SCALE THROUGH AUTOMATION

Market Analysis

To date, the market can be categorized as having three main types of vendor.

Vendors and Categories

Vendors and Categories

Self Serve and Community	Platform and Services	Customer Engagement Suite
Self Serve API Giants • Google • Amazon Web Services • Microsoft • Tencent • IBM **Open Source** CMUSphinx	**Speech Platform and VAS** • Almawave • Cedat 85 • Speechmatics • Spitch • Intelligent Voice • VoiceBase • AISpeech • Yactraq **Broader NLP Suites and Services** • Nuance • SoundHound • Omilia • Sestek	**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 Amazon, Google, IBM, Microsoft and Tencent (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.