- Whitepaper

Bringing Intelligence to Voice Bots

to improve Customer Experience

clickworker

Contents

Introduction to the Power of AI-Powered Chatbots

Chatbots—both textbots and voicebots—are here to stay because customers like them, and they can slash the cost of support for organizations. AI-powered chatbots, also known as conversational agents, are transforming the way organizations and their users communicate.

Consider these data points:

- ***80% of the 265 Billion Customer Service Calls Could be Handled by Chatbots.***
  According to IBM, every year, 265 billion customer service calls are made, and businesses spend $1.3 trillion on servicing them, and 80% of them could be served by AI chatbots without human involvement.[1]

- ***$142 Billion in Sales via Chatbots.***
  By 2024 consumer retail spend via chatbots worldwide will reach $142 billion, according to Insider Intelligence.[2]

- ***Conversational AI Market to Nearly Triple in Size.***
  The conversational AI market is projected to grow from a current $6.8 billion to more than $18 billion over the next four years—with small and medium-sized companies representing the greatest drivers of growth, according to a report published in Forbes.[3]

Preventing the Chatbot "Spiral of Misery"

All of this means that the future is bright, very bright, for organizations seeking to tap into the immense power and value of AI-powered chatbots. However . . . organizations must exercise diligence in creating their chatbot deployments to avoid what the New York Times has called the chatbot "spiral of misery." [4]

Preventing the spiral of misery is especially important because increasingly consumers prefer self-service solutions. "Two-thirds of customers believe service through online channels and mobile devices should be faster, more intuitive, and better able to serve their needs. Organizations should seize the opportunity with improved front-end robotics or "virtual agents" to handle repetitive, transactional requests as well as to guide customers through a logical menu of topics and intentions to address issues." according to the consultancy McKinsey.[5]

The Good News

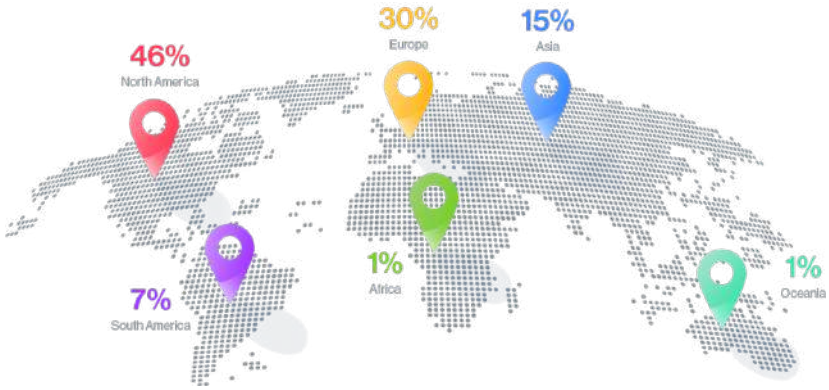
Organizations can reduce the cost of customer service, enhance the quality of customer interactions, and enable automated sales through use of quality AI-powered chatbots—both voicebots and textbots.

The McKinsey report found that "Companies that have incorporated such technologies are seeing significant returns: in fact, effectively deploying conversational AI can create a twofold improvement in customer experience; reduce cost to serve by 15 to 20 percent; improve churn, upsell, and acquisition by 10 to 15 percent; and result in a fourfold increase in employee productivity."

The Bad News

Poorly executed chatbots can deeply frustrate users, creating negative experiences that can tarnish a brand and drive away customers.

Clickworkers
We have 6 million Clickworkers spread across the globe.



Needed: Robust Training Data for AI

For an AI-powered chatbot to function well in the real world, it requires real-world data for training. The more precise and reality-based the data, the more successfully the algorithms driving the AI can be trained, tested, and perfected. Many AI projects have foundered because of a lack of diverse, real-world training data.

Clickworker Value Proposition: Creating Real-World Training Data for AI

Clickworker brings the human touch to AI training. With over 6 million Clickworkers spread across the globe, we are ready to help organizations get more out of their algorithms by generating, labeling, and validating unique AI datasets, specifically tailored to needs as well as to providing a solution for analyzing an AI's output results.

Whatever scenarios are required for training an AI algorithm, our Clickworkers can formulate a range of user problem cases—capturing the diverse manners in which real people may try to explain a problem or react to a suggestion. We can assemble a pool of Clickworkers to meet demographic, geographic, language, cultural, and other needs.

Taking a Closer Look at Voicebots

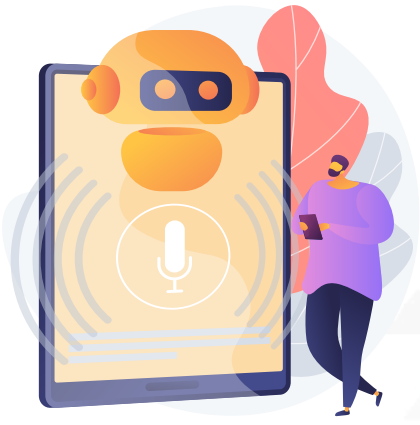
Considerable AI is required to create either a robust textbot or voicebot, with special effort required to achieve the second. Both textbots and voicebots have a similar task: parsing input, formulating an appropriate response, and delivering the response to the user.

This is a great challenge for both textbots and voicebots, as the task involves natural language processing (NLP), which enables an AI application to decipher—some might say understand—the input, and then create and deliver a logical response.

The Challenge for Voicebots

For voicebots, this challenge is considerably greater, as it takes us into the surprisingly complex realm of voice recognition. This involves translating the analog sound of voice to digital—which includes breaking sound waves into distinct sound blocks called phonemes.

The challenge then turns to assembling phonemes into words, figuring out where one word ends and the next begins, and validating that the stream of assembled words make sense contextually.



A succinct example of this challenge is provided in a paper from Institute of Acoustics Chinese Academy of Sciences, which reads in part:

"This process is even more complicated for phrases and sentences—the system has to figure out where each wovrd stops and starts. The classic example is the phrase "recognize speech," which sounds a lot like "wreck a nice beach" when you say it very quickly. The program has to analyze the phonemes using the phrase that came before it in order to get it right.

Here's a breakdown of the two phrases:

r eh k ao g n ay z s p iy ch "recognize speech"	r eh k ay n ay s b iy ch "wreck a nice beach"

Why is this so complicated?

If a program has a vocabulary of 60,000 words (common in today's programs), a sequence of three words could be any of 216 trillion possibilities. Obviously, even the most powerful computer can't search through all of them without some help."

While it is beyond the scope of this white paper to dive into the mathematics and code of NLP, it is critical to stress the importance of providing robust training data so the NLP can work with the very broad spectrum of human voices, dialects, vocabularies, speech patterns, and even manner of thinking.

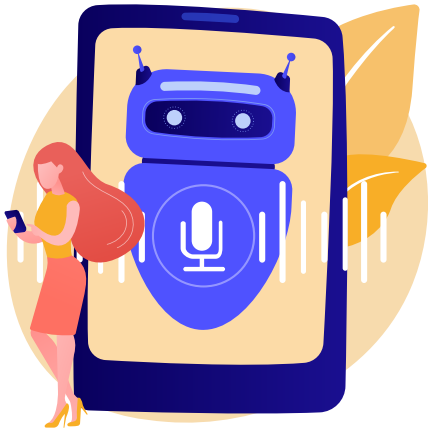
For any given task, how will the way one person phrase a question, compare to another? Will they say: "I need help" . . . or . . . "Need your assistance" . . . or . . . "How is this supposed to work?" . . . "This ain't working" . . . or . . . "Why haven't you fired your interface designer?"

Apart from that, and a world of other variations in how inputs might be phrased, how well will your program be able to handle voice queries from English speakers from Korea or Norway, or France, or Nigeria, or Vietnam, or from rural Kentucky? Add to that, gender, age, and variations in speech, as well as level of education and social experience.

The potential variations in vocabulary, phrasing, enunciation, dialect, and other variables abound. The degree to which such elements can be captured and put to use as training data to a large degree determines the intelligence the AI-powered voicebot can actually show.

Worth the Effort: Many Users Prefer Voicebots to Textbots

While it is easier to create a textbot, many users prefer interacting with a voicebot because speech is a more efficient and natural way of communication. No typing required.



"The existing research has shown that, in general, humanized chatbots benefit companies," Dr. Cammy Crolic, Associate Professor of Marketing at the Saïd Business School at the University of Oxford, told The Wall Street Journal. "Humanized chatbots have been shown to be more persuasive, increase enjoyment and provide the added benefit of social presence. Consumers are more likely to trust humanlike technology interfaces because they believe them to be more competent and less prone to violations of trust."

The Risk of Disappointing Users with Inefficient AITextbots

The risk of creating a humanized voicebot experience, is that users tend to expect more from a voicebot than a textbot. The more human you make the AI experience, the higher are the performance expectations from the user.

Dr. Crolic's research reported in The Wall Street Journal found high customer satisfaction when voicebots were able to handle common tasks such as balances, transactions, and other account information. Yet the researchers found that disappointment can be amplified when the voicebot falls short.
"Often humanized chatbots are simply unable to meet expectations, which leads to disappointment," said Dr. Crolic. "Both angry and nonangry customers feel the disappointment of unmet expectations, but angry customers are more sensitive to this disappointment and prone to act on it. They hold the humanized chatbot more responsible, respond aggressively, and "punish" the chatbot and the company it represents through lower ratings and plans to purchase."

Incorporating robust training data from real-world people from across the globe and across the spectrum of human variations in vocalization, vocabulary, language background, social and democratic attributes can help ensure that voicebots maximize the user experience.

How Clickworker Creates Great AI Training Data

As AI systems, voicebots can only come alive when they are given powerful, real-world datasets that reflect the human-centered environment in which they will operate. Data scientists have long warned "garbage in, garbage out" when it comes to analytics. This sentiment is even more true when training the AI systems that power voicebots.

This means going beyond simply feeding in sound files. The data flowing into these system must go beyond phonemes to capture the reality of human reason and human behavior—the quirky ways in which we all differ in how we speak and think.

Capturing the Rich Diversity of Human Interactions

Clickworker captures the rich diversity of human interactions through its global Clickworker community. According to the needs of our clients, we can produce precision sound files of how thousands of users from around the globe would utter a command or ask a question—including open-ended questions such as: "What would you say if you needed this . . . or didn't understand that?"

Rich Diversity
Our global community of Clickworkers provides the rich diversity required to create powerful AI training data.

Every voice recording differs—even in the same language—due to the individual choice of words, the word order, as well as every single Clickworker's specific pronunciation. The result is a wealth of real-world training data.

Some of our clients make use of more than 100,000 instances of our Clickworkers interacting with a system. Many of our clients regularly return for additional training data as they continue to fine tune their AI products or need new data to test added functions and features.

Proven Workflows

We work closely with our clients to define the needs for each project, and then put our Clickworkers to work using our well-proven workflows. For example, one client required voice training data for a new AI application that would be distributed in nine countries.

The client identified 150 training scenarios to be captured 600 times for each of the nine languages, resulting in 810,000 sound recordings, captured in MP4 format.

We distributed the task to Clickworkers in the nine specified countries.
All used the same three-part workflow:

- **Task 1: Creation of Voice Recordings**

  Create the audio recordings

  Transcribe the recordings

  Analyze and evaluate the recordings, including identifying keywords

- **Task 2: Quality Assurance**

  Pass the results to a second Clickworker who performed the same validation processes for quality control

  In this case, it involved checking the transcriptions of 810,000 voice recordings made by native speakers

- **Task 3: Analysis and Evaluation**

  Calculation of the keywords and their frequency per scenario and language

  Filtering the phrases including frequency per scenario and language

From here the results are transmitted to the client for training. Transfers are often made daily as the work continues, depending on client needs.

Providing Real-World Solutions

Organizations make use of our global community of Clickworkers to aggregate high-quality AI training data for a spectrum of needs. In this section we'll look briefly at some of our examples.

Needed: Training Data for an In-Home Emergency Systemtions

One of our clients created an in-home emergency system to detect and respond to cries for help from the elderly (and others in need) but can't reach a phone. The device market included the U.S. Germany, France, Spain, and Portugal.

- **The Need:**

  Data to train the device to automatically call an ambulance whenever a call for help was detected.

- **Clickworker Solution:**

  Involved Clickworker community members between the ages of 18 to 65 from each of the target countries to record variations on a cry for help. More than 2,000 voice recordings were captured, incorporating client-specified technical requirements for background noise and other variables. Recordings were uploaded daily to the client. Clickworker community diversity included divergent ages, genders, and accents.

Needed: Training Data for Voice Biometrics and Sound Identification

A maker of medical devices required AI training data for its biometric solution that enables voice activation and voice commands.

- **The Need:**

  Data to train the device to respond to voice commands. The client had created its own online sound collection device, but needed a large and diverse set of training data across a number of languages.

- **Clickworker Solution:**

  Involved Clickworker community members between the ages of 18 to 65, with the client-specified 50-50 gender mix. To incorporate all client-requested languages, we involved Clickworker members fluent in English (as spoken in the U.S., the UK, and Australia); Spanish (as spoken in Spain, the U.S., Mexico, and South America); French and Canadian French; German, Swiss German, and Austrian German; Portuguese, Brazilian Portuguese; and Italian.

  So far, we have delivered more than 60,000 voice recordings, with the happy client regularly returning for more. Throughout we have worked with the client to meet their technical specifications for voice recording, including excluding laptop microphones and anything that includes noise cancellation.

Needed: Training Data for Automotive Voice Interactions

A maker of automotive components needed training data for its AI application enabling humans to request information from, and deliver commands to, a vehicle's onboard computer systems.

- **The Need:**

  Training data for phrases such as "Where's the nearest gas station?" as well as a wide range of less predictable queries and some 5,000 control commands.

- **Clickworker Solution:**

  Involved Clickworker community to record 5,000 different commands, in 18 languages. More than 90,000 recordings were aggregated for the client, with results uploaded daily to them.

Summary

There is a growing demand for AI-powered voicebots because, when well crafted, many users prefer automated systems to dealing with a human, or having to wait for a human to be available to help. This is one reason why sales via chatbots are expected to hit $142 billion by 2024. And as IBM has noted, chatbots can handle an estimated 80% of the 265 billion customer service calls—which cost businesses about $1.3 trillion annually.

So, there is push from the consumer side and the business side to create great chatbots, with voicebots providing the most natural and efficient means of interaction. To avoid what The New York Times has called the "spiral of misery" that can chase away customers, robust AI is needed—which requires real-world training data that reflects how the great diversity of humanity around the world reasons and speaks.

We created Clickworker as a global organization to precisely meet these needs.

About Clickworker

Clickworker is a full-service provider and offers both standard and custom solutions for the implementation of data-oriented projects. These projects are automatically broken down into micro jobs and processed by qualified Clickworkers from the crowd. The results are then reassembled and transmitted to the customer in a quality-assured manner.

Clickworker also offers a self-service solution via the online marketplace for smaller and standardized tasks in the areas of text creation, surveys, and sentiment analysis. These practice-proven services offer reliable and high-quality results with high throughput, outstanding scalability and significant cost savings.

For more information, please visit us at www.clickworker.com.

[1]https://abdulquadri-oshoare.medium.com/how-chatbots-can-turbo-charge-your-business-7e9ef29823d8

[2]https://www.insiderintelligence.com/insights/chatbot-market-stats-trends

[3]https://www.forbes.com/sites/quickerbettertech/2021/12/01/on-crm-chatbots-are-becoming-a-10-billion-market/?sh=374e4098238b

[4]https://www.nytimes.com/2022/03/03/technology/ai-chatbot.html

[5]https://www.mckinsey.com/business-functions/operations/our-insights/technology-and-innovation-building-the-superhuman-agent

[6]http://english.ioa.cas.cn/psk/201302/t20130204_98876.html

[7]https://www.wsj.com/articles/chatbot-consumers-11637336456