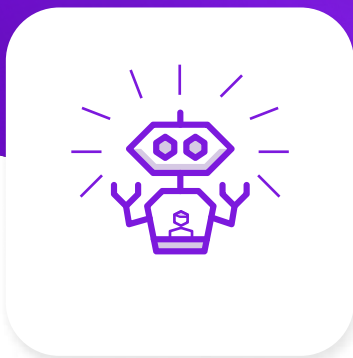# Conversational AI: From Hypothetical to Reality

Conversational AI: From Hypothetical to Reality

An Innovative Approach to Implementing Conversational Artifcial Intelligence

There is no shortage of hype around artificial intelligence (AI), Machine Learning (ML) and Chatbots. The goal of this paper is to discuss the current state of conversational bots, and provide a practical way for a company to leverage and grow into the new technology without sacrificing the customer experience.

NIRVANA

When executives read about conversational AI, they picture an AI engine that responds to every customer's request with a perfect, personalized, accurate message in real-time. This is an ideal scenario where every company can deploy AI engines that pass the "Turing Test", which is the default standard used to define artificial intelligence that can pass for being a real human.

CX ISSUES

While nearly all major technology companies (e.g. Google, IBM, Microsoft) are investing billions of dollars into AI engines to power the future of B2C conversation, the industry is still years from passing the Turing Test. This has led enterprises to develop the following:

Logical AI - This is essentially a form that is filled out and structured as a conversation. There is no natural language processing (NLP) involved. This is not real AI; it more closely resembles a modern web form.

Deutsche Bank   PSECU   ORION   MAPFRE   BERKSHIRE HATHAWAY   M 1st MEMBERS 1st FEDERAL CREDIT UNION

NLP - This is real AI; however, the failure rate is quite high. It is blatantly obvious every time this fails as it will say: "Please rephrase that", "Do you mind asking that in another way", "Hold on while I try to transfer you to an agent". This type of AI is currently so unreliable that there are reports of Facebook Messenger bots having a 70% failure rate.

A CONVERSATIONAL AI STRATEGY

While the current state of conversational AI has plenty of room for improvement, there is a practical approach that companies can take to leverage this technology and its benefits without

1	2	3	4
Humans	AI-Powered Humans	Human-Powered AI	Full AI
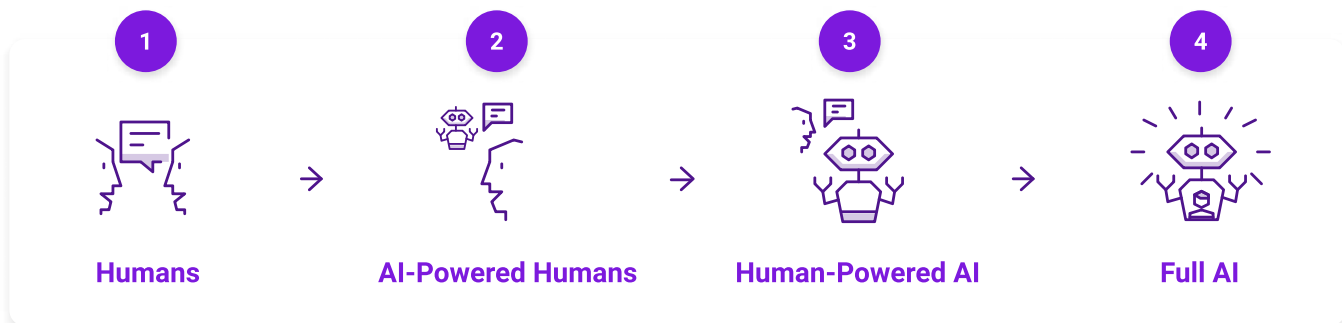
sacrificing the customer experience, as inevitably happens with the high failure rate of NLP. This approach is called the "AI Kaizen Launch Plan". Over time, this approach will allow a company to evolve with the technological advancements of conversational AI, and continue to provide an exceptional customer experience. That said, skipping any phase in this approach could have a negative impact on the customer experience.
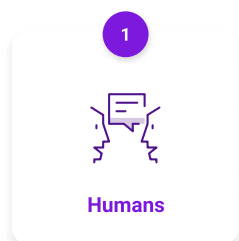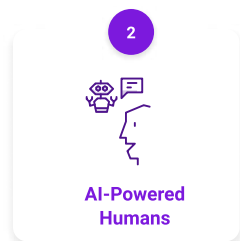
PHASE 1 – HUMANS

The first phase is centered around gathering conversational data. A company must understand what its customers are asking. With this data in hand you may be thinking that you're ready to feed the bot.

Humans

Doing so prematurely leads to the high failure rate and negative customer experience mentioned previously. Instead, use the insights gathered from synthesizing FAQs to empower your staff.

Don't Skip: Skipping this step will cause your agent experience in phase 2 to be negatively affected.

2

AI-Powered Humans

PHASE 2 – AI–POWERED HUMANS

The second phase exploits the learnings from phase 1 to super-charge your agents for operational efficiencies, but maintains a good experience for the end-customer. In this phase, Conversational AI achieves an accuracy level around 90%.

Part A: Take the 5-10 conversation topics that were gathered in phase 1 and use them to generate a list of simple keywords that would trigger preferred responses for your agents to use. Input these responses into your AI engine and point that engine's responses at your agents. Once you have a statistically relevant sample of real conversations, run an analysis to determine how often your agents used the recommended responses. From there, you will be able to tweak keywords and responses, add new ones, and run the process until agents are utilizing the responses over 50% of the time.

Part B: Expand from keywords to phrases. Phrases are important because they help the AI engine use context to distinguish between various intent. For example, if one of your keywords is "car", the phrases 'I have a car' or 'I don't have a car' will most likely require fundamentally different responses. Tweak the phrases and responses, add new ones, and run the process

Part C: Incorporate threaded conversations into the AI engine. The flow should mimic that of a verbal conversation, with several topics and important context. For example, if a customer is talking about

their computer and says it's broken, a human agent understands that the customer is referencing the computer. This is more difficult for an AI engine, like Amazon Alexa, to comprehend. If you were to ask Alexa "When is the next Yankees game?" and follow up with "How about the Mets?", you'll likely get "I'm sorry, I don't understand the question" in response, because Alexa is unable to connect the dots between those two questions. Every request is isolated, and therefore ends immediately after the question is asked or the request is made. Tweak the threaded conversations until agents are using 80% of the conversational AI engine's suggestions.

Part D: Incorporate custom data into the conversation to provide context. This involves having your AI engine integrate with your business data repositories (e.g. CRM). This allows the AI engine to incorporate knowledge such as account balance or claim status when interacting with a customer.

Part E: Once we have threaded, contextualized conversations, and the agent is utilizing nearly all of the suggestions, it's time to incorporate greetings and closing messages into the AI engine.

Don't Skip: Skipping this phase will result in your agents constantly having to correct the AI engine, which could lead to a bad customer experience.

3
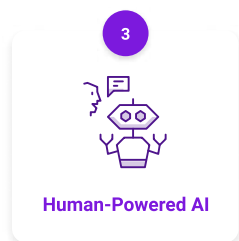
Human-Powered AI

PHASE 3 – HUMAN-POWERED AI

By this phase, human agents should be clicking accept for more than 50% of the AI-suggested responses. Due to the fact that the AI engine is becoming increasingly accurate, agents would probably start asking their managers why they have to bother clicking accept. This level of accuracy should lead to auto-sending the reply.

Auto-sending can be configured to automatically send the message when a response has 90% accuracy. At this point, the agent is primarily monitoring the bot, chiming in only when the AI engine cannot produce an answer with at least 90% accuracy. The lower level of involvement required on behalf of the agent enables them to easily manage multiple engagements at once. With agent capacity and efficiency maximized, monitoring should be focused on capturing customer sentiment and survey results where you'll see that the customer experience is not only maintained, but positively impacted by the use of Humanpowered AI.

Now, a company can start to measure the number of completed conversations that did not require human intervention. As this level increases, the company can start to prepare for.

Dont Skip: Skipping this phase will result in a bad customer experience. Channel switching or asking a customer to "rephrase" a question leads to a bad experience.

Success in this phase is measured by the percentage of conversations where no human intervention is required. Lower levels of human intervention signify to the company that they have engineered a higher quality, and more independent AI engine, and they're ready to start preparing for phase 4.

Despite having implemented well-threaded conversations in phase 2, skipping phase 3 carries the same risks we see with channel-switching today, even before the advent of AI engines. Even with 80-90% accuracy, 10% of the tens (or hundreds) of thousands of customers contacting you each month will hit a point of failure requiring them to channel switch, resulting in added costs and decreased loyalty.

4
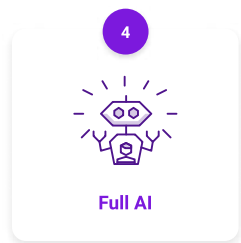
Full AI

PHASE 4 – FULL AI

In the final phase of this evolution, we see AI engines handling customer interactions and executing end-to-end conversations with an accuracy level of at least 95% for each response. But even in this phase, the human has not been removed from the conversation entirely.

CONCLUSION

In this phase, the confidence interval of each AI-generated response is weighted against the availability of human agents to maximize the customer experience. For example, if live agents are available and the confidence level of the AI engine for a particular response is at 80%, then a conversation will be initiated with a human agent. Whereas, in the event that there are limited human agents available, the AI engine may send the message with only a 80% confidence interval.

Notice that even in the phase of Full AI, we're still not at 100% automation. While the long-term goal may be to phase out human agents, this idea really only exists in a state of nirvana. So while AI is not quite mature enough to pass the Turing test or handle the increasingly complex long tails of customer conversations, it is important for companies to embrace the technology that is available and begin to prepare for the technology that is to come. Enterprises that implement this phased approach will find fewer instances where they need to course-correct because of entirely avoidable mechanical errors and, instead, see more predictable and desirable impacts, including improved customer satisfaction rates and overall decreases in costs.

ABOUT GLIA

Glia is an enterprise software company based in New York City that develops engagement solutions to meet or exceed the in-person customer experience online. Through our solutions, companies are able to identify high-value website visitors, communicate with them through

Deutsche Bank   PSEG   ORION   MAPFRE   BERKSHIRE HATHAWAY   MEMBERS 1st FEDERAL CREDIT UNION

live chat, audio, or video, and engage them through best-in-class CoBrowsing - all without any downloads or installations. Our solutions empower leading financial services institutions, retailers, and manufacturers to increase online sales and conversions, improve customer support and loyalty, and provide an overall better customer experience.

Written by Justin DiPietro
Justin DiPietro is a Co-Founder and Chief Operating Officer at Glia, a SaaS startup that provides Communication and Visual Engagement solutions to Enterprise organizations worldwide.

Deutsche Bank   PSECU   ORION   MAPFRE   BERKSHIRE HATHAWAY   M1st MEMBERS 1st FEDERAL CREDIT UNION