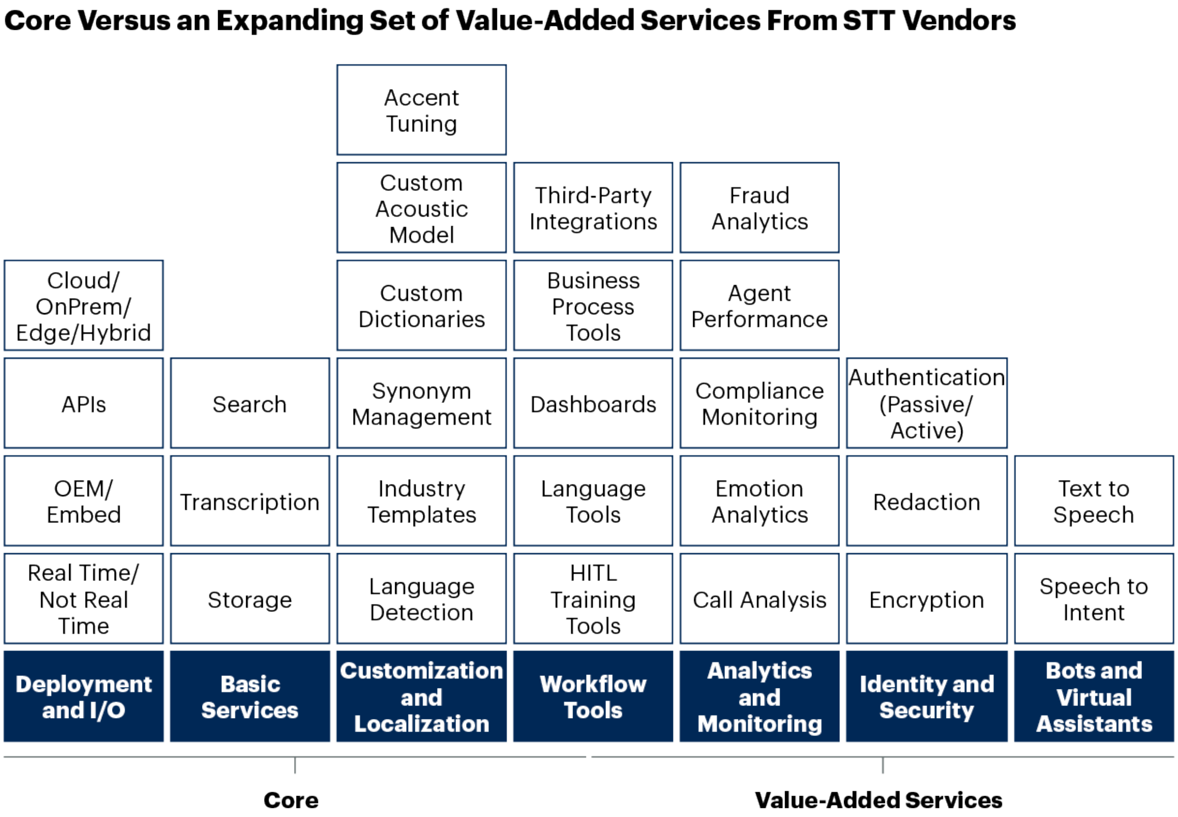
Source: Gartner

451355_C

- **Audio signal chain.** Typically for STT providers, the initial audio signal chain or file is something provided by third-party systems. This may be live such an interactive voice response (IVR) or virtual assistant being run by the client, a mobile app, a meeting room system or a real-time audio feed. It could also be a recorded source provided by a content management system, a URL from an API or from call archives.
- **Front-end integrations.** This element of the platform deals with integrating the STT providers' environment with the speech source.
- **STT engine/platform.** This is the core capability of the platforms where external language assets such as acoustic models and semantic assets (lexicons, language models) are used in conjunction with deep learning to translate raw speech into text.
- **Supporting assets.** There are a number of approaches used to support STT. The elements below are the most common:
  - **Acoustic models.** A model that learns from audio records what the relationship is between an audio signal and the phonemes (or other linguistic units) that comprise speech.
  - **Lexicons.** These store the words and other related information to support the STT process such as pronunciations and parts of speech. Providers typically allow end users to create their own lexicons and use them alongside the core application lexicons.
  - **Language models.** Support the speech recognition engine in determining how likely a word sequence is, independent of the acoustics.

- **Outputs.** Outputs vary from STT systems. At a minimum they will include the verbatim or transcription of the speech and associated metadata around the conversation. Some platforms have up to 50 types of metadata that can be provided alongside the transcript including behavioral and emotional data.
- **Value-added services (VAS).** These can provide other outputs such as call routing, authentication and other services. These VAS are growing rapidly for this market — see Figure 2 for illustrative VAS (see Table 1).
- **Integrations.** Depending on the use of the STT system, platforms can integrate with other enterprise assets such as CRM systems and other applications.
- **Onward processing.** End users may take the materials and outputs provided to support further business processes such as chatbots/virtual personal assistants (VPAs) or analytics applications.

Figure 2. Core Versus an Expanding Set of VAS From STT Vendors



Core Versus an Expanding Set of Value-Added Services From STT Vendors

Deployment and I/O	Basic Services	Customization and Localization	Workflow Tools	Analytics and Monitoring	Identity and Security	Bots and Virtual Assistants
		Accent Tuning				
		Custom Acoustic Model	Third-Party Integrations	Fraud Analytics		
Cloud/ OnPrem/ Edge/Hybrid		Custom Dictionaries	Business Process Tools	Agent Performance		
APIs	Search	Synonym Management	Dashboards	Compliance Monitoring	Authentication (Passive/ Active)	
OEM/ Embed	Transcription	Industry Templates	Language Tools	Emotion Analytics	Redaction	Text to Speech
Real Time/ Not Real Time	Storage	Language Detection	HITL Training Tools	Call Analysis	Encryption	Speech to Intent

Source: Gartner

451355_C

Table 1. Description of VAS

Value-Added Service	Service Definition
Deployment and I/O	
Cloud/on-premises/edge/hybrid	Mode of solution deployed in customer environment.
APIs	Interface to provide programmatic access to service functionality and data within an STT application or a database.
OEM/embed	Mode of solution deployed in device or application environment.
Real time/non real time	Ability to handle batch (non-real time) files as well as real-time speech and transcription.
Basic Services	
Search	Search aggregates the results of a user-initiated search and presents those results back to the user.
Transcription	Transcription to extract insights from real-time or prerecorded voice streams.
Storage	Data management options to run on a server, storage network device or storage device to aid in managing and protecting the data.
Customization and Localization	
Custom acoustic model	A customized model that learns from audio records what the relationship is between an audio signal and the phonemes or other linguistic units that comprise speech.
Custom dictionaries	A usually large collection of textual/statistical data that can be customized to infer and validate business rules.
Synonym management	Management of like and similar words.
Industry templates	Prebuilt elements designed for specific industries covering training assets, model variations, custom workflows.
Language detection	Automatic detection of language used in a multilanguage scenario.
Workflow Tools	
Third-party integrations	Integrations with external enterprise applications.
Business process analysis tools	Primarily intended for use by business end users looking to document, analyze and streamline complex processes, thereby improving productivity, increasing quality, and becoming more agile and effective.

Value-Added Service	Service Definition
Dashboards	Reporting mechanisms that aggregate and display metrics and key performance indicators (KPIs), enabling them to be examined by users before further exploration via additional business analytics (BA) tools.
Language tools	Determine which languages are supported. It is also useful to know which languages were developed with native corpora. It should be noted that quality of STT by a vendor may vary widely between languages.
HITL training tools	The ability to leverage and manage domain experts via UIs to support development of speech models.
Analytics and Monitoring	
Fraud analytics	Fraud detection protects customer and enterprise information, assets, accounts and transactions through the real-time, near-real-time or batch analysis of activities by users.
Agent performance	Real-time and postcall personalized guidance for agents based on customer interactions, particularly in customer service or sales calls.
Compliance monitoring	Either real-time or near-real-time solutions that protect communications sent over applications, while also ensuring regulatory compliance.
Emotion analytics	Recognizing a person's emotional state — for example, anger, confusion or deceit across both voice and nonvoice channels.
Call analytics	To classify calls, trigger alerts/workflows, and drive operational and employee performance across the enterprise.
Identity and Security	
Authentication (passive/ active)	Authentication may be natively supported in product or services (including other security tools).
Redaction	Refers to the permanent removal of information, not the masking or obfuscation of data.
Encryption	Process of systematically encoding a bit stream before transmission, so that an unauthorized party cannot decipher it.
Speaker recognition	Speaker recognition is the process of automatically recognizing who is speaking using speaker-specific information in speech waves.
Bots and Virtual Assistants	
Speech to intent	Where the processed speech input (utterance) is matched to the appropriate handler of the request. This usually uses machine learning.
Text to speech	Natural language generation (NLG), or the ability to generate natural language responses in speech.

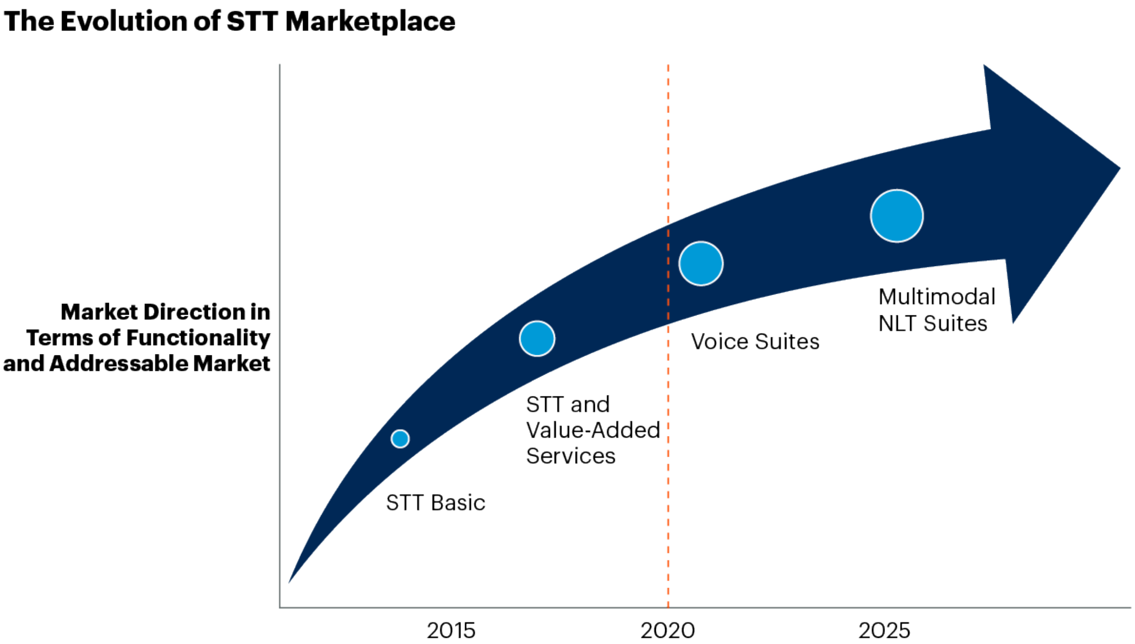
Source: Gartner (April 2020)

Market Direction

While the core technology for STT, namely converting speech to text with foundational deep learning approaches, has reached the Plateau of Productivity in the Gartner "Hype Cycle for Artificial Intelligence, 2019," the underlying technologies and market approaches haven't stood still. The recent evolution is the shift to richer solutions — rather than simply sell raw speech-to-text solutions, vendors are increasingly looking to provide workflow and services across an increasing range of conversational and analytics experiences.

Over the next five years, we expect to see a further evolution of offers into broader voice services presented as synergistic suites. Niche solutions will continue to play a role as best-of-breed offers, supporting less common languages (e.g., Malayan or Sinhalese) or applications targeting specific niche requirements. Meanwhile, broad suites from the very large artificial intelligence (AI) cloud providers will increasingly dominate the NLT ecosystems. These will share language and acoustic assets for broader performance across STT, text to speech (TTS), speech mining, translation, conversational platforms, NLG and others (see Figure 3).

Figure 3. The Evolution of STT Marketplace


The Evolution of STT Marketplace
Market Direction in Terms of Functionality and Addressable Market
STT Basic
STT and Value-Added Services
Voice Suites
Multimodal NLT Suites
2015
2020
2025

Source: Gartner

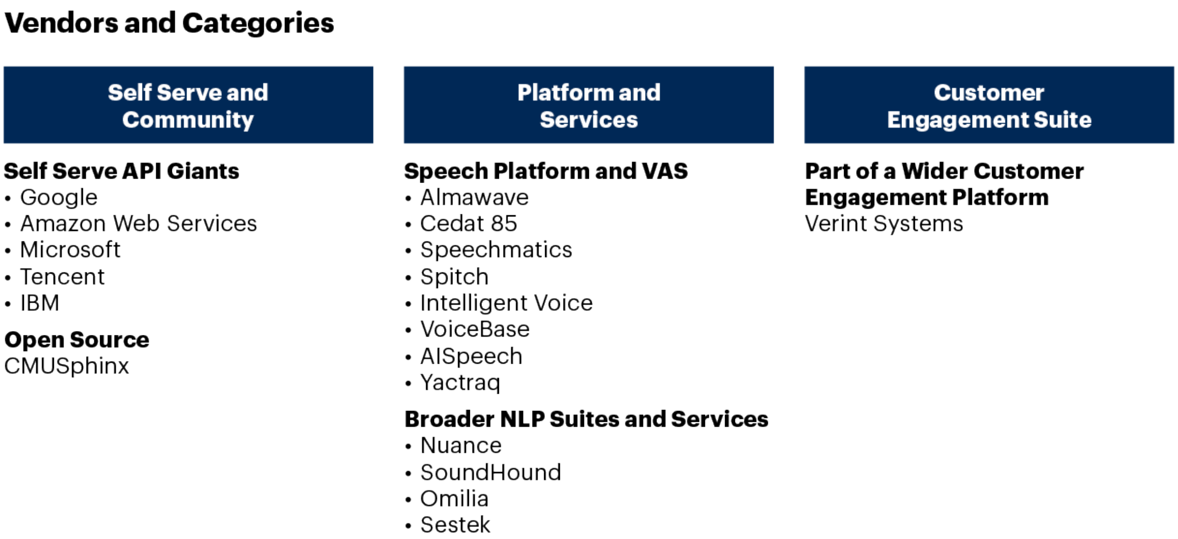
451355_C

Unlike the chatbot market, with low barriers to entry, the STT market does not have thousands of players. A major barrier is the need for high volumes of speech training data, skills and investment in speech and language modeling. Today, we often see STT provided (often white-labeled) as part of some broader CX, CRM and engagement platforms.

Market Analysis

To date, the market can be categorized as having three main types of vendor (see Figure 4).

Figure 4. Vendors and Categories



Vendors and Categories

Self Serve and Community

Self Serve API Giants

- Google
- Amazon Web Services
- Microsoft
- Tencent
- IBM

Open Source

CMUSphinx

Platform and Services

Speech Platform and VAS

- Almawave
- Cedat 85
- Speechmatics
- Spitch
- Intelligent Voice
- VoiceBase
- AISpeech
- Yactraq

Broader NLP Suites and Services

- Nuance
- SoundHound
- Omilia
- Sestek

Customer Engagement Suite

Part of a Wider Customer Engagement Platform

Verint Systems

Source: Gartner

451355_C

We see three broad categories that can be useful when grouping the market.

1. **Self-serve (and community).** Aimed widely at developers and data scientists. The barrier to using STT technology has come down with the self-serve approaches used by Google, Tencent, Amazon and Microsoft (see "Magic Quadrant for Cloud AI Developer Services") — aimed at developers. Giants in this space often have a raft of AI services alongside STT. There are open-source players in this space, such as CMUSphinx, although their adoption rates are very low.

2. **Platform (and services).** Vendors here typically focus solely on STT. This class of vendors has the most well-developed value-added services (see Figure 4 above). They differentiate from the tech giants by speech features, ability to deliver edge-based models, domain customization and wraparound system integration support.

3. **Customer engagement suite.** This is where the STT capability is provided as part of a customer service and support suite. Vendors in this space often don't have their own STT engine and often make use of third-party, white-label licenses.

While the self-serve heavyweights have been slow to create domain-specific offerings, we now see them increase efforts around markets like automotive and healthcare.

- Amazon Transcribe Medical uses machine learning to extract relevant medical information from transcriptions, such as medical condition, medication, dosage, strength, and frequency. Nuance also has a well-designed cognitive driving experience that uses speech and other modal signals (vision, touch, facial expression) to shape service delivery.

- Nuance has created a spinoff for their automotive industry product called Cerence, providing richer multimodal/multisensory car experiences.
- Tencent has extended its STT capabilities to the communications and media industry, to digitize and transcribe native folk songs, nursery rhymes and poetry for media and content publishers.

The broader natural language technologies market has rapidly evolved on many fronts over the last two years. This metamarket will shape the evolution of STT solutions. We expect to see the following evolutions in the market:

- **Better out-of-the-box performance and integration.** While today's systems require lots of effort to deeply embed them inside your knowledge and business systems, this is something that will become easier over time. The improvement of out-of-the-box capabilities and services will be propelled by massive repetitions of learning cycles and huge volumes of training data.
  - For example, rolling out a new language or dialect has always been a heavily configured piece of work. By using transfer learning, Amazon recently demonstrated the bootstrapping of other models. As a new wave of users come online to use speech-to-text services, this longtail of developers will rapidly increase the use of this technology and value of the overall market. See the Gartner "Magic Quadrant for Cloud AI Developer Services."
- **Strategic alliances with tech heavyweights as the conversational market consolidates.** We expect greater strategic alliances by focused providers with the tech heavyweights as a result of conversational platform consolidation.
  - For example, Amazon and Salesforce at the end of 4Q19 developed real-time call analysis with STT provided by Amazon and integrated live into the UIs and business rules systems of Salesforce. Speech will be another driver to reduce the number of players in the chatbot and virtual agent market. And users increasingly want capabilities that they can deploy across both text- and speech-based ecosystems. Interactions LLC, for example, has pivoted to embed a wider set of voice services into their conversational platform.
- **An increase in the partner services specifically around voice experience design.** Simply having access to the baseline STT technologies does not make for a good voice experience design. Design agencies, who have evolved through web, social and mobile will partner with practitioners in STT to deliver richer cognitive design services.
- **Greater collaboration of semantic assets between vendor STT and chatbots and NLG projects.** As the number of NLT projects increases inside an organization, it becomes clear that having uncoordinated language and knowledge models about your business across multiple vendors is a major inhibiter to AI maturity. We expect to see STT players become more present in the knowledge engineering space. The accuracy of verbatim will continue to be a major focus but the richer metadata looking at behavioral models will be a major differentiator.

The barrier to using STT technology has come down with the self-serve approaches by Google, Tencent, Amazon, Microsoft (see “Magic Quadrant for Cloud AI Developer Services”) — aimed at developers.

Pricing Models

Pricing of STT-related services is highly variable and is often not published. When it is published, the published price is often not what ends up in the deal. Gartner sees the published pricing for basic STT at about $0.024 (2.4 cents per min). However, pricing for high volume (tens of millions of minutes per month) will be considerably lower depending on features and turnaround time SLAs.

One factor that makes pricing difficult is that there are many features such as custom vocabulary, PCI redaction, predictive models and call metrics. These become very important, and are offered at varied added cost or are sometimes bundled. As a result, there is effectively a spectrum of pricing that goes from basic transcription to full-blown speech analytics with many steps along the way depending on the use case.

When considering solutions, note that price may not be as important as quality. Additionally, quality evaluations may change significantly depending on how well a vendor can accommodate customization.

Representative Vendors

The vendors listed in this Market Guide do not imply an exhaustive list. This section is intended to provide more understanding of the market and its offerings.

The inclusion criteria below were used for longlisting vendors in this report:

- Have a native speech-to-text engine (rather than dependent on a third-party engine). Without this criterion it was automatic exclusion.
- Can take input from a number of formats (e.g., audio, video).
- Handle multiple languages and improve solution accuracy through the use of custom dictionaries and acoustic models.
- Provide analytics on speech content.
- Can deploy the solution across a variety of channels, e.g., IVR, mobile apps, virtual assistants, embedded.
- Provide transcriptions of speech content.

This Market Guide is not a qualitative scoring of vendors against one another. Vendors in this report are illustrative of aspects of the market that we deem to be important. Readers should use this report to understand the current market dynamics and the evolution of this market.

Market Introduction

While quality of text transcription is the most critical function for the vendors in this market, STT vendors also differentiate themselves by offering different features and focus areas. The tables below summarize vendor capabilities in four areas:

Table 2 provides an overview of the industries where the vendors reported they had particular experience, for instance, banking and securities or healthcare.

Table 3 provides information regarding the deployment options, for instance, real time and/or batch execution, or on-premises and/or cloud.

Table 4 indicates the languages and dialects that are supported.

Table 5 provides an overview of vendor support for key features.

Table 2. By Industry

Vendor	Aerospace and Defense	Banking and Securities	Business Services	Communications and Media	Education	Energy	Healthcare	Insurance	Manufacturing	Media Monitoring	Professional Services and Consulting	Public Sector	Not for Profit	Retail	Transportation	Utilities	Wholesale Trade	Others
AISpeech				x														x
Almawave		x		x		x	x	x		x	x	x			x	x		
Amazon Web Services	x	x	x	x	x	x	x	x	x	x		x	x	x	x	x	x	x
Cedat 85		x	x	x	x	x	x	x	x	x	x	x			x	x		x
CMUSphinx																		
Google		x		x			x											x
IBM		x					x	x						x	x			
Intelligent Voice	x	x	x				x	x		x		x						
Microsoft	x	x	x	x	x	x	x	x	x	x	x	x		x	x	x	x	x
Nuance	x	x	x	x	X	X	X	x		X		X		X	X	X		

Vendor	Aerospace and Defense	Banking and Securities	Business Services	Communications and Media	Education	Energy	Healthcare	Insurance	Manufacturing	Media Monitoring	Professional Services and Consulting	Public Sector	Not for Profit	Retail	Transportation	Utilities	Wholesale Trade	Others
Omilia		x		x		x	x	x				x		x	X	X		x
Sestek		x		x			x	x			x			X				
SoundHound				x					x									x
Speechmatics				x						x	x							
Spitch		x	x	x			x	X	x		x	x		x	x	x		X
Verint Systems		x	x	x	x	x		x	x			x		x	x	x		x
VoiceBase		x	x	x	x		x	x	x	x	x	x	x	x	x	x		x

Source: Gartner (April 2020)

Table 3. By Deployment

Vendor	Real Time/Batch	Deployment Types: On-Premises, Cloud (Private/ Public)	Embedded and Edge Capabilities
AISpeech	Both	On-device	Internet of Things (IoT) devices, onto a Mobile
chanAImawave	Both	On-premises, cloud (private and public)	Embedding on mobile/IoT, on-premises or cloud
Amazon Web Services	Offline/batch/ asynchronous	Cloud only	OEM via cloud only
Cedat85	Both	Both	Mobile SDKs for apps, embedded for offline STT in SBC (single board computers)
CMUSphinx	Open source	No cloud, local (on-device/on-premises) only	Yes, via PocketSphinx
Google	Both	Cloud only	Edge-based and embedded models require cloud
IBM	Both	Public cloud, premium cloud	Only available via streaming over web sockets
Intelligent Voice	Both	On-premises, cloud (private and public)	Capable of running all current languages on an iOS device
Microsoft	Both	Both (via containers)	Deploy to edge devices using containers and web — models can't be downloaded and executed locally. Speech Devices SDK is a pretuned library that's paired with purpose-built, microphone array development kits
Nuance	Both	On-premises, cloud	Embedded
Omilia	Real time, batch (partial)	On-premises, cloud	Mobile device integration is possible through code embedding during application development
Sestek	Both	On-premises, cloud	SDK provided for offline use with mobile/IoT
SoundHound	Real time	Both	Yes, can embed STT without need for cloud
Speechmatics	Both	Both	No
Spitch	Both	On-premises, cloud (private and public)	Software development kit (SDK) for cloud-connected mobile/IoT. Some solutions

Vendor	Real Time/Batch	Deployment Types: On-Premises, Cloud (Private/ Public)	Embedded and Edge Capabilities
			(including biometrics) can be embedded with no need for cloud
Verint Systems	Postcall and real time	Cloud, on-premises, managed service, hybrid	No
VoiceBase	Offline batch	Core service — cloud (private/public) analytics output supports on-premises and cloud	Support for mobile or web application development using the API platform

Source: Gartner (April 2020)

Table 4. By Language

Vendor	Auto Language Detection	Languages (and Dialects)
AISpeech	Yes	Chinese, English
Almawave	No	Modern Standard Arabic (Jordan, Kuwait, Libya, Oman, Palestine, Saudi Arabia, Sudan, Syria, United Arab Emirates, Catalan), Chinese (Mandarin), Danish, Dutch, English (Australia, Caribbean, India, Ireland, U.K., U.S.), Farsi, Flemish, French (Canada, France, Switzerland) German (Germany, Switzerland, Austria), Greek, Hindi, Italian (Italy, Switzerland), Japanese, Portuguese (Brazil, Portugal), Russian, Spanish (Spain, Columbia, Mexico), Swedish, Turkish, Urdu
Amazon Web Services	No	Arabic (Gulf Arabic, Modern Standard Arabic), Chinese Mandarin (Mainland), Dutch, English (Australian, British, Indian, Irish, Scottish, U.S., Welsh), Spanish (Spain, U.S.), French (France, Canada), Farsi, German (German, Swiss German), Hebrew, Indian (Tamil, Telegu, Hindi), Indonesian, Italian, Japanese, Korean (ko-KR), Malay (ms-MY), Portuguese (Brazil, Portugal), Russian, Turkish
Cedat85	Yes	Italian, English (Irish, U.K., U.S.), Spanish (European, North American, South American), French, German, Brazilian Portuguese, Russian. Industry-specific: U.K. Medical Language Model, Italian Medical Language Model, U.K. Legal and Finance, Italian Legal, Finance, Utilities, Telco, Religious
CMUSphinx	Unknown	Spanish (Mexico, Spain), Portuguese, Mandarin, English (Indian, U.S.), Catalan, German, Greek, French, Dutch, Italian, Hindi, Kazakh, Russian
Google	Yes	120 different dialects across 64 different languages
IBM	Yes	Arabic (Modern Standard), Brazilian Portuguese, Chinese (Mandarin), Dutch (Beta), English (United Kingdom, U.S.), English (United States), French (France), German, Italian, Japanese, Korean, Spanish (Argentina, Castilian, Chilean, Colombian, Mexico, Peru). NB: many models are classified as being in beta
Intelligent Voice	Yes	English (UK, U.S., SA, AUS), Spanish (MEX, EU), Catalan, German (DE, Swiss), Portuguese (BR, EU), Dutch, Norwegian, Danish, Japanese, French, Russian, Korean, Mandarin, Tagalog, Cantonese
Microsoft	Yes	Arabic (UAE, Bahrain, Egypt, Kuwait, Qatar, Saudi Arabia), Catalan, Danish, German, English (AUS, Canada, UK, India, New Zealand, US), Spanish (Spain, Mexico) Finnish, French (Canada, France), Indian (Gujarati, Hindi, Marathi, Tamil, Telegu), Italian, Japanese, Korean, Norwegian, Dutch, Polish, Portuguese (Brazil, Portugal), Russian, Swedish, Thai, Turkish, Chinese (Mandarin, Cantonese, Taiwanese)
Nuance	Yes	Arabic (Egypt, Jordan, Saudi Arabia, UAE), Afrikaans, Austrian (German), Basque, Catalan, Galician, Valencian, Finnish, Spanish (Argentina, Columbia, Chile, Ecuador, Guatemala, Mexico, Peru, Spain, USA, Venezuela), Portuguese (Brazil, Portugal), Mandarin (China, Taiwan), Cantonese (HK), English (Australian, Canadian, England, Welsh, Indian, New Zealand, Singapore, South African, USA), Greek, Hungarian, Icelandic, Indian (Assamese, Bhojpuri, Bengali, Gujarati, Hindi,

Vendor	Auto Language Detection	Languages (and Dialects)
		Kannada, Malayalam, Marathi, Oriya, Punjabi, Tamil, Telugu, Urdu), Bengali (Bangladesh), Flemish (Dutch, Belgian), Bulgarian, French (Canada, France), Czech, Danish, Korean, Malay, Nepali, Dutch, Norwegian, Urdu, Polish, Romanian, Russian, Serbian, Slovak, Swedish, Swiss (German Swiss), Thai, Turkish, Vietnamese, Indonesian, Hebrew, Italian, Japanese
Omilia	No	English (USA, Canada, U.K., Jamaica, South Africa), French (Canada, France), Greek, Polish, Russian (Russia, Belarus, Kazakh, Ukrainian), Portuguese (Portugal), Spanish (Spain, Latin America), German, Turkish, Kazakh, Ukrainian, Mixed Ukrainian-Russian/Ukraine
Sestek	Yes	Turkish, English (American, British, Indian), Spanish, French (Metropolitan, Belgian), Azerbaijani, Arabic (Levantine, Gulf, Kingdom of Saudi Arabia, Kuwait, UAE, Qatar, Bahrein, Iraq, Oman, Egyptian), German, Russian, Urdu, Flemish, Dutch, Kurdish, Czech, Italian, Indian-Hindi, Ukrainian, Kazakh
SoundHound	Not disclosed	14 languages
Speechmatics	No	Arabic, Bulgarian, Catalan, Croatian, Czech, Danish, Dutch, English, Finnish, French, German, Greek, Hindi, Hungarian, Italian, Japanese, Korean, Latvian, Lithuanian, Mandarin, Malay, Norwegian, Polish, Portuguese, Romanian, Russian, Slovak, Slovenian, Spanish, Swedish, Turkish
Spitch	No	English (U.K. and U.S.), German (High German, Swiss German), Italian, French, Russian
Verint Systems	Yes	Arabic (Bahraini, Egyptian, Gulf, Morocco, Levantine), Mandarin (China, Hong Kong, Singapore, Taiwan), Cantonese, English (Hong Kong, Malaysia, Singapore, Arab countries, Canada, Caribbean, New Zealand, South Africa, Australia, India, U.K., U.S.), French (Canada, France), Portuguese (Brazil, Portugal), Spanish (Argentina, Chile, Colombia, Costa Rica, Dominican Republic, Mexico, Peru, Uruguay, U.S., Venezuela, Spain)
VoiceBase	No	English (U.S., U.K.), AU, SA, German, Italian, Dutch, Portuguese (Brazil), French, Italian, Spanish (Spain, Latin America)

Source: Gartner (April 2020)

Table 5. By Speech and Language Features

Vendor	Real-time captioning	Supports specific speech input types (e.g., numerical, currencies, units of measures, etc.)	STT processing at word, sentence level or both	Automated punctuation and capitalization (explicit/implicit)	Time-stamping of utterances (words and sentence)	Supports real-time confidence scoring	Diarization for multiple speakers	Real-time alerts	Indexing and searchability of audio or speech	Provides behavioral or emotional analysis	Voice biometrics — passive, active or both
AISpeech	Yes	Yes	Sentence	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
Almawave	Yes	Yes	Both	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Amazon Web Services	Yes	Yes	Both	Yes	Yes	No	Yes	Yes	No	No	No
Cedat85	Yes	Yes	Both	Yes	Yes	Yes	Yes	Yes	Yes	No	Under development
Google	Yes	Yes	Both	Yes	Yes	Yes	Yes	Yes*	Yes*	Yes	No
IBM	Yes	Yes	Sentence	No	Word	Yes	Yes	No	No	Via IBM Tone Analyzer	No
Intelligent Voice	Yes	Yes	Sentence	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Both
Microsoft	Yes	Yes	Both	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Active
Nuance	Yes	Yes	Both	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Both
Omilia	Yes	Yes	Both	No	Yes	Yes	Yes	Yes	Yes	Partial	Both

Vendor	Real-time captioning	Supports specific speech input types (e.g., numerical, currencies, units of measures, etc.)	STT processing at word, sentence level or both	Automated punctuation and capitalization (explicit/implicit)	Time-stamping of utterances (words and sentence)	Supports real-time confidence scoring	Diarization for multiple speakers	Real-time alerts	Indexing and searchability of audio or speech	Provides behavioral or emotional analysis	Voice biometrics — passive, active or both
Sestek	Yes	Yes	Both	Yes	Yes	Yes	Yes	Yes	Yes	Partial	Both
SoundHound	Yes	Yes	Sentence	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Speechmatics	Yes	No	Both	Yes	Yes	Yes	Yes (batch processing)	No	No	No	No
Spitch	Yes	Yes	Yes, both	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes, both
Verint Systems	Yes	No	Word	No	Yes	Yes	Yes	Yes	Yes	Yes	Both
VoiceBase	No	No	Word	Yes	Yes	Yes	No	No	Yes	Yes	No

Source: Gartner (April 2020)

- CMUSphinx (not listed in this table) has a basic feature set but is often extended by platforms like LIUM and VoxForge.
- Google* — offers features in conjunction with other GCP products

Table 6. By Security Features

Vendor	Encryption	Data redaction, including PCI redaction	Data confidentiality/ protection/GDPR	Management and alerting of compliance watch terms	Data residency/ federation of data from clients
AISpeech	Yes (audio and text, online and offline)	No	GDPR	Yes	Yes
Almawave	Yes (protocol level, application level, API received/sent)	Yes (personal)	GDPR	Yes	Yes (all)
Amazon Web Services	Yes (streaming does not archive and batch offers opt-out of data retention)	No	GDPR, HIPAA, PCI, SOC1/2/3	Yes, our service provides custom masks to filter words in the speech output	Data residency by regions
Cedat85	AES data encryption (audio/text) both in the cloud and on-premises. No search when encrypted	Yes (personal, financial)	GDPR	Yes	Yes
CMUSphinx	No	No	No	No	No
Google	Yes (when outside Google boundaries)	Yes (PCI decision support system [DSS] certification and PCI redaction in conjunction with other Google Cloud Platform (GCP) services — cloud data loss prevention	GDPR, Health Insurance Portability and Accountability Act (HIPAA)	Yes — in conjunction with other services (Dialogflow)	No
IBM	No	Yes	GDPR	No	No
Intelligent Voice	Yes (searchable when encrypted)	Yes	GDPR	Yes	Data federation
Microsoft	Yes	No (our service is PCI DSS compliant, and it doesn't log the data after recognition is completed; it is up	GDPR	Yes (our service provides profanity masks to filter profanity	By Geography/ region

Vendor	Encryption	Data redaction, including PCI redaction	Data confidentiality/ protection/GDPR	Management and alerting of compliance watch terms	Data residency/ federation of data from clients
		to developers to handle the transcribed data)		words in the speech output)	
Nuance	Encryption for both on-premises and cloud deployments	Yes	GDPR protection and data protection via encryption both in transit and rest	Yes (STT will transcribe in real time and pass to NL model)	Data residency by region
Omilia	Yes (encryption is supported on both on-premises and cloud solutions)	Yes	PCI DSS, ISO/IEC 27001, GDPR, SOC 2 Type 2	Yes	Yes (all)
Sestek	Yes (audio is encrypted; both cloud/on-premises are supported; searchable while encrypted)	Yes	Yes	Yes	Yes
SoundHound	No	No	GDPR	No	No
Speechmatics	Software as a service (SaaS) data at rest	No	Yes	No	No
Spitch	Partial	Yes (personal, financial)	GDPR	Yes	Data residency by region
Verint Systems	End-to-end encryption; media and transcription data are encrypted both on disk and during transmission.	Yes (personal, financial)	Personally identifiable information (PII), GDPR, PCI	Yes	Yes (data sovereignty by region)
VoiceBase	Yes (all customer data is encrypted at rest and in flight; searchable when encrypted)	Yes	GDPR, PCI, HIPAA, EU-US and Swiss-US Privacy Shield	Alerting supported via analytics output	Yes (all)

Source: Gartner (April 2020)

Vendor Profiles

AISpeech

aispeech.com

Headquarters: Suzhou, China

Products: Smartic TH1520

AISpeech was previously profiled in Gartner's Market Guide for Conversational AI Vendors in China (see "China Summary Translation: 'Market Guide for Conversational Platforms'"). It is a primarily R&D-focused company that specializes in its own TH1520 intelligent audio chips with voice algorithms. The TH1520 has on-chip storage, utilizes low standby and full-state power consumption; and supports more than 200 voice instructions when offline — making it well-suited for rapid deployment on portable and mobile devices. AISpeech chips and software are designed to enhance speech recognition and integrate acoustic models to cancel out background noise, such as in-vehicle wind noise. AISpeech is one of China's major vendors that has an end-to-end natural language platform targeted at IoT use cases.

Supporting predominantly both Chinese and English languages, it is a market leader in China that specializes in native language support for Chinese dialects with regional accents tuning. AISpeech provides an open-source natural language processing (NLP) developer platform, dialogue user interface (DUI), aimed at IoT applications for smart homes, wearables, smartphones and vehicles.

Almawave

almawave.com

Headquarters: Rome, Italy

Products: Iride Call, Iride Voice, Iride Verbal Order, Iride RT, Iride Wavebot, Audioma Box, Audioma RT, PSW (PerVoice Subtitling Workstation), FlyScribe

Almawave is a European vendor for speech technologies, focused on advanced analytics, multimedia broadcast, reporting and subtitling, to support CX management. It has been mentioned multiple times by Gartner since 2014, predominantly in the field of customer service technologies. Almawave's Iride product suite consists of multiple modules, where Iride Voice is the most relevant product to deliver core STT capabilities. Iride Voice integrates natively with its spoken language translation for both offline and real-time transcription.

Almawave's other capabilities run a comprehensive list from speech IVR, outbound calling, voice analytics, chatbots, semantic call routing and other speech services. Its considerably mature speech technology spans over 32 different Indo-European languages with features, such as conversational agents for text and voice, RPAs and guidance, knowledge management, and text analytics.

Amazon Web Services (AWS)

aws.com

Headquarters: Seattle, Washington

Products: Amazon Transcribe

Amazon Transcribe is a self-service automatic speech recognition service provided as part of the wider portfolio of Amazon natural language technology offerings.

Support of verticals is mostly down to end users self-serving with Amazon's custom vocabulary and pronunciations. However, in 2019 Amazon partnered for more dedicated offerings with Salesforce for contact center applications, and in 2019 Amazon released a variation on the core product called Amazon Transcribe Medical.

Amazon supports 30+ languages. The service offers word-level time stamps and word-level confidence scores, plus implicit capitalization and punctuation. The service is available in both real-time and batch API offerings.

Amazon is rapidly and deeply integrating their STT capability across their wider NLT offerings such as Amazon Translate, Amazon Lex (virtual assistants), Amazon Sumerian (avatars), and Amazon Polly (TTS) multimodal authoring suite. Amazon is also heavily researching new techniques such as transfer learning which will enable more rapid development of speech services across different languages and industries.

Cedat 85

cedat85.com

Headquarters: Rome, Italy

Products: Speech-i ASR (engine), Cabolo® (portable secure transcription device)

Cedat 85 has developed a modern, flexible, customizable platform around automatic speech recognition (ASR) core technology, composed of several modules that can be activated according to specific needs. These modules include capabilities like biometrics, real-time speech analytics, and voice to intent.

Cedat 85 has been operating since 1985 and has over 500 customers. The vendor has specialized in creating richer solutions and modules to support different industries and use cases. Digital4Democracy solution for government offers combinations of secure cloud and edge solutions to automate workflows and reporting. They have also developed dedicated offerings for media, telco, finance/compliance and utilities and are currently converting the British Library archives.

Cedat 85 recently released a product called Cabolo for frontline workers to support secure, flexible, portable transcription in an edge-based hardware box (that also consent live subtitling).

CMUSphinx

cmusphinx.github

Headquarters: Pittsburgh, Pennsylvania

Products: CMUSphinx toolkit

CMUSphinx is a speaker-independent large vocabulary continuous speech recognizer released under BSD style license — originally designed at Carnegie Mellon University. It is a collection of open-source tools and resources that allow researchers and developers to build speech recognition systems, with specific focus on low-resource platforms. The project is mastered on GitHub. Along with other open-source players in this space, including Kaldi and HTK, CMUSphinx has a basic feature set but is often extended by platforms like LIUM and VoxForge.

Google

google.com

Headquarters: Mountain View, California

Products: Google Cloud Speech-to-Text API

Google aims their speech-to-text (STT) services at self-serve developers. They offer a selection of prebuilt models to support modalities including: short voice commands/search, phone calls, audio from video, and a default model. Speech models can be customized to work in any vertical or domain through the real-time speech adaptation feature.

Google's STT service supports 127 different dialects across 71 different languages. Outside of core STT, Google Cloud provides value-adding services such as spoken and unspoken punctuation, speaker diarization and language identification. Google's approach to STT allows industries to customize models based on their needs, in real time, as opposed to using vertical-specific models with individual maintenance. They do however offer some precustomized models.

Google's STT seamlessly integrates with Google's other leading AI technologies for natural language understanding (NLU) and data processing such as Dialogflow, AutoML Natural Language, or Video Intelligence API.

Users that permit data logging to help improve speech models are offered a reduction in costs.

IBM

ibm.com

Headquarters: Armonk, New York

Products: Watson Speech to Text (Watson STT), IBM Watson Speech Services

IBM offers a broad set of AI-related capabilities under the Watson brand. The speech recognition functionality is in the Watson Speech to Text solution set. Watson STT is part of Watson's speech services which include Watson Assistant for Voice Interaction (WAVI), Watson API Kit for CloudPak for Data, IBM Watson Media and Watson speech APIs.

IBM Watson STT runs on- and off-premises, in private and public cloud environments and customers own their personal data. The solution works with real-time or uploaded batch files. The IBM Data Science and AI Elite team gives clients an accelerated track to implementation using tested, agile AI methodologies. IBM has professional services and partners globally. IBM is profiled in the Gartner report "Market Guide for Conversational Platforms."

Intelligent Voice

intelligentvoice.com

Headquarters: London, UK

Products: Intelligent Voice, IV Live, JumpTo

Intelligent Voice specializes in proactive compliance and e-discovery technology solutions for voice, video and other media. Intelligent Voice is in use across multiple clients in government agencies, banks, security firms, call centers, litigation support providers, and insurers. These include U.S. Department of Justice Antitrust Division and The Pensions Advisory Service, all involved in the management of risk and the meeting of multijurisdictional regulation.

Additional platform capabilities include biometrics, encrypted search, NVIDIA partnership for GPU-accelerated STT as well as edge deployments.

Intelligent Voice supports 17 languages with a variety of dialects across English, Spanish, Catalan, German and Portuguese.

Intelligent Voice's partners include Hewlett Packard Enterprise (HPE), NVIDIA, Zendesk, ZyLab, Proofpoint and Relativity.

Microsoft

microsoft.com

Headquarters: Redmond, Washington

Products: Azure Speech to Text

Microsofts STT service is a part of their broader Azure Cognitive Speech Services, also covering Speech Translation and Text to Speech. The API service is available to any app via a REST API. For easier adoption by developers, Microsoft also provides several client libraries to support integration with apps written in C#, Java, JavaScript and Objective-C. Developers can code applications to deliver recognition results in real time to users for either clarification or subsequent processing/communication.

Microsoft has rich support for creating custom speech recognition models to include speaking styles and domain-specific terminology. These custom models have an additional fee.

Uniquely, Microsoft can automatically generate custom models using an end user's Office 365 data to optimize speech recognition accuracy for specific organization-dependent terms.

A recent innovation is the Microsoft Conversation Transcription service that can improve the transcription from live gatherings using three speakers on separate smartphones or laptops. Microsoft has also added support for a speaker-verification service that confirms the identity of speakers based on their voice.

Nuance

nuance.co.uk

Headquarters: Burlington, Massachusetts

Products: The Nuance Intelligent Engagement Platform offers STT in various ways: as part of broader integrated virtual assistant, analytics and security solutions for text and phone channels, as a containerized microservice engine available via Conversational AI Services, and through their Speech Suite 11.

Nuance is a long-standing leader in speech recognition technology and solutions. Through the 90s and 2000s, through internal R&D and acquisitions, Nuance became the leading speech recognition provider in multiple markets, including contact centers and transcription. Since then, it has continued to expand its portfolio and technology, focusing many of its products on customer service, and building out security and biometrics functionalities. Nuance is also reviewed in the Gartner report "Market Guide for Virtual Customer Assistants."

Nuance stand-alone speech-to-text offers include: Nuance Speech Suite, Conversational AI Services for STT, and Nuance Transcription Engine. Nuance offers its own professional services and also has many service partnerships. The Nuance solution integrates with all leading contact center platforms. Elements of the Nuance solution are available either as cloud or on-premises; other parts are only available as a cloud offering.

Omilia

omilia.com

Headquarters: Cyprus, EMEA

Products: Omilia, deepASR (DNN Powered Automatic Speech Recognition Engine), Omilia Cloud Platform (OCP) Transcription Service (Omilia Cloud Platform cloud-based transcription service), OCP Conversational ASR Service (Omilia Cloud Platform cloud-based ASR service, optimized for IVR virtual assistants)

Omilia focuses on customer care experiences and provides both STT and conversational AI solutions that can be used independently or together, enabling end-to-end conversational experience across multiple channels.

Omilia provides off-the-shelf pretrained models for banking, insurance, telco and healthcare industries and supports the development of custom models. They have 40 customers to date and among them 27 production-grade contact centers globally. Omilia has partnerships with NICE inContact, Genesys, Concentrix, Connex, Talkdesk, Serenova, among others.

The platform supports 21 languages with dialect variations for English (Jamaica, South Africa, Canada) and French (Canada), Spanish (Spain, Latin America). Additional features include behavioral analysis, biometrics, and compliance. Omilia has mature experience in creating wider conversational applications spanning the use of STT, IVR and virtual assistant technologies especially within the financial sector.

Sestek

sestek.com

Headquarters: Istanbul, Turkey

Products: Sestek Speech Recognition, Conversational AI and Analytics, Conversational Biometrics, FreeTalk (medical)

Sestek provides a full suite of speech and conversational services specializing in banking, with particular focus on STT deployments in legacy IVRs.

Sestek supports 17 languages with a unique focus in Arabic dialects (Levantine, Gulf [KSA, Kuwait, UAE, Qatar, Bahrein, Iraq, Oman), Egyptian Arabic). It also supports dialects for English (US, U.K., Indian) and French (metropolitan, Belgian). Other capabilities include: biometrics, behavioral analytics, monitoring/compliance, and dual language support.

SoundHound

soundhound.com

Headquarters: Santa Clara, California

Products: Houndify, Hound

SoundHound offers both consumer- and developer-facing speech and sound products. The Hound consumer product is a mobile-device-based voice assistant that provides users with answers to their inquiries. The Houndify solution is a cloud-based voice AI platform service. The company also offers SoundHound which enables people to discover and share music.

The SoundHound features that leverage its STT include speech recognition, speech intent identification and NLU, audio and music identification, custom trigger phrase, multilingual TTS and customer commands. SoundHound has multiple contracts where it is incorporated as the solution for hands-free control, for functions such as automotive navigation, phone calls and stock updates. Customers incorporating this capability include Mercedes-Benz, Pandora, Hyundai and Honda, among others.

Speechmatics

speechmatics.com

Headquarters: Cambridge, UK

Products: Automatic speech recognition technology

The Speechmatics any-context speech recognition engine is used across multiple industries including contact centers, financial services, and broadcast providers. Features include advanced punctuation, custom dictionaries, speaker and channel diarization. The real-time transcription operates best at a latency of less than one second. The STT technology operates at the word and sentence level and so it can take in the context even when audio is unclear. The company provides accuracy comparisons which are publicly available on their website.

Spitch

spitch.ch

Headquarters: Zurich, Switzerland

Products: CodyFi, VeryFi, SentyFi, SignyFi, Lingware Suite, Dialogue Composer, Data Curator

Spitch was selected as a Gartner Cool Vendor for speech and natural language in 2019. The company provides an omnichannel conversational platform primarily targeted at the financial services sector (banking and insurance), with a unique voice-to-voice dialogue system. Its VeryFi voice biometric platform uses deep neural networks and transfer learning to improve both passive and hybrid identification/authentication, benefiting the customer experience of tasks, such as payments and onboarding.

Its key STT products, CodyFi and SignyFi, utilize verbal metadata, e.g., suprasegmental (stress, tone, intonation) and emotion — to increase speech recognition accuracy rates and provide a unique voice-to-intent solution. Spitch's sentiment analysis is cascaded, which means "chunks" of conversation are analyzed for sentiment and then semantically interpreted as a whole. This differentiated feature renders a more fine-grained analysis of how a conversation progresses and completes.

Spitch's core competencies include speech IVR, outbound calling, voice analytics, development kits for speech applications (e.g., apps), voice-first omnichannel chatbots, semantic call routing and other speech services in different European and North American languages.

Verint Systems

verint.com

Headquarters: Huntington, New York

Products: Verint Speech Analytics, Verint Text Analytics, and Verint Interaction Analytics.

Verint Systems focuses on multiple products and services that leverage its speech, text, and analytics roots. These include security, surveillance, business intelligence and customer engagement.

Verint's flagship analytics product is the Verint Interaction Analytics, alongside Verint Speech Analytics and Verint Text Analytics. Gartner also reviews Verint's Engagement Management platform in the "Magic Quadrant for the CRM Customer Engagement Center" and the "Magic Quadrant for Workforce Engagement Management."

Verint offers a number of advanced speech-related capabilities as well. Their Verint Speech Transcription and their Verint Automated Quality Management (AQM) are both based on the STT capabilities. Verint offers security functions such as both passive and active voice biometrics, end-to-end encryption, and multiple data confidentiality functions. Verint offers cloud and hybrid deployment options.

VoiceBase

voicebase.com

Headquarters: San Francisco, California

Products: The VoiceBase API product enables end-to-end controls for ingesting, transcribing, analyzing, and reporting on the data.

VoiceBase has an extensive partner network and offers prebuilt integrations with Amazon Connect, Genesys, Tableau, and Twilio. Analytics service supports data output to Snowflake, Redshift, Azure SQL Server, and Exasol. VoiceBase's cloud-based transcription approach is primarily word-based which allows it to perform cost-effective rapid transcription projects. However, this makes it less effective at transcriptions where the full sentence context is critical.

The VoiceBase solution is cloud-based. However, data is encrypted at rest and in flight, and the solution offers data confidentiality and protection. Use cases for VoiceBase include: customer service speech analytics, sales agent coaching and top agent modeling, compliance via PCI and personally identifiable information (PII) redaction, and scoring of interactions.

Market Recommendations

Buyers should select vendors that meet not just their immediate use cases, but should consider the likely pipeline of speech-powered solutions over the next 12 months (IVR, biometric authentication, chatbots etc.). Use Table 2 to consider your deployment type and features needed.

Given the rate of development of this technology, buyers should avoid lengthy contract tie-ins and ensure that language and acoustic models they have developed will remain an asset should the vendor no longer be viable. This will allow them to retrain or bootstrap another solution reducing technical debt risk. More broadly, consider STT solutions as part of a wider natural language technology portfolio where language assets and models are able to be shared among implementations such as chatbots, knowledge management, text mining and of course speech.

Ensure you understand what will be needed to reach an optimum level of performance for your industry and budget for setting up human-in-the-loop processes and UIs to enable ongoing curation of the speech experience. A strong analytics and supervised learning loop is essential for improving the implementation over time; any product without it should not be considered.

Acronym Key and Glossary Terms

AQM	automated quality management
ASR	automatic speech recognition
CX	customer experience
CRM	customer relationship management
HIPAA	Health Insurance Portability and Accountability Act
IoT	Internet of Things
IVR	interactive voice response
KPIs	key performance indicators
NLG	natural language generation
NLP	natural language processing
NLT	natural language technology
NLU	natural language understanding
PII	personally identifiable information
SaaS	software as a service
SBC	single board computers
SDK	software development kit
STT	speech to text
TTS	text to speech
VAS	value-added services
VPA	virtual personal assistant

Gartner Recommended Reading

Some documents may not be available as part of your current Gartner subscription.

“Using Conversational AI Middleware to Build Chatbots and Virtual Assistants”

“Magic Quadrant for Cloud AI Developer Services”

“Evolving IVRs to Conversational Platforms — Critical Organizational Issues”

“Evolving IVRs to Conversational Platforms — Critical Technology Issues”

“Evolving IVRs to Conversational Platforms — Critical Leadership Issues”

“Evolving IVRs to Conversational Platforms — Intent Modeling Issues”

Note 1 Representative Vendor Selection

The inclusion criteria below were used for longlisting vendors in this report:

- Have a native speech-to-text engine (rather than dependent on a third-party engine). Without this criterion it was automatic exclusion.
- Can take input from a number of formats (e.g., audio, video).
- Handle multiple languages and improve solution accuracy through the use of custom dictionaries and acoustic models.
- Provide analytics on speech content.
- Can deploy the solution across a variety of channels, e.g., IVR, mobile apps, virtual assistants, embedded.
- Provide transcriptions of speech content.

This Market Guide is not a qualitative scoring of vendors against one another. Vendors in this report are illustrative of aspects of the market that we deem to be important. Readers should use this report to understand the current market dynamics and the evolution of this market.

Note 2 Gartner’s Initial Market Coverage

This Market Guide provides Gartner’s initial coverage of the market and focuses on the market definition, rationale for the market and market dynamics.

GARTNER HEADQUARTERS

Corporate Headquarters
56 Top Gallant Road
Stamford, CT 06902-7700
USA
+1 203 964 0096

Regional Headquarters
AUSTRALIA
BRAZIL
JAPAN
UNITED KINGDOM

For a complete list of worldwide locations,
visit http://www.gartner.com/technology/about.jsp

This publication may not be reproduced or distributed in any form without Gartner's prior written permission. It consists of the opinions of Gartner's research organization, which should not be construed as statements of fact. While the information contained in this publication has been obtained from sources believed to be reliable, Gartner disclaims all warranties as to the accuracy, completeness or adequacy of such information. Although Gartner research may address legal and financial issues, Gartner does not provide legal or investment advice and its research should not be construed or used as such. Your access and use of this publication are governed by Gartner Usage Policy. Gartner prides itself on its reputation for independence and objectivity. Its research is produced independently by its research organization without input or influence from any third party. For further information, see "Guiding Principles on Independence and Objectivity."