Four Channels for Conversational Technologies

Published: 22 January 2018 **ID:** G00341438

Analyst(s): Van Baker

Application leaders that want to deploy conversational platforms have four alternatives available to them. They must choose the alternative that fits their level of investment, and the business and technical needs of the enterprise.

Key Challenges

- Many enterprises underestimate the requirements for developing and deploying a conversational, platform-based chatbot because they assume that the service is either available as an off-the-shelf offering or can be built using a small dataset.
- Building a functional chatbot requires varying amounts of enterprise data and differing skill sets that can range from data scientists with machine-learning expertise to skills working with integrators on outsourced projects, depending on the channel selected. Domain experts are also necessary.
- Enterprises that use off-the-shelf services for natural-language processing must nonetheless provide data to build supplemental models in order to add domain-specific ontologies, taxonomies and vocabularies to the chatbot.

Recommendations

Application leaders responsible for artificial intelligence need to:

- Set a narrow scope of capabilities that your chatbot will address prior to its development and deployment to improve its efficiency.
- Match the channel for development and deployment of chatbots to the skill set of the enterprise to ensure a successful initial deployment of conversational technology.
- Deliver a clean, well-organized, metadata-tagged dataset to train any chatbot that requires domain-specific expertise to ensure good performance of the chatbot.

Table of Contents

List of Figures

Strategic Planning Assumption

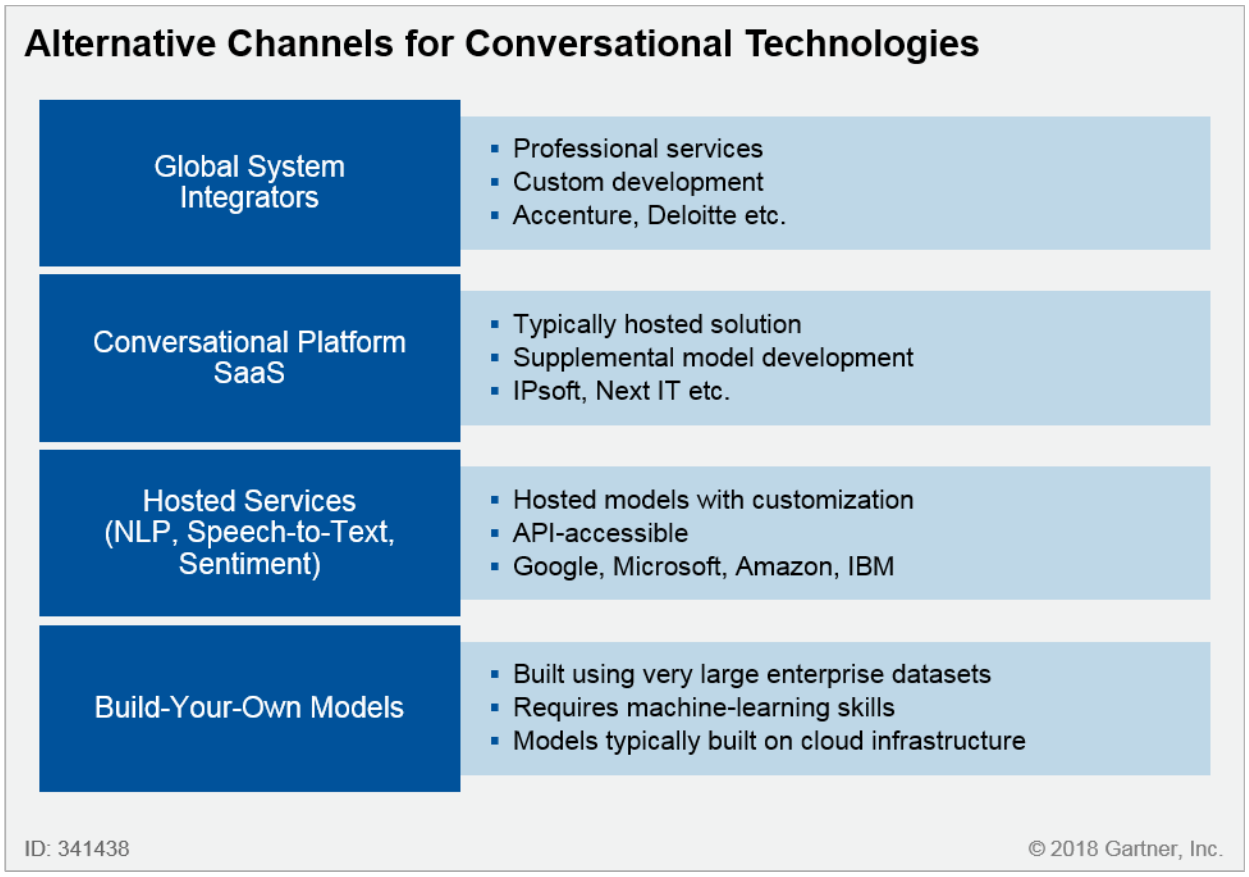
By 2020, over 50% of medium to large enterprises will have production chatbots deployed.

Introduction

Chatbots are on the mind of many enterprises. They are exploring the technology and a small percentage (6% to 7%) have proofs of concept or pilot programs going forward. An even smaller percentage (3% to 4%) of enterprises have production chatbots in place. Chatbots are expected to become widespread within the next two to three years, because the value proposition from a user perspective (ask a question, get an answer) is compelling. Unfortunately, many sources, including both tool providers and the media, have caused enterprises to believe chatbots are easy to create and deploy. Some small development shops and framework providers lead one to believe that you can input a couple of hundred Q&A pairs and have a chatbot up and running. While this is technically true, you will end up with a chatbot that users will not be happy with due to very poor performance. This will reflect poorly on your company and should be avoided.

Implementing conversational technologies in the enterprise, typically as a chatbot, can be accomplished via a few alternative channels. These alternatives vary based on the amount of effort required by the enterprise to implement. They range from complete control of the processes, technologies and data to complete outsourcing of all the tasks associated with the implementation of the technologies. The alternatives are shown in Figure 1.

Figure 1. Alternative Channels for the Implementation of Conversational Technologies



Source: Gartner (January 2018)

The amount of effort and skills required by the enterprise is greatest if it chooses to build its own models. This approach requires considerable expertise, including machine-learning capabilities and in-house data scientists, to cleanse the data, and get what will be a large dataset prepared for use to train models for the conversational solutions. It may also require skills in using deep neural network frameworks such as TensorFlow, Microsoft Cognitive Toolkit (CNTK), Apache MXNet and Caffe.

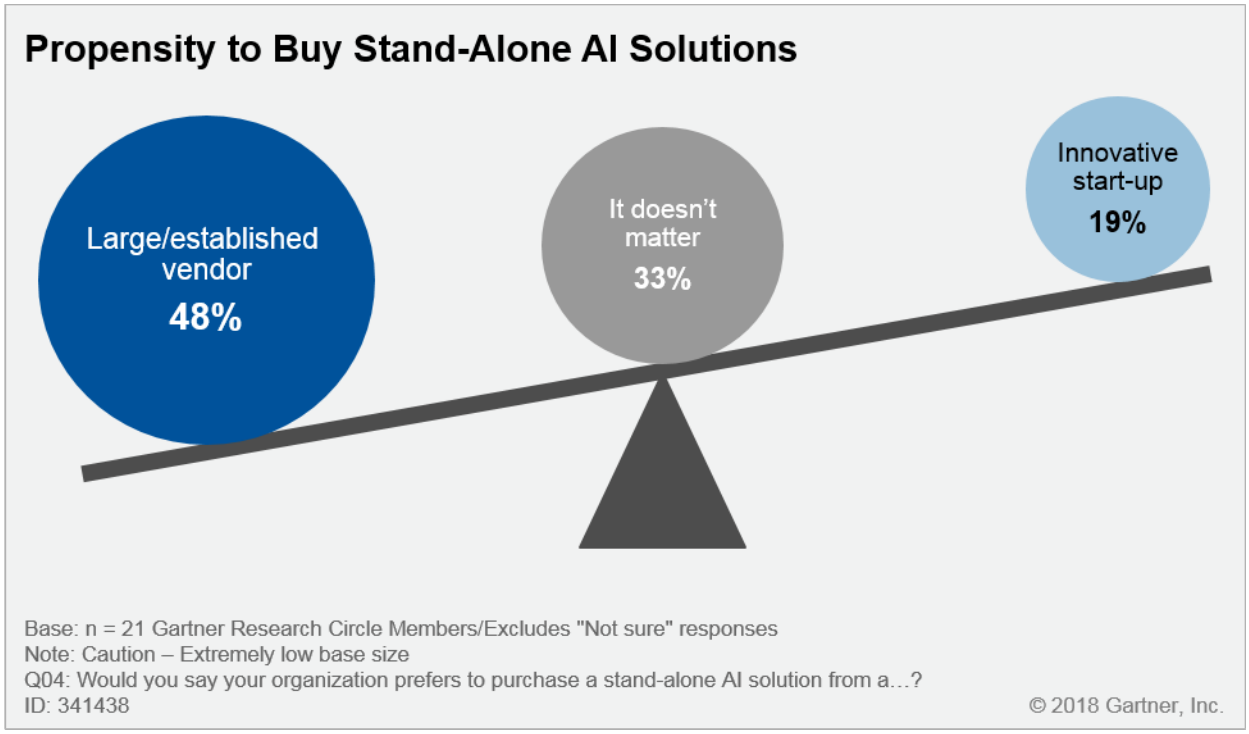
Some of the early-adopter enterprises are taking this approach using chatbot frameworks such as Amazon Lex, api.ai from Google, the Microsoft Chatbot Framework, Wit.ai from Facebook, or IBM Watson Conversation. With these frameworks, enterprises inherently use the hosted services for natural-language processing (NLP) models that each provider has developed. While these API-accessible NLP models perform well for general language, they lack the domain-specific ontologies, taxonomies and vocabularies necessary to address most enterprise use cases. This requires incremental models to be built and work in conjunction with the existing NLP algorithms.

Many providers in the market are offering conversational platform SaaS models to deliver chatbot capabilities to enterprises wanting to deploy these technologies. These providers, in many cases,

will use the underlying NLP models from the hosted service providers, but some will use their own NLP models. These providers will build the supplemental models needed to add domain-specific capabilities. They typically host the models in the public cloud, but some will implement on-premises solutions. These vendors also maintain and update models as needed to improve the performance of the chatbots.

Global system integrators will also develop and deploy conversational technologies for enterprises. These solutions can utilize any or all of the technical elements discussed previously. In short, this is the approach that requires the least effort on the part of the enterprise, but it is also a solution that tends to be the most expensive alternative. There is also a risk that skills will not transfer to the organization, resulting in difficulty in sustaining the models needed to support the chatbot. However, if an enterprise has no skills in conversational technologies, machine learning or data management, this may be the most viable alternative.

Figure 2. Propensity to Buy Stand-Alone AI Solutions



Source: Gartner (January 2018)

Analysis

Set a Narrow Scope of Capabilities That Your Chatbot Will Address Prior to Its Development and Deployment

For most application leaders, the best initial approach is to start small with a well-defined use case for the chatbot that you want to implement. Deploying a chatbot that is narrow in scope increases

the likelihood that the chatbot will perform well. Deploying a chatbot that is broad in scope and requires significant integration into back-end data sources is likely to result in a chatbot that performs poorly and users that are not satisfied with the capabilities of the chatbot. This focused approach also dictates that the use case in question should address simple requests with consistent, structured responses. Trying to deal with very complex questions increases the possibility of the chatbot performing poorly.

While narrowly focused chatbots with simple use cases are recommended, it is important to make users of the chatbot aware of its capabilities so they are not disappointed in the responses they receive. This requires that the chatbot's capabilities be spelled out when it is invoked, and that the chatbot can respond to the user when asked what it can do.

Match the Channel for the Development and Deployment of Chatbots to the Skill Set of the Enterprise

It is important to assess the skills that are present in the enterprise, or available to the enterprise from established sourcing arrangements, prior to deploying conversational solutions. Consider, for example:

- Does the enterprise have data scientists on staff?
- Does the business have expertise in deep learning frameworks and machine-learning capabilities?
- Does the infrastructure available to the enterprise offer the compute power and memory needed to build AI models? (This is especially true if the enterprise wants to employ unsupervised deep learning techniques to build their models, as this approach requires a much larger dataset than supervised learning.)
- Are any of these capabilities available from existing partners?
- Does the enterprise have the dedicated personnel needed to maintain and enhance these models on an ongoing basis?

All of these questions and others will help the enterprise determine which of the channels outlined above should be used to deploy their initial and subsequent chatbots.

It is important that application leaders understand the difference between training AI models and the inference that is done by the algorithms that result from the training. Much of what is written about chatbots focuses on the inference or execution of the algorithms that machine-learning models create without discussing the training needed to deliver the algorithms that make chatbots work. Training requires large datasets and a lot of effort to create a dataset that can be used to train the models that deliver the algorithms that facilitate the inference needed to extract intent from utterances or questions that are presented to the chatbot. Almost all chatbots that are deployed will require custom models to be trained to incorporate the domain-specific expertise needed by the chatbot to perform acceptably. The training of these domain-specific models or complete models is a very different task than the deployment of algorithms that provide the inference commonly used by chatbots to extract intent from utterances.

Deliver a Clean, Well-Organized, Metadata-Tagged Dataset to Train Any Chatbot That Requires Domain-Specific Expertise

The training of any chatbot to understand the unique terminology in the utterances or questions that it is presented with will utilize datasets that come from the enterprise. With the simplest of chatbot implementations, these are Q&A pairs, which can get very large due to the variations that can be presented by users to the chatbots. Very simple questions may have 10 variants in terms of how the question could be posed by the user. More-complex questions could have up to 350 or more variations in the way the question could be presented to the chatbot. Multiply this by the number of questions that you would like the chatbot to answer and the datasets can get quite large. This is further compounded by the fact that supervised learning requires these datasets to have metadata tags to improve the training of the model and resultant algorithms that support the chatbots. This scenario describes a very simple implementation where the chatbot merely matches an answer to the intent or question. More-sophisticated implementations, where the intent may be used to invoke an enterprise search or a SQL query, will require a greater degree of training. If the organization wants to build its own NLP models, the number of datasets needed increases by several orders of magnitude.

Regardless of the channel selected, the vast majority of enterprise use cases will need models to be built in order for the chatbot to perform adequately to address user needs. The datasets used to train supplemental models are an important asset for the enterprise. Conversational technologies are evolving rapidly and the segment will be very fluid for the foreseeable future. This, by definition, translates to a market where vendors will come and go. In this environment, the datasets are the assets that will allow enterprises to migrate conversational solutions from one vendor to another. Application leaders should maintain and improve these training datasets as they will be very valuable to the business. They should also make sure the organization retains ownership of these important datasets.

Gartner Recommended Reading

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

"Four Use Cases for Chatbots in the Enterprise Now"

"Seven Decision Points for Success With Virtual Customer Assistants"

"Market Guide for Virtual Customer Assistants"

"Architecture of Conversational Platforms"

"Cool Vendors in AI for Conversational Platforms, 2017"

"Lack of Focus on AI Licensing Will Result in Higher Costs, Risks and Long-Term Headaches"

"Innovation Insight for Conversational Commerce"

Evidence

The analysis presented in this document is based on conversations with various vendors in each of the channels that are described in the document, as well as conversations with AI-leading authorities from varying disciplines. Additional sources included many published sources in the area of AI, with an emphasis on language-related technologies.

More on This Topic

This is part of an in-depth collection of research. See the collection:

- Deliver Artificial Intelligence Business Value: A Gartner Trend Insight Report

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