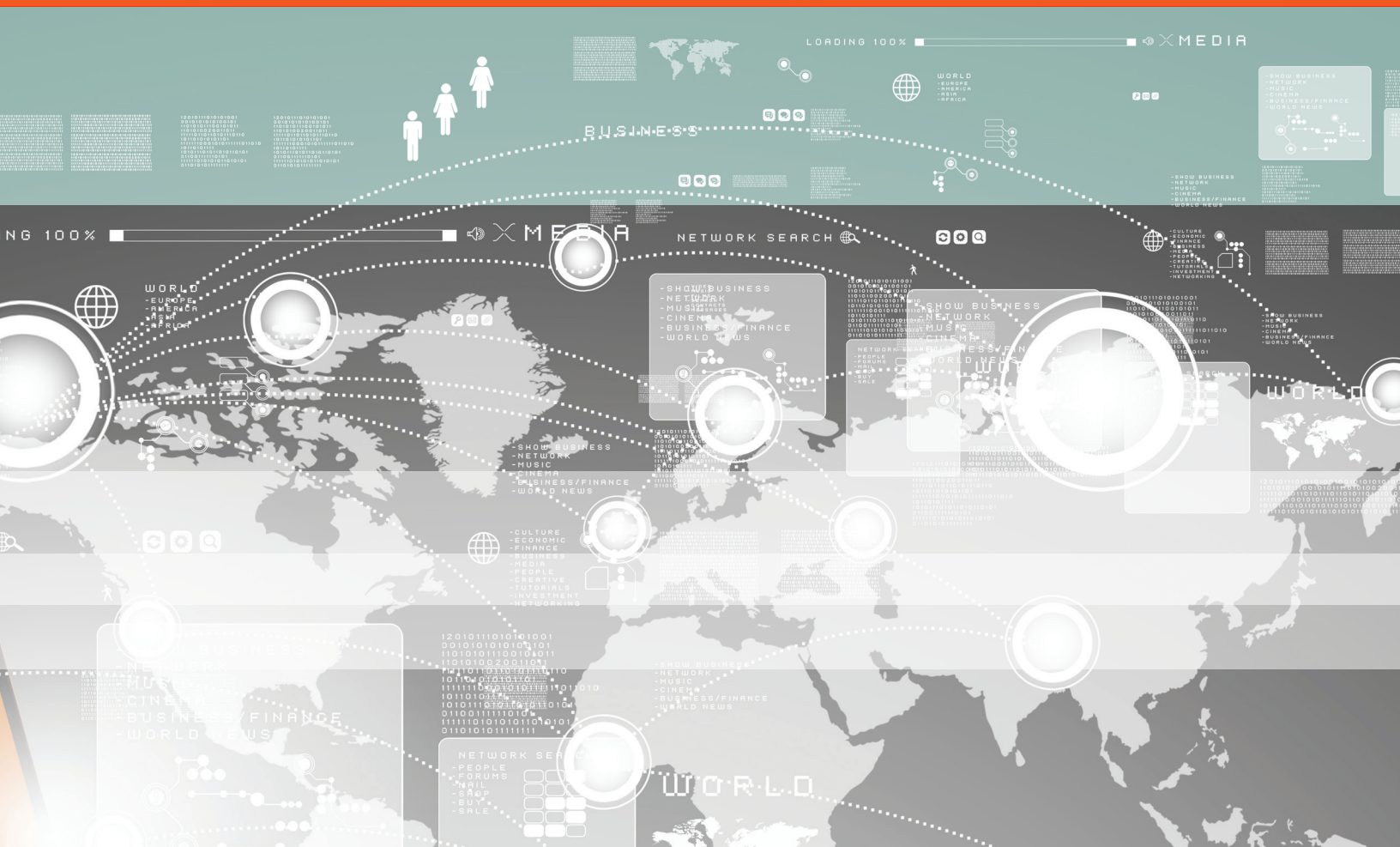
Conversational Voice Response:

Bringing the Best Tech Forward

opusresearch



opusresearch | Report »

Conversational Voice Response:

Bringing the Best Tech Forward »

Now that artificial intelligence is a religion and we all worship at the altar of Big Data and analytics, it is heresy to mention that the foundations for conversational, cognitive interactions between brands and their customers often reside in the switch closet (or cloud) that houses your Interactive Voice Response (IVR) systems.

»

June 2018

Dan Miller, Lead Analyst & Founder, Opus Research

Opus Research, Inc.
350 Brannan St., Suite 340
San Francisco, CA 94107

www.opusresearch.net

Figure 1: Conversational Evolution

"IVR": An abbreviation known to send chills down the spines of inbound callers as they anticipate that their supple entry of digits on a keypad will be followed by an hour of "music on hold," while waiting to be routed to the wrong department or an ineffectual CSR (the TLA for customer service representative). As counter measures, these callers reflexively (and repeatedly) press zero, or loudly say "Agent!" while clasping their smartphones in a white-knuckled chokehold.

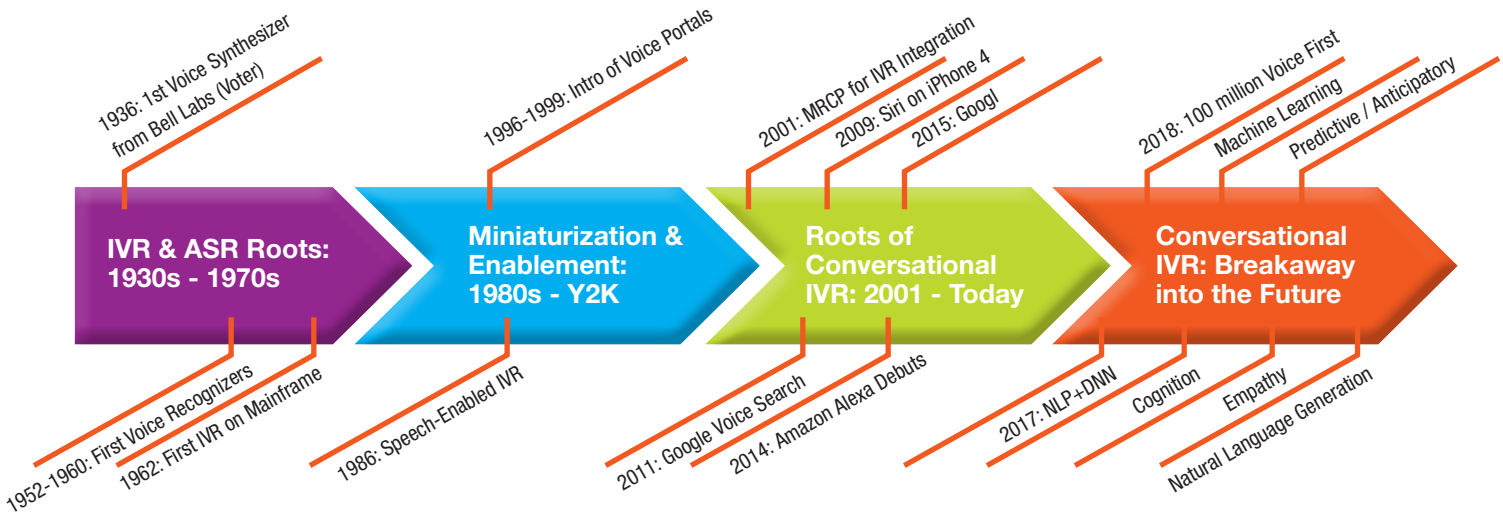
In this modern metaphor, an IVR is the obligatory, yet unpleasant, stopping point on each customer's journey from search and discovery to fulfillment. Its staying power within a company's IT and communications infrastructure is due to its proven performance in managing large incoming call queues and documented cost-effectiveness, resulting from efficient management of labor costs.

The image of the IVR as a technological backwater is an inaccurate trope. Instead, it should be seen as the technology platform that spawned development and implementations of the core technologies that make today's "Voice-First Personal Assistants" possible. Siri, Alexa, Google Assistant and others would not be where they are today were it not for the much-maligned IVR systems that reside in the critical path of nearly every voice conversation between companies and their customers.

Specifically:

- **Automated Speech Recognition (ASR)** learned how to play well with Neural Networking, Intent Recognition, and Natural Language Processing as they played their role of in speech enabling IVRs in the last century.
- **Text-to-Speech (TTS) and voice synthesis** capabilities also evolved from highly robotic sounding output to extremely lifelike personas while handling billions of calls through speech enabled IVRs.
- **Scripting languages (like VoiceXML) and "dialogue modules"** were cultivated for IVRs but came to greater fruition as the "stock phrases" and other raw material for the latest generation of developers to build chatbots and other virtual agents.
- **Tools** for developing, deploying, monitoring and tweaking intelligent assistants, likewise have their roots in the IVR domain.

Figure 2: Accelerating Pace of Conversational IVR



This timeline dramatizes the long legacy of IVR and the accelerated pace of innovation as speech processing technologies joined forces with advances in search, analytics and Deep Neural Networking (DNN) to to reach breakaway velocity in the 2017 time frame. The IVR was the platform where both technology and technologists honed their conversational skills. In the early days of speech-enablement (meaning the late1980s), they made the most of Neural Networks and Statistical Language Models, including the mysterious Hidden Markov Model (HMM - look it up).

Today, heralding the breakaway from IVR to CVR, plain vanilla neural networks have been replaced with Deep Neural Networks, machine learning is being augmented by cognitive resources and the introduction of new virtual agents and personas is being accelerated by deploying methodologies that incorporate Big Data, analytics, and cloud-based architectures. Yet many of the key technological components will be carried forward and enhanced by more modern CVR platforms.

Keeping Up With the Optichannel Engagement Model

The value of the IVR in an omnichannel or, more accurately, optichannel world has been called into question because they are generally defined by their core function, which has been centered on “voice” (the “V word” in IVR). Late in 2016, [24]7.ai published results of research that culminated in its seminal “Customer Engagement Index.” It compiled data that shows stark but unsurprising differences between the behavior among Millennials, Gen Xers, Baby Boomers and The Greatest Generation. While there were differences in the level of digital device ownership, like smartphones and tablets, there were commonalities in the sequence of channels individuals pursued when they were shopping or carrying out other forms of digital commerce.

The survey revealed that, while customer journeys most often start on a company’s website (which are often on mobile web), nearly a third of them turn to their phones to call a company if their initial efforts are unsuccessful.

What's more, when their journeys take them one step further along the digital path, (for instance if they fail to resole an issue through a website or over the phone), they will use the phone channel with the specific mission of talking to a live agent. At that point in the customer's journey, the IVR is like a bouncer in front of a popular nightspot. It is all that stands between a caller and a conversation with an agent that he or she seeks. No wonder they are characterized as rigid, ineffective, unresponsive or worse.

In an age of "digital transformation" and "optichannel communications" the IVR has languished while brands are investing in Alexa skills, text-based chat bots and even avatars.

Opus Research has long observed that self-service solutions must support customers "using their device-of-choice, over their channel-of-choice at their time-of-choice." IVRs have been evaluated and, in general, management has not deemed to make the necessary upgrades to bring them into the digital age. But this is a mistake. In addition to being the proving ground for ASR, TTS, NLP, MT and conversational interactions, IVRs were also the platforms on which companies developed new approaches to digital commerce that combined voice with text and other visual content (under variations of the term "visual IVR") all performed in real-time and in the context of providing a better experience for individual customers or users.

Augmenting basic IVRs with the latest technologies creates a new category of self-service platform: Conversational Voice Response (CVR).

Conversational Voice Response Fills a Much Needed Role

CVR is making its appearance on the conversational commerce landscape just in time to meet new challenges created by the ubiquitous presence of voice-first devices. It starts with about 2.5 billion smartphones globally, and now adds voice-first smart speakers in over 50 million households, with the likes of Amazon and Google, as well as car manufacturers like Mercedes-Benz, BMW and Ford defining how every one of the hundreds of millions of cars on the road is outfitted to support voice-activated services.

Note that these devices are called "voice-first" or "voice-activated," not "voice-only." That's because the conversations they support may start with a wake-up word or a spoken question and they have the power to derive intent or recognize the meaning when an individual speaks in his or her own words. Yet a growing number of these devices are physically connected or associated with screens. Smartphones and tablets are touch-sensitive, flat screens with built-in microphones that support ASR and speakers that support lifelike TTS. Cars are equipped with both microphones and touch-sensitive screens. Five short months after Amazon introduced the Echo smart speaker, the Echo Show made its debut, demonstrating the value of full on video closely mated to a CVR. It doesn't take long to see how speech-enabled remote controls on smart TVs will rise to support a visual CVR model.

As these services evolve and become routinely accepted by end users, it is important to remember that the basic IVR was the original development run-time platform for conversational interactions. The scripts that were originally written to support such conversations and were refined over the years to remain accurate and *au courant* are now the raw material that inform the "stock answers" that an Virtual Personal Assistants provide.

The rules that govern when and how to answer queries or for routing calls to other resources are also of great value. There is no need to reinvent this wheel.

Support Optichannel Engagement Models Today and Tomorrow

When customers or prospects pick up the phone to reach a company, their issues are urgent. They've already tried one or two other sources and they want rapid response or action. Old ways die hard and, when callers detect that they are in contact with an IVR, they may immediately and ritualistically start seeking contact with an agent. In many cases, this might be the right action on their part. In spite of the advancements in speech recognition, text-to-speech, natural language processing, machine learning and predictive analytics, the vast majority of "old-guard" IVRs are not speech-enabled and are limited in function.

Over the years, IVRs have proven to be highly reliable and, therefore untouchable sets of resources. Architecturally, today's IVR could not be better positioned. It is the focal point for every voice-based contact, yet it is no longer the bouncer that everyone wants to get around. Instead, it is a combination of triage specialist and concierge. It is programmed to engage in conversation immediately (no music on hold), understand natural language input and extract the purpose of the call (based on decades of successful interactions), and determine whether it can resolve the issue itself (based on its knowledge base and confidence that it has the answer).

If it is not able to resolve the issue immediately, it can use its knowledge to determine what resource, including live agents, has the highest probability of providing the proper answer or taking the correct action.

Recognizing the Context Inside Each Journey

Traditional IVRs would have had the ability to tailor responses based on automated number identification (ANI). CVRs augment basic ANI with factors such as information about prior searches on a company website, location, "device profiles" (indicating the screen type and operating system of the callers device) and other relevant factors. This helps to tailor responses that are personal in nature and more precisely tuned to the customer's intent.

Ingesting Data and Metadata from "Everywhere"

Speech-enabled IVRs often sound very lifelike, but the range of their responses is the product of orchestrated

scripts that craft their end of the conversation from recorded voices, synthesized voices according to instruction sets or "rules." Conversations get broken when callers go out-of-bounds. The rules don't apply. And no voice file or text output is appropriate.

Big Data by its very nature is almost anything that is machine-readable. It is often described as a "combination of structured and unstructured data", meaning that it can be read from columnar databases, like product catalogues or price lists, or it may be free form like the transcriptions of conversations between agents and customers. What has been most important lately has been the ability for a CVR, playing the role of Intelligent Assistant, to find the precise answer that a caller seeks from inside unstructured data. Ingesting voluminous amounts of data did not always equate to delivering precise answers, but now it can.

Incorporating Human Input and Recognizing Human Intent

Amid long-standing concern that Artificial Intelligence will take the place of human intelligence and that virtual agents will take the place of real agents, CVRs provide ongoing roles for humans. Once, Voice of the Customer programs invited callers to say whether their issues were resolved properly. In the CVR world, other humans are the source of the best answers and there are programmatic ways to incorporate human feedback to identify the best answers to future questions.

Continuous Learning and Improvement – Supervised, of course

Best practices indicate that human input is also a pre-requisite for training Intelligent Assistants, including CRVs. Algorithms can be led astray by malicious end-users and companies must assign some of their most knowledgeable employees to monitor responses, often in real time, to provide callers with the best answer.

A Tactic for Keeping Up With The Pace of Change

We've entered an era when the refresh cycles for enterprise IT and contact center infrastructures mimic those for smartphone apps. Meanwhile, both customers and prospects are initiating voice-first conversations without picking up a phone and dialing a phone number. CVRs answer the challenges posed by high-volumes of voice-based conversations by invoking resources that anticipate the purpose of that call and render audio responses in real-time that combine the latest, dynamic data from a variety of sources along with scripts or transcripts that are the product of time-tested, successful interactions.

Enterprises cannot and should not just walk away from their IVRs. Taking a CVR approach enables them to weave their best aspects and attributes into omnichannel conversations. While consumers increasingly interact with Alexa and Google Assistant for simple, common tasks, and use messaging platforms to chat with friends and companies, it's imperative for enterprises to maintain that customer relationship and keep control of their brand.

With a CVR approach, enterprises have the ability to build conversational services that anticipate, recognize and respond to each customer's intent in a voice that is consistent with each brand's image and, of course, "Voice".

About Opus Research

Opus Research is a research-based advisory firm providing critical insight and analysis of enterprise implementations of software and services that support multimodal customer care and employee mobility strategies. Opus Research calls this market "Conversational Commerce" with tailored coverage and sector analysis that includes: Self-Service & Assisted Self-Service, Voice & Call Processing, Web Services, Personal Virtual Assistance, Mobile Search and Commerce and Voice Biometrics.

For sales inquires please e-mail info@opusresearch.net or call +1(415) 904-7666