**How chatbots work**

To interact with the human user, chatbots must be able to:

* parse the user input
* interpret what it means
* provide an appropriate response or output

For example, a user query could be, “Show me hotels in Los Angeles for tomorrow.”

A good chatbot will be able to identify the intent and entities of the query. The intent is the purpose or category of the user query, such as to retrieve a list of hotels. Entities are extra information that describes the user’s intent. In this case, the entities are “Los Angeles” and “tomorrow.” With these pieces of information, chatbots should be able to respond to the user with a list of available hotels for the correct location and date.

**Response architecture models**

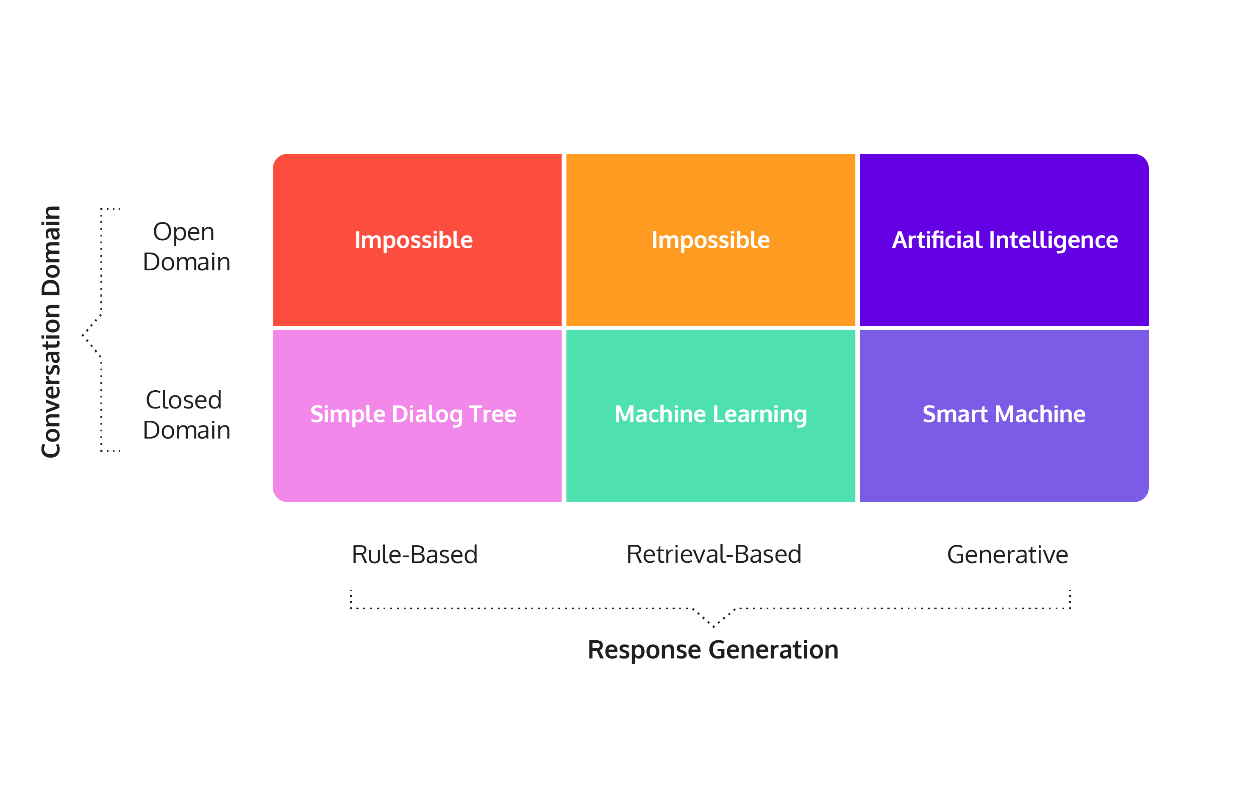
The simplest approach is the rule-based model, where chatbot responses are entirely predefined and returned to the user according to a series of rules. This includes ***decision trees*** that have a clear set of possible outputs defined for each step in the dialog.

Next, there is the retrieval-based model, where chatbot responses are pulled from an existing corpus of dialogs. Machine learning models, such as statistical NLP models and sometimes supervised neural networks, are used to interpret the user input and determine the most fitting response to retrieve. Like rule-based models, retrieval-based models rely on predefined responses, but they have the additional ability to self-learn and improve their selection of response over time.

Finally, generative chatbots are capable of formulating their own original responses based on user input, rather than relying on existing text. This involves the use of deep learning, such as LSTM-based seq2seq models, to train the chatbots to be able to make decisions about what is an appropriate response to return.

While generative models are very flexible and powerful in that they are not confined to a predefined set of rules or responses, they are also significantly more challenging to implement. Training these chatbots require an abundance of data, and it is often unclear what gets used for their decision-making, making them more prone to grammatical errors and nonsensical replies. By contrast, retrieval-based models can guarantee the quality of the responses since they are predefined, but these chatbots are in turn restricted to language that exists within the training data.

Thus, chatbots often use a combination of the different models in order to produce optimal results. For example, a customer support chatbot may use generative models for creating open-ended small talk with the user, but then are able to retrieve professional, predefined responses for answering the user’s inquiries regarding the business or product.



**Conversation domains**

Chatbots can also be categorized based on the range of conversation topics they are able to cover. Closed domain chatbots, or dialog agents, are restricted to providing responses with a particular focus, such as booking a hotel room. Because they are designed with a specific goal in mind, these chatbots are often very efficient and have great success in accomplishing what they are intended to accomplish. The user-perceived quality is also high, because users don’t expect the chatbots to provide responses outside of the pre-established domain.

On the other hand, open domain chatbots, or conversational agents, are capable of exploring any range of conversation topics, much like how a human-to-human interaction would be. Many of these “companion bots” have filled the roles of a friend or therapist, allowing the user to connect with them on an emotional level. While they have great potential, open domain chatbots are challenging to implement and evaluate.

*Another way chatbots can be categorized is by which side – the user or bot – is able to take initiative on the conversation. Looking back at our previous example on hotel search, notice how the user is free to provide their request in their own words, and the chatbot is able to identify and piece together the relevant keywords to answer the query. This is an example of a mixed-initiative system, representative of a normal human-to-human conversation where all participants have the chance to take initiative.*

By contrast, a system-initiative system is one where the chatbot controls the conversation and explicitly asks for each piece of information, such as the date and location for the hotel booking. While this system is more straight-forward to implement because user response can be anticipated, it lacks the flexibility and naturalness that characterize a normal human dialog.

**Ethics of Chatbots**

From the user’s perspective, a big ethical consideration when it comes to technology is transparency. In other words, is the user aware of all aspects involving the chatbot and the consequences of interacting with one?

Be open and clear about data usage, ownership, and protection. One way to maximize transparency includes implementing a data regulation system like the European Union’s GDPR, which gives individuals more control over their personal data.

***Did you know:***[*Google Duplex*](https://blog.research.google/2018/05/duplex-ai-system-for-natural-conversation.html) *system, which can carry out convincingly natural, human-like phone conversations for specific tasks like booking appointments.*

Additionally, we need to take caution when training the bot to ensure that it behaves appropriately. If it is not properly trained, the chatbot could be at risk of displaying racism, sexism, or use of abusive language.

This is exactly what happened to Microsoft’s Tay, a bot the company created for use on Twitter that generated its responses based on how users interacted with it. When various users began posting offensive tweets towards the bot, Tay reciprocated by emulating that same language in its replies.

**This type of behavior can be prevented with more effective training of the bot, such as using supervised learning to ensure the quality of training data and better predict the outputted responses.**

how can a chatbot demonstrate compassion and empathy towards the user? If the user is expressing suicidal thoughts, would the bot be able to offer help? These are all ethical questions that should be considered*.*

***Did you know:*** *Some chatbots, like Woebot, are specially trained to be able to help users with their mental health.*

**The History of Chatbots**

In October of 1950, Alan Turing proposed an approach for evaluating a computer’s intelligence and famously named his method, The Imitation Game. The premise is that an interrogator talks to two people through a “typewritten” machine (today we would refer to this as instant messaging). The catch is that only one of the conversations is with a real person – the other is with a computer.

Turing posited that, by the turn of the century (the year 2000), in a well-controlled experiment, a computer should be able to fool the interrogator 70% of the time. While we’ve made a lot of progress since 1950, no algorithm has consistently reached this bar. However, there has still been substantial progress in the field of chatbot development, which has led to a multi-billion-dollar industry and dozens of profitable products.

**Some examples of chatbots**

**ELIZA**

ELIZ was developed by Joseph Weizenbaum at MIT Laboratories in 1966 and was the first chatbot that made a meaningful attempt to beat the Turing Test. It used pattern-recognition to pick out patterns in a person’s speech, then repeated the words back to the person in a premade template. The most famous implementation of the ELIZA chatbot is DOCTOR. In this implementation, the chatbot acts like a psychotherapist, responding to a patient’s statements by selecting a phrase from the respondent and parroting them back in the form of a question. This form of chatbot is rule-based, because the program responds to a person based on rules that a developer establishes in a predefined script.

**PARRY**

Kenneth Colby developed PARRY while at Stanford University in 1968. Colby built PARRY with a similar rule-based method to ELIZA. However, it was designed to model the behavior of a person with diagnosable paranoia. PARRY also had a richer response library and was able to simulate the mood of a person by shifting weights of mood parameters. PARRY would respond differently based on the distribution between mood parameters for anger, fear, or mistrust. PARRY passed a modified Turing Test by fooling people who tried to distinguish the difference between it and a person with paranoia.

**A.L.I.C.E.(Artificial Linguistic Internet Computer Entity)**

Richard Wallace started developing A.L.I.C.E. in 1995, shortly before leaving his computer vision teaching job. Wallace improved upon ELIZA’s implementation by continuing to watch the model while it had conversations with people. If a person asked an A.L.I.C.E. bot something it did not recognize, Wallace would add a response for it. In this way, the person who designs an A.L.I.C.E.-powered device could continuously modify it by adding responses to unrecognized phrases. This means that well-developed A.L.I.C.E. bots can respond to a variety of questions and statements based on the developer’s needs. In a chapter from the 2009 book, Parsing the Turing Test, Richard Wallace described this process as supervised learning, because the developer – who he calls the botmaster – can supervise the learning of the model.

In 2000, 2001, and 2004, A.L.I.C.E., won the Loebner prize, which is a contest that was started in 1980 to award computer programs that are the most human-like. The Loebner prize is awarded to the bot that performs the best on the Turing Test. The judges of the Loebner competition have two conversations simultaneously. One with a real person and the other with a bot. The winner of the competition is the one that tricks a judge the highest percentage of the time.

**Jabberwacky**

Jabberwacky was developed by Rollo Carpenter in the 1980s and was launched on the web in 1997. It was the first chatbot that tried to incorporate voice interaction. Two versions of Jabberwacky won the Loebner prize in 2005 and 2006. Jabberwacky has undergone continuous development since it debuted on the web. When it launched, it used a similar rule-based approach to previous models, like ELIZA and PARRY. However, in 2008, the model was renamed Cleverbot and updated to include a method for learning without the supervision of a botmaster. Cleverbot can parse and save human responses to questions, and respond similarly if a human asked it the same question.

**Mitsuku**

Mitsuku was developed by Steve Worswick during the early 2000s and first won the Loebner prize in 2013. The model is still actively developed and has won the Loebner Prize in 2016, 2017, 2018, and 2019, making it the most human-like chatbot available. Mitsuku works more like the A.L.I.C.E. It is a supervised learning model, where developers actively tweak the rules to make interactions with Mitsuku more human-like.

Alexa falls well short of competing with Mitsuku in carrying out a conversation.

With the growing market for chatbot-driven technologies, there is far more money and interest in chatbot development than just a few years ago.

In 2017, Amazon started the Alexa challenge, where teams compete to develop a chatbot that has the best conversation with users. The winners of this challenge receive $500,000, and Amazon tests each model’s ability to converse on Alexa devices.