ChatBots

Antoine PALISSON

TABLE OF CONTENTS

Introduction

Definition

A **chatbot** is a program designed to simulate human conversation and interact with users through text-based, sometimes voice-based, interfaces. It uses natural language processing and artificial intelligence techniques to understand and answer to user inputs in a conversational manner.

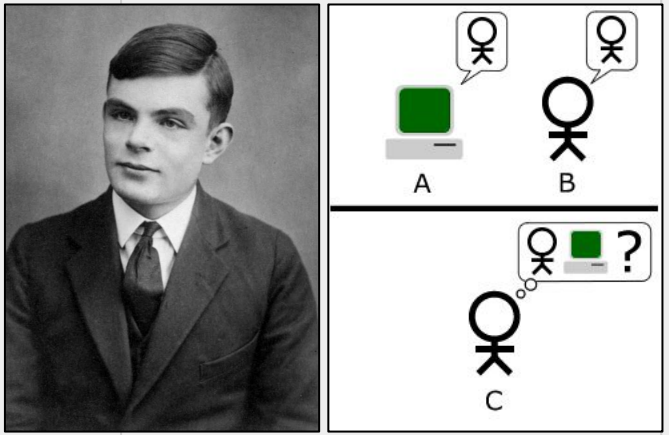
Which problem do Chatbots solve ?

Chatbots provide scalable and 24/7 customer support, reducing response times, and improving cost efficiency. They offer personalized recommendations, enable self-service assistance, and enhance user engagement. Additionally, they can automate routine tasks, handle large volumes of inquiries simultaneously, and operate around the clock. Overall, they are generally design to increase customer satisfaction, and create more efficient and engaging experiences.

Antoine PALISSON

History

The history of chatbots dates back to 1950 when **Alan Turing** proposed the idea of a program that could communicate with humans without being recognized as artificial. The concept, known as the Turing test, sparked the development of chatbots.



Alan Turing(1) & the Turing test(2)

A chat with ELIZA

ELIZA	Hello, I am Eliza. I'll be your therapist today.
ME	Are you a real therapist ?
ELIZA	Why are you interested in whether or not I am a real therapist?
ME	I prefer to know who I am talking to.
ELIZA	Do you believe it is normal to be talking to?
ME	Yes ... your question is strange.
ELIZA	Why are you so concerned about my question is strange?
ME	Because it doesn't make any sense
ELIZA	Can you elaborate on that?

The first chatbot, **ELIZA**, was created in 1966 and simulated a psychotherapist by returning the user's sentences in an interrogative form. However, ELIZA's knowledge was limited, and it could only carry out a specific range of discussions and could not learn from the conversation.

(1): wikipedia.com (2): wikipedia.com

History

In 1972, **PARRY** was developed as a more advanced version of ELIZA. It acted as a patient with schizophrenia and had a more sophisticated control structure that responded based on a system of assumptions and emotional responses. Nevertheless, it also had its limitations and was considered low in language understanding and the ability to express emotions.

PARRY	People get on my nerves sometimes.
ELIZA	I am not sure I understand you fully.
PARRY	You should pay more attention.
ELIZA	Suppose you should pay more attention.
PARRY	You're entitled to your own opinion.
ELIZA	What makes you think I am entitled to my own opinion?
PARRY	It bothers me just to be around people in general.
ELIZA	Please go on.
PARRY	I went to the track at Bay Meadows a while back.
ELIZA	What does that suggest to you?
PARRY	I went to the races.
ELIZA	What else would you like to discuss?
PARRY	Horse racing interests me.
ELIZA	Why do you say that just now?
PARRY	There's a lot of crooked gambling going on there.
ELIZA	I am not sure I understand you fully
PARRY	It's clear to me if not to you.
ELIZA	Do you think it's likely that not to I?
PARRY	I went to the races.
ELIZA	Please go on.
PARRY	I gamble on horses.

PARRY's encounter with ELIZA - 1972

phrasee.co

History

The first chatbot to utilize Artificial Intelligence was **Jabberwacky** in 1988. It used contextual pattern matching to respond based on previous discussions. However, it could not reply at high speeds and could not work with a vast number of users.

In 1995, **ALICE** was developed, which was an online chatbot inspired by ELIZA that also used pattern matching . ALICE had about 41,000 templates and related patterns which was significantly higher than ELIZA's 200 keywords and rules. It had the ability to discuss any topic on the web.

Both **ALICE** and **Jabberwacky** later won the **Loebner Prize**. It was an annual competition in the field of artificial intelligence (AI) that focuses on testing the capabilities of chatbot systems and especially to determine the chatbot that exhibits the most human-like conversational abilities.

The Loebner gold medal that has never been awarded yet

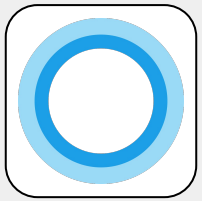
History

In 2001, **SmarterChild** was developed and made available on messaging platforms like AOL and MSN. It was unique because it was the first chatbot that could perform practical daily tasks like retrieving information about movie times, sports scores, stock prices, and even the weather. This marked a significant development in both machine intelligence and human-computer interaction as information systems could be accessed via interaction with a chatbot.

The late 2000s to 2010s saw a massive leap in chatbot development with the advent of smart personal voice assistants. These voice assistants, like Apple's Siri, Google Assistant, Microsoft's Cortana, Amazon's Alexa, and IBM's Watson, revolutionized the way users interacted with their devices.

Siri

Google Assistant



Cortana

Alexa

IBM Watson

History

Launched in 2010, was groundbreaking as it facilitated user interactions using voice commands, integrated with various internet services, and even adapted to users' language usage and preferences. However, it had limitations such as a lack of support for several languages, and difficulty in understanding heavily accented speech or in noisy environments.

Launched in 2011, was remarkable in its understanding of natural human language which was demonstrated when it won a game of the Jeopardy Game show against human champions.

Launched as Google Now in 2012, which later evolved into Google Assistant in 2016. It was designed to deliver information proactively, predicting user needs. However, it faced criticism for its lack of personality and potential privacy violations.

Launched in 2014, was designed to recognize voice commands and perform tasks like identifying time and location, supporting reminders, sending emails, and managing lists.

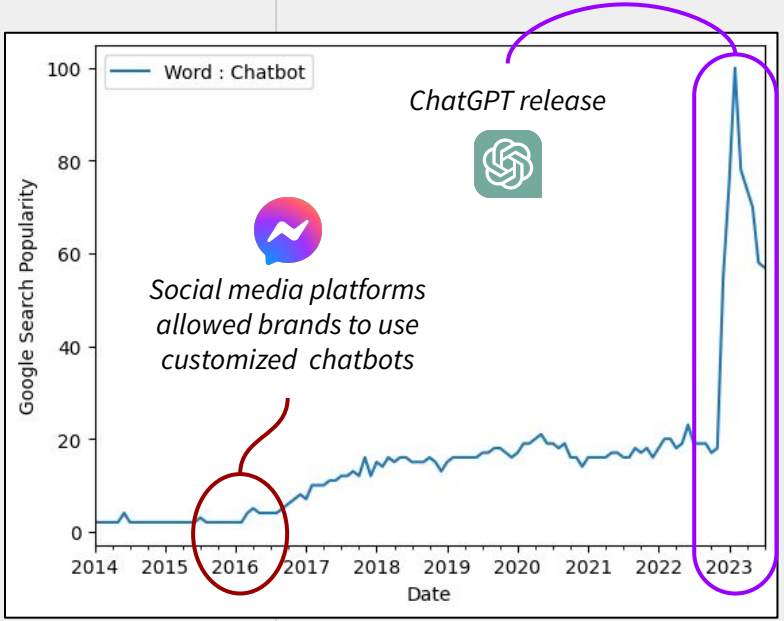
Also introduced in 2014, went a step further by integrating with home automation and entertainment devices, making the Internet of Things (IoT) more accessible to humans. Developers could even use the Alexa Skills Kit (ASK) to create and publish skills for Alexa.

History

In 2016, social media platforms opened up for developers to create chatbots for their brands or services, which dramatically changed how customers interacted with businesses. Tens of thousands of chatbots were developed for popular messaging platforms, industrial solutions, and research purposes.

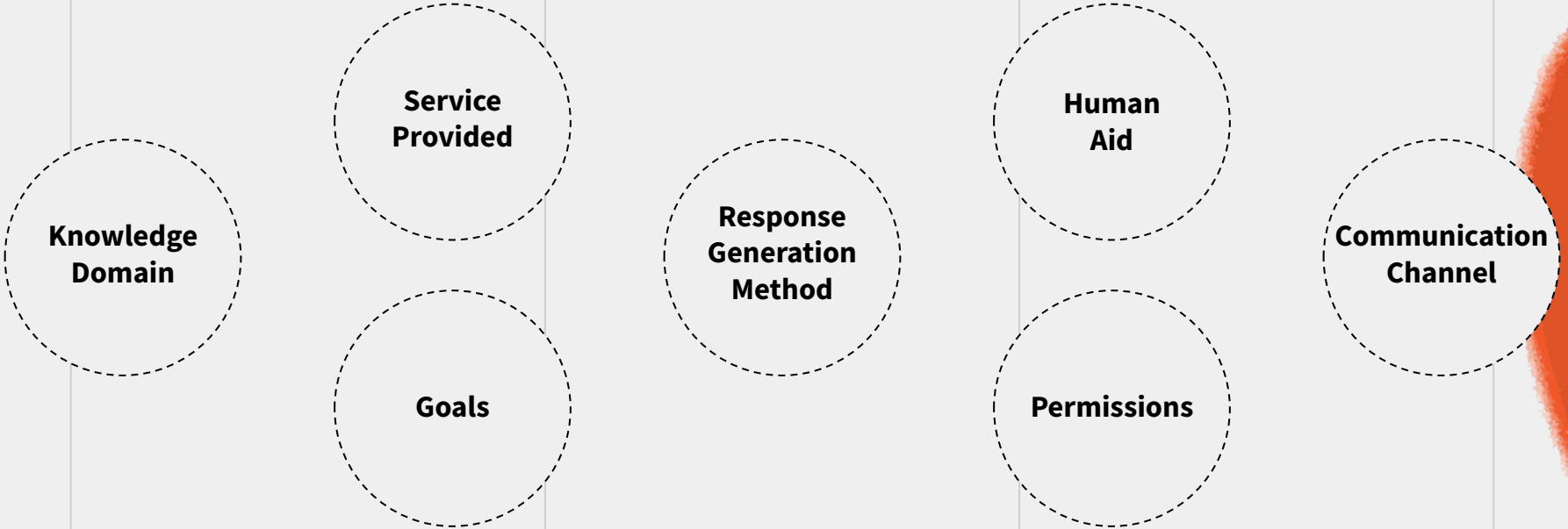
In recent years, chatbots have become increasingly sophisticated thanks to the advances in Large Language Models (LLM). In 2022, OpenAI introduced **ChatGPT**, a chatbot based on their Generative Pretrained Transformer model (GPT) that was able to generate impressively human-like text.

This release was quickly followed by other similar LLM chatbots such as **Google Bard** in 2023.



Categories

According to Adamopoulou et al., 2020, chatbots can be described with 7 criteria.



Categories

The **knowledge domain** of a chatbot refers to the breadth and depth of information it can understand and converse about. This can be broadly divided into two categories:

Open domain - They are more advanced than generic ones. They are designed to handle a wide variety of topics, but unlike generic chatbots, they can delve into greater depth on many topics. They use machine learning and natural language processing to better understand the context and nuances of a conversation.

Closed domain - These chatbots are experts in a specific domain or subject. They are designed to provide detailed information or perform tasks related to a specific topic.

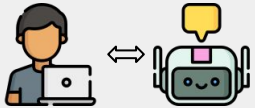
For example, a chatbot designed by a bank might only be able to discuss topics related to medicine.

Categories

The **service provided** by a chatbot refers to the type of interaction and service it offers to users or other systems. It can be categorized into three types:

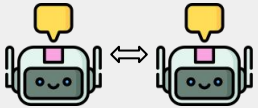
Interpersonal - They assist in interactions between people in order to facilitate conversation, help solve problems, and generally work to ensure smooth communication between different parties.

For example, social media bots helping users interact with each other, or customer service bots that links customers and the company.



Intrapersonal - They are designed to interact directly with an individual user for a variety of personal tasks.

For example, personal assistants that help schedule meetings and set reminders or therapy bots that assist with mental health.



Inter-agent - They are designed to interact with other bots or digital systems.

For example, a bot might gather information from databases, interact with other bots to carry out tasks, and then present the results to a user.

Categories

The **goals** of a chatbot refer to the main objectives that they are designed to achieve. They are generally categorized into three types:

Informative

They are designed to provide users to any requested information such as answering customer inquiries about a product, provide news updates, weather forecasts and so on.

Conversational

They are designed to engage users in a dialogue, often for entertainment or social interaction. They aim to emulate human conversation as closely as possible, allowing users to have natural conversations with them.

Task-based

They are designed to help users complete specific tasks. Tasks could be setting reminders, scheduling appointments, assisting in online shopping or even booking tickets.

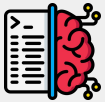
Categories

The **response generation method** of a chatbot refers to how a chatbot creates its responses to user inputs. There are generally three main categories:

Rule-based

They operate based on a predefined set of rules and heuristics. When a user inputs a message, the chatbot matches the input against its set of rules and delivers a predetermined response. The rules are explicitly programmed by developers.

Retrieval-based



They select responses from a predefined set of responses. The selection process is typically powered by machine learning algorithms, which predict the most appropriate response based on the context of the conversation and the user's input.

Generative

They generate responses on-the-fly using natural language processing techniques, typically powered by machine learning. They don't choose from a predefined set of responses but create new responses from scratch.

Categories

The **human aid** criterion refers to the degree of human involvement in the chatbot's operation. This is typically classified into two categories:

Autonomous

They operate without direct human intervention by using artificial intelligence and machine learning to understand and respond to user inputs. These chatbots can be designed to learn from their interactions and improve their responses over time.

Human-mediated

The chatbot works with human oversight or intervention. A human operator might be involved in guiding the chatbot's responses, especially in complex or uncertain situations.

> *For example, customer services could use a chatbot to handle basic inquiries but answer more complex issues thanks to a human representative.*

Categories

The **permissions** of a chatbot refer to the legal or licensing framework under which a chatbot operates. This can be divided into two main categories:

Open source - They are those whose codebase is publicly available for anyone to use, modify, and distribute. They are typically developed in a collaborative manner with contributions from various developers around the world. The primary advantage of open source chatbots is their transparency and the ability for anyone to customize and improve upon them.

Commercial - They are proprietary, and their codebase is typically not available for public use. Users typically need to pay to use these chatbots or to access premium features. The advantage of commercial chatbots is that they often come with dedicated support and regular updates, and they may have more advanced features or capabilities.

Categories

The **communication channel** of a chatbot refers to the mode through which the chatbot interacts with users or other systems. These can generally be categorized into three types:

Voice

Voice assistants such as Alexa, Google, Siri …

Image

Understand visual input from the user (photos, pictures …)

Text

Most common communication channel for chatbots

These communication channels are not mutually exclusive. Many chatbots support multiple communication channels.

For example, a virtual assistant might primarily use voice communication but also support text input for situations where voice isn't convenient or appropriate.

Applications

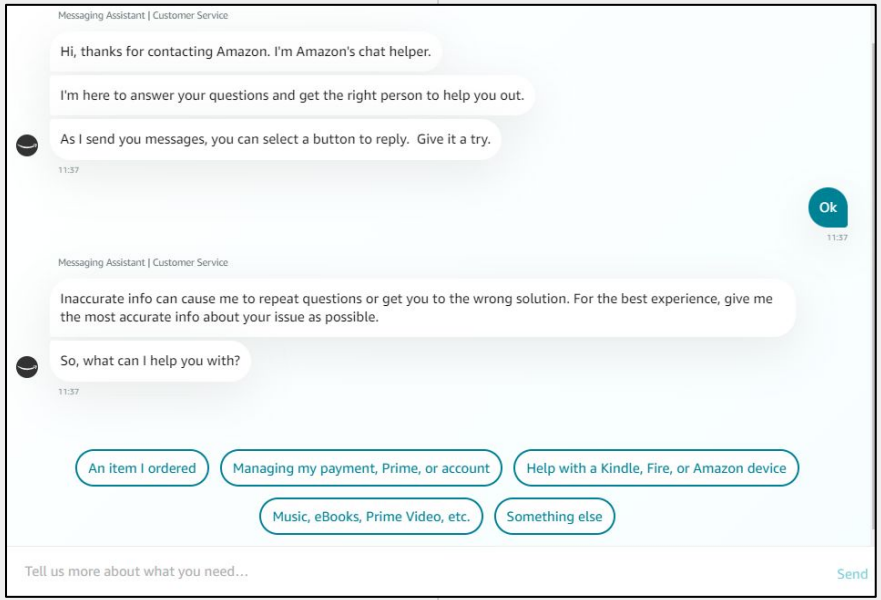
Chatbots have been widely adopted across multiple sectors due to their ability to automate tasks, provide instant customer service, and handle multiple inquiries at once.

- Customer Service
- Healthcare
- Banking & financial services
- Travel and Hospitality
- Education
- Communication and IT
- Human Ressources

Applications

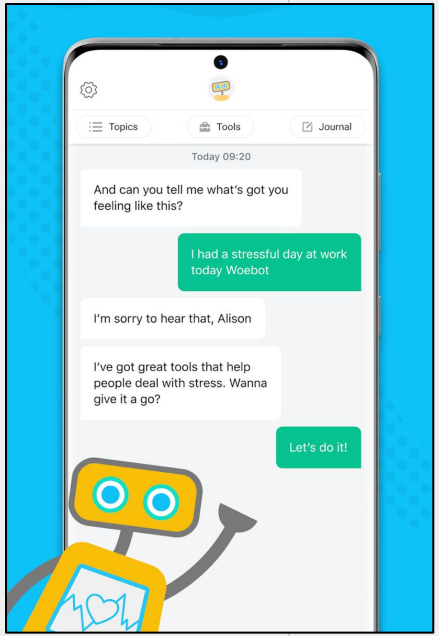
Customer Service Chatbots typically handle customer inquiries, guide customers through product selections, make recommendations based on user preferences, and even facilitate transactions. They greatly improve the shopping experience by providing immediate assistance and personalizing the user's shopping journey.

Amazon's chatbot assists with order tracking, returns, and other customer service inquiries.



Applications

Healthcare Chatbots are used to answer common patient inquiries, facilitate appointment booking, send medication reminders, and provide basic health advice. Some more advanced healthcare chatbots can conduct initial patient screenings and assist with mental health therapy.



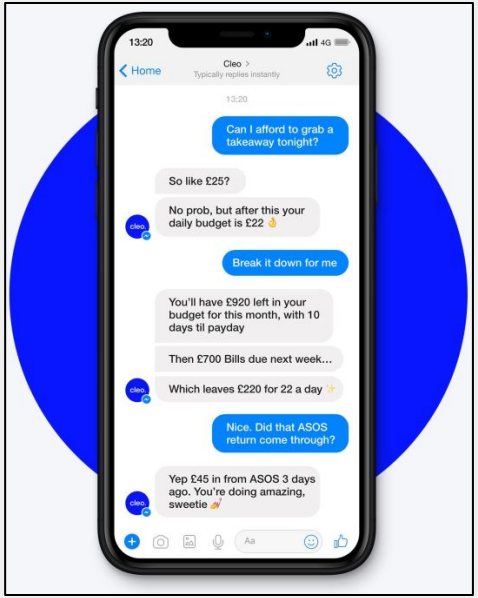
Woebot Health™

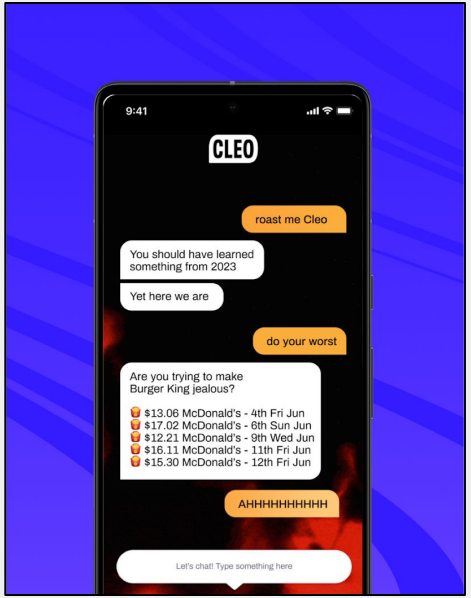
Woebot, developed by psychologists at Stanford University, is a therapy chatbot that helps users manage their mental health.

Applications

Many **financial** institutions have adopted chatbots for customer service, transaction facilitation, and providing financial advice. Chatbots in this sector can answer customer queries about account balances, recent transactions, loan eligibility, etc., and can also assist with transactions like money transfers.

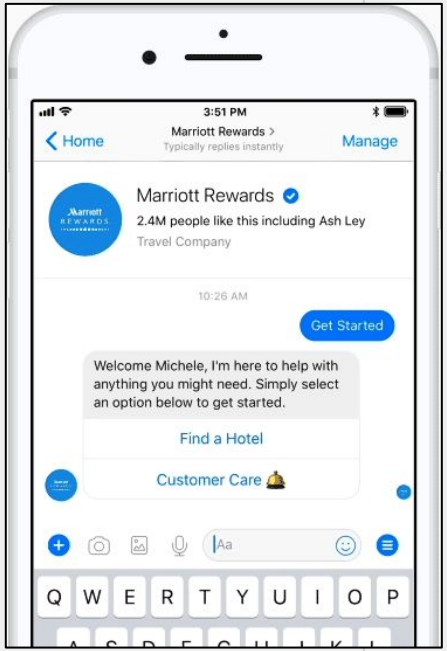
Cleo is an AI assistant for personal finance management.





Applications

Travel and Hospitality Chatbots can provide information on bookings, availability, pricing, and local attractions. They can help users with flight bookings, hotel reservations, car rentals, and even suggest personalized travel itineraries.



Marriott International's chatbot assists with booking hotel rooms.

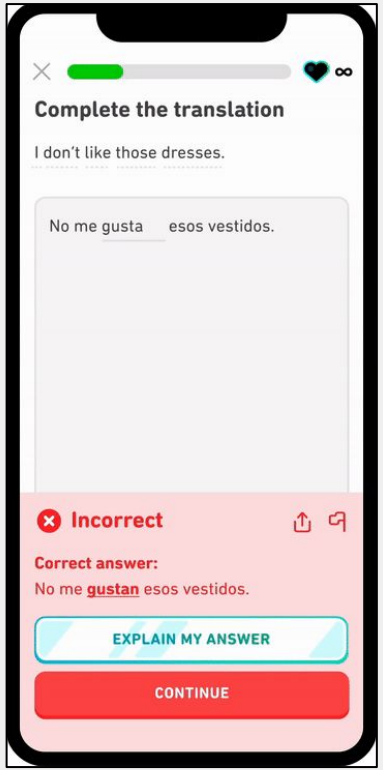
***Expedia**, another booking website, is currently experimenting with a ChatGPT plugin to help planning trips.*

Applications

Educational institutions and e-learning platforms use chatbots to assist students with common inquiries, provide study resources, and facilitate course selections. Some advanced chatbots can even help with homework and grade assignments.

duolingo

Duolingo, a language-learning platform, uses a chatbot to help learners practice conversation in different languages. It uses a GPT-4 model which is currently one of the most advanced chatbot.



User Experience

The aim is to make a chatbot's interactions with the user as natural, seamless, and efficient as possible. Indeed, designing a great **conversational UX** (*user experience*) isn't just about the technology used, but about understanding human communication and providing value in a way that feels natural and intuitive to the user.

1. **Keep Conversations Natural and Contextual**
2. *Clarity & Brevity*
3. *Error Handling*
4. *Setting Expectations*
5. *Guidance*
6. *Feedback Mechanism*
7. *Personalization*
8. *Respect for User's Time*
9. *Ethics & Privacy*

Keep Conversations Natural and Contextual

Chatbots should be designed to mimic human-like conversations as closely as possible. They should understand and respond based on the user's input, keeping track of the context throughout the conversation.

User Experience

The aim is to make a chatbot's interactions with the user as natural, seamless, and efficient as possible. Indeed, designing a great **conversational UX** (*user experience*) isn't just about the technology used, but about understanding human communication and providing value in a way that feels natural and intuitive to the user.

1. *Keep Conversations Natural and Contextual*
2. **Clarity & Brevity**
3. *Error Handling*
4. *Setting Expectations*
5. *Guidance*
6. *Feedback Mechanism*
7. *Personalization*
8. *Respect for User's Time*
9. *Ethics & Privacy*

Clarity and Brevity

Conversational UX design should be clear and concise. Too much information can overwhelm the user and lead to miscommunication. Responses should be easy to read and should not contain unnecessary jargon or complex language.

User Experience

The aim is to make a chatbot's interactions with the user as natural, seamless, and efficient as possible. Indeed, designing a great **conversational UX** (*user experience*) isn't just about the technology used, but about understanding human communication and providing value in a way that feels natural and intuitive to the user.

1. *Keep Conversations Natural and Contextual*
2. *Clarity & Brevity*
3. **Error Handling**
4. *Setting Expectations*
5. *Guidance*
6. *Feedback Mechanism*
7. *Personalization*
8. *Respect for User's Time*
9. *Ethics & Privacy*

Error Handling

A good chatbot should be able to handle errors gracefully. When it fails to understand the user's input, it should be able to request for clarification without frustrating the user. It should also learn from these interactions to improve over time.

User Experience

The aim is to make a chatbot's interactions with the user as natural, seamless, and efficient as possible. Indeed, designing a great **conversational UX** (*user experience*) isn't just about the technology used, but about understanding human communication and providing value in a way that feels natural and intuitive to the user.

1. *Keep Conversations Natural and Contextual*
2. *Clarity & Brevity*
3. *Error Handling*
4. **Setting Expectations**
5. **Guidance**
6. *Feedback Mechanism*
7. *Personalization*
8. *Respect for User's Time*
9. *Ethics & Privacy*

Setting Expectations

Chatbots should inform users of their capabilities and limitations at the beginning of the interaction.

Guidance

Chatbots should guide the conversation by providing suggestions or offering prompts.

User Experience

The aim is to make a chatbot's interactions with the user as natural, seamless, and efficient as possible. Indeed, designing a great **conversational UX** (*user experience*) isn't just about the technology used, but about understanding human communication and providing value in a way that feels natural and intuitive to the user.

1. *Keep Conversations Natural and Contextual*
2. *Clarity & Brevity*
3. *Error Handling*
4. *Setting Expectations*
5. *Guidance*
6. **Feedback Mechanism**
7. **Personalization**
8. *Respect for User's Time*
9. *Ethics & Privacy*

Feedback Mechanism

Chatbots should provide a way for users to give feedback.

Personalization

Personalizing the user experience can enhance engagement. This can involve using remembering past interactions, or customizing recommendations based on the user's feedback.

User Experience

The aim is to make a chatbot's interactions with the user as natural, seamless, and efficient as possible. Indeed, designing a great **conversational UX** (*user experience*) isn't just about the technology used, but about understanding human communication and providing value in a way that feels natural and intuitive to the user.

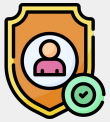
1. *Keep Conversations Natural and Contextual*
2. *Clarity & Brevity*
3. *Error Handling*
4. *Setting Expectations*
5. *Guidance*
6. Feedback Mechanism
7. Personalization
8. **Respect for User's Time**
9. *Ethics & Privacy*

Respect for User's Time

Responses should be swift, but not at the cost of accuracy. A well-designed chatbot understands when a user needs a quick response or when it's appropriate to take more time to provide a thorough, detailed answer.

Ethics & Privacy

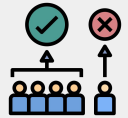
Ethical considerations are very important when it comes to chatbot design and deployment. Two of the main ethical issues related to chatbots are **bias** and **data privacy**:



Data Privacy

Chatbots often handle sensitive user data, including personal information and conversation history. The chatbot must be designed with strong data protection measures to prevent misuse or unauthorized access. Users should be informed about what data the chatbot collects, why it is collected, how it is stored, who has access to it, and how it can be deleted.

Bias



The potential for bias arises when a chatbot learns from data that includes biased human decisions, or when the data used for training is not representative of all users. This can lead to unequal treatment or discrimination. It's important for developers to carefully curate and review the data used for training to ensure it is as unbiased and representative as possible.

Ethics & Privacy

In addition to these issues, other ethical considerations can include transparency, accountability and respectful interaction.

Transparency

Users should be aware that they are interacting with a chatbot and not a human. This should be made clear at the beginning of the interaction. It's also important to be transparent about the chatbot's capabilities and limitations.

Accountability

If a chatbot makes a mistake, there should be mechanisms in place to correct it and to prevent similar errors in the future. Developers and organizations must take responsibility for the chatbot's actions and any harm they may cause.

Respectful Interaction

The chatbot should be designed to interact respectfully with users, avoiding any form of harassment, hate speech, or other inappropriate behavior.

Future Trends

With the growing popularity of smart speakers and virtual assistants like Amazon's Alexa, Google Home, and Apple's Siri, there's a rising trend of **voice bots**. These bots can engage in natural language conversations, performing tasks such as setting alarms, making purchases, and answering queries. The next evolution might include more complex interactions, such as interpreting different accents and dialects or understanding complex commands.

Chatbots of the future might not just rely on text or voice; they might also be able to interpret and respond to **visual inputs**. This could be especially useful in situations where visual understanding can help provide better responses, such as helping users assemble a piece of furniture or identifying a plant species from a photo.

As more and more devices become connected to the **Internet of Things**, chatbots could serve as an interface between users and these devices. For example, a user could instruct a chatbot to preheat their smart oven or adjust their smart thermostat.

Future Trends

The next generation of chatbots could also include **Emotional AI** which involves machines recognizing, interpreting, and responding to human emotions. This could mean understanding the tone and sentiment behind a user's message and responding appropriately. For instance, if the bot detects frustration or anger in a user's text or voice, it could alter its responses to try to appease the user. As the technology improves, emotional AI could become a standard feature in customer service and mental health bots.

Future chatbots might also have more advanced **self-learning capabilities**, allowing them to improve their performance over time without the need for manual retraining. They may even be able to recognize when they've made a mistake and correct themselves.

Exercises

Exercise I

Pick a business sector and analyze the chatbots of two leading companies in that sector.

Go to the Applications 🔗 section to see example of sectors

Instructions : Assess how effective the chatbots are, the technology used, the user experience, and how the chatbots integrate with the companies' broader marketing strategies.

Exercise II

Create a simple chatbot design blueprint for a specific business scenario.

Go to the Categories 🔗 section to see how to build a chatbot.

Instructions : Identify the type/category and the persona of the bot, the types of queries and the potential responses. You can base your chatbot on the ones studied in the first exercise.

Natural Language Processing for chatbots

NLP and Chatbots

A key element that enables chatbots to function efficiently is Natural Language Processing. It generally helps with four main chatbots functions:

Understanding User Input

It’s the first step in any conversation a chatbot has with a user. NLP enables chatbots to break down the text into understandable and analyzable pieces.

Intent Recognition

After the initial processing of the user's input, NLP helps the chatbot identify the user's intent, that is, what the user wants to achieve with the conversation.

Generating Correct Answers

Once the chatbot understands the user's intent, NLP plays a crucial role in ensuring that the generated response is contextually appropriate and grammatically correct.

Continuous Learning

Chatbots can learn to provide more accurate responses and handle a wider variety of queries over time. NLP facilitates this learning process.

Preprocessing

The first chatbot’s operation is the comprehension phase. During that phase, the chatbot interprets the user's input and prepares it for further processing. It generally consists of four preprocessing NLP techniques:

Tokenization

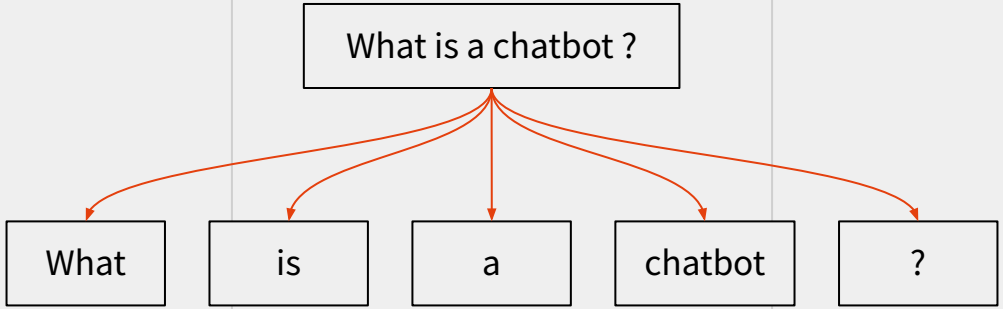
Lemmatization

Stemming

Part-of-speech Tagging

Preprocessing

Tokenization is the process of breaking down a user's input into individual words or tokens. It is the most basic NLP operation and is generally the first step.



Each word and punctuation of the original sentence have been separated into tokens.

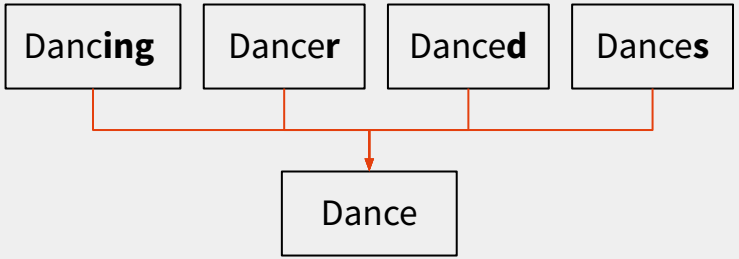
Antoine PALISSON

Preprocessing

Both **Stemming** and **lemmatization** process and normalize words from the tokenization step. They reduce words to their root form to focus on their underlying meaning.

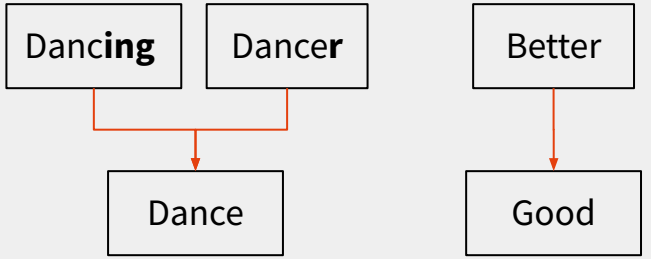
Stemming

It is the process of reducing inflected or derived words to their root form. The purpose of stemming is to bring word variants to a common base that still holds meaning.



Lemmatization

It is the process of reducing inflected or derived words to their base form by analyzing their morphology and using a vocabulary.



It generally offers better results where understanding the context is essential.

Preprocessing

Part-of-speech tagging is the process of marking a word in a text as corresponding to a particular part of speech (noun, verb, adjective and so on). It comes after the tokenization and the stemming/lemmatization processes and helps to understand the grammatical context in which the words are used.

DT	Determiner
JJ	Adjective
NN	Noun (singular)
NNS	Noun (plural)
NNP	Proper Noun
RB	Adverb
SYM	Symbol
VB	Verb (base)
VBD	Verb (past tense)

Some of the most commonly used part-of-speech tags.

Hello ! Are you a chatbot or a human ?

Word	Tag	Meaning
Hello	UH	Interjection
!	.	Punctuation
Are	VBP	Verb *(present, non-3rd person)*
you	PRP	Personal Pronoun
a	DT	Determiner
chatbot	NN	Noun *(singular)*
?	.	Punctuation

Antoine PALISSON

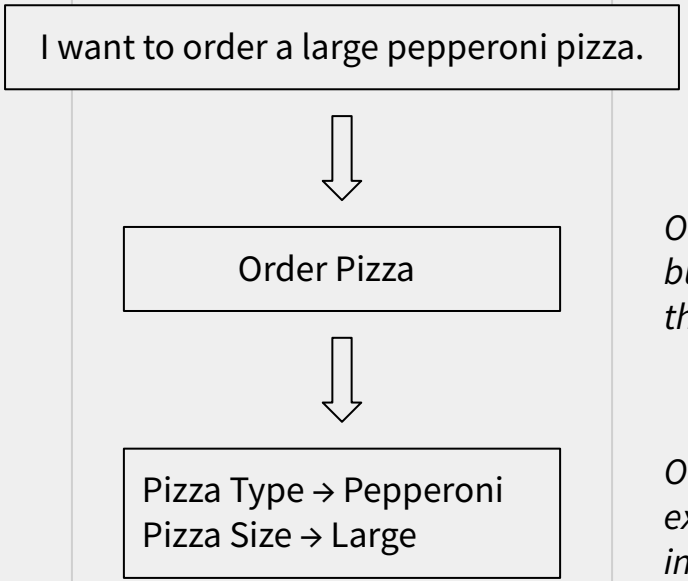
Intent Recognition

Intent recognition is a crucial process by which a chatbot determines what a user wants to accomplish based on their input. It is generally divided in three components:

Utterance
The entire sentence that the user inputs.

Intents
The user's goal according to their query.

Entities
Information that are relevant to the user's intent(s).



Other intents might exist but were not found in the existing utterance.

Other entities might exist but were not found in the existing utterance.

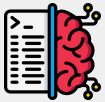
Intent Recognition

While intent recognition does not dictate the choice of chatbot type, the type of chatbot chosen can affect the way intent recognition is implemented and how its results are used.

Rule-based

They use explicit intent recognition methods, often relying on techniques like keyword matching or rule-based systems. Sometimes, Named Entity Recognition can be used.

Retrieval-based



They use a mix of explicit and implicit intent recognition that relies on Machine Learning techniques with some form of explicit intent classification as part of their process to select the appropriate response.

Generative

They usually have implicit intent recognition, as they generate responses in real-time based on the user's input using Deep Learning techniques.

Named Entity Recognition

Named Entity Recognition (NER) is a subtask of information extraction in Natural Language Processing that focuses on identifying and categorizing specific pieces of information, also called entities, in a text. These entities often represent a specific, unique concept such as:

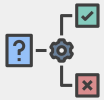
- Date
- Organization
- Product
- Product Type
- Location
- Job Title
- Person
- Capability

On August 8, 2025, the AI startup ChatTech launched Jarvis, their latest chatbot software, at a major tech event in Silicon Valley. The CEO, Dr. John Doe, stated that Jarvis can understand and respond in over 30 languages and can handle complex tasks better than their previous models.

Named Entity Recognition

Named Entity Recognition can be done in three different way:

Rule-based



These approaches use a set of predefined rules to identify and classify entities. The rules might be based on grammar, context, formatting, list of known words or patterns in the text.

Machine Learning

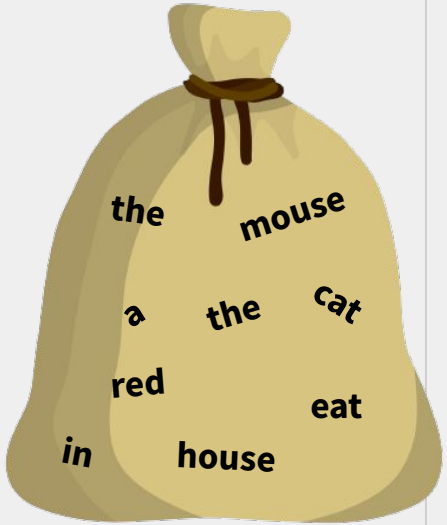
These approaches involve training a machine learning model to recognize entities based on labeled examples. Traditional machine learning algorithms can be used but will not be as efficient as deep learning ones.

Deep Learning

Deep Learning models such as RNNs or Transformers can learn to recognize entities based on the context in which they appear, making them more flexible and accurate than other methods.

Further Preprocessing

Depending on the type of chatbot used, further preprocessing could be involved to transform the user's utterance before sending it to a model. Methods such as **Bag of Words**, **Term Frequency-Inverse Document Frequency** (TF-IDF), and **Word Embeddings** are often used in text preprocessing steps.



Bag of Words is used to convert text data into numerical features. It creates a vocabulary of all the unique words in the text, and then represents each word by a vector indicating its frequency in the vocabulary.

Bag of words does not capture any information about the order of words.

Further Preprocessing

TF-IDF improves the Bag of words model by weighting each word's frequency by how often it appears in the text. Words that appear frequently in a single document but rarely in the rest of the dataset are given more weight. TF-IDF does not capture any information about the order of words.

Significance of a word t in a document d given all the documents D

$$TFIDF_{t,d,D} = TF_{t,d} * IDF_{t,D}$$

Frequency of a word t in a document d

$$TF_{t,d} = \frac{\text{number of times the word t appears in d}}{\text{total number of words in d}}$$

Significance of a word t in all the documents D

$$IDF_{t,D} = log\left(\frac{\text{total number of documents}}{\text{number of documents containing the word t}}\right)$$

Further Preprocessing

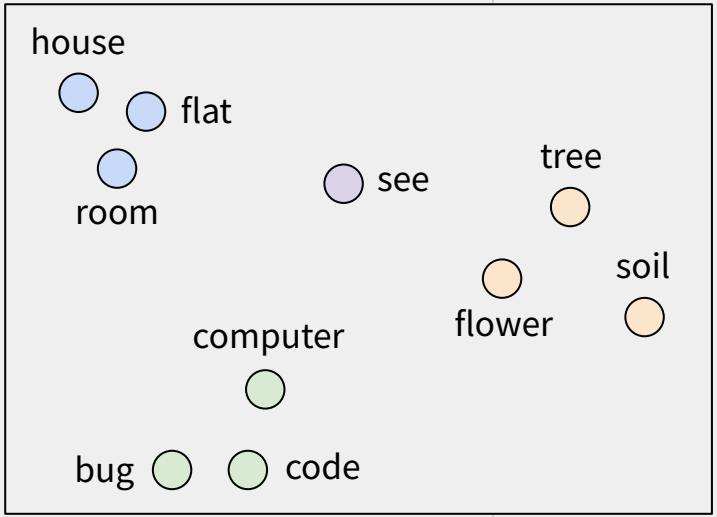
Let's imagine three texts, A, B and C that contain 10 words each.

The TFIDF of the words "the", "eat", "cat" and "is" is calculated below.

	TF			IDF	TFIDF		
Words	***A***	***B***	***C***		***A***	***B***	***C***
the	2/10	1/10	3/10	0	0	0	0
eat	0	0	1/10	0.48	0	0	0.048
cat	2/10	1/10	0	0.30	0.06	0.03	0
is	1/10	2/10	1/10	0	0	0	0

Further Preprocessing

Unlike Bag of words and TF-IDF, **word embeddings** represent words as vectors in a high-dimensional space where the semantic relationships between words are captured, such as synonyms being close together and words appearing in similar contexts also being close.



A 2D space where the words are projected.
The original vector space could much bigger.

The house is red

The	**[-0.8 0.7 … -0.1 1.8]**
House	**[0.2 0.5 … 0.9 -2.1]**
is	**[-0.1 1.6 … -0.4 0.1]**
red	**[0.4 -0.5 … -1.3 0.7]**

Contextual Appropriateness

Once the chatbot understands the user's intent, the next step is generating a contextually appropriate response.

First, to ensure **contextual appropriateness**, it's important for the chatbot to dynamically ask for further details or clarification based on the context and the user's input. This helps prevent misunderstandings and error, and ensures that the chatbot gather all the necessary information to fulfill the user's request.

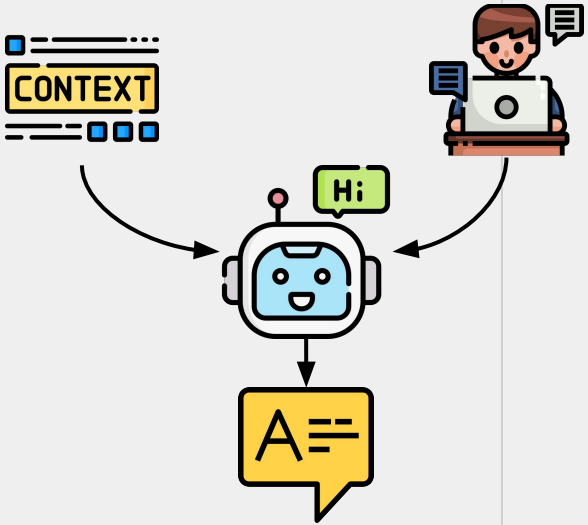
But how does the chatbot know when to ask for details ?

Generally, once the intent of a user's message is detected, it's common to have certain mandatory entities that the chatbot needs in order to fulfill the intent. Thus, determining which entities are required for each intent and implementing the logic to gather these entities is a crucial part of designing a chatbot.

Contextual Appropriateness

Additionally, to ensure **contextual appropriateness** the chatbot should take into account the entire history of the conversation, not just the last message.

For instance, if the user asked about the weather in Paris earlier and now asks, "And what about tomorrow?", the chatbot should understand that the user is asking about the weather in Paris for the next day.



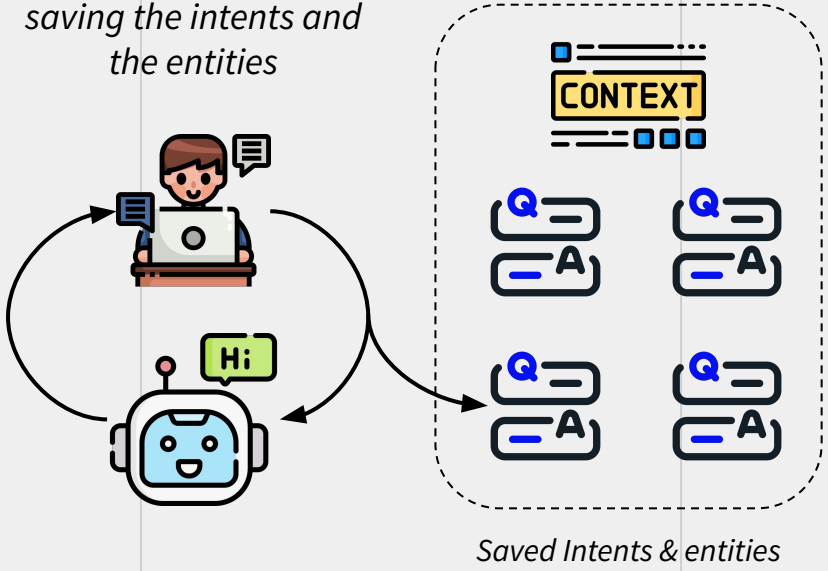
Maintaining context and conversation history is easy for generative chatbots because they work by understanding the context of the user's utterance.

However, it is a challenge in rule-based or retrieval-based chatbots since these types of models don't inherently understand the conversation's context.

Contextual Appropriateness

Both rule-based and retrieval-based chatbots track the intents and entities recognized throughout the conversation to avoid duplicating questions or actions, and to provide a coherent and contextually appropriate dialogue.

Building the context by saving the intents and the entities



Saved Intents & entities

Contextual Appropriateness

Sometimes, the chatbot doesn't understand the user's input. When this happens, it is generally recommended to implement a **fallback strategy** to help maintain a positive user experience even when the chatbot reaches the limits of its capabilities. Some of the most common strategies include:

I'm sorry, I didn't understand that. Could you please rephrase or provide more information?

Clarification request

Suggestions

I'm sorry, I didn't quite understand. Were you asking about [Topic A] or [Topic B]?

I'm sorry, I'm currently unable to provide a answer. However, I can provide information on [Topic A] or [Topic B]. Which one would you prefer?

Topic Switch

Hand-off to human

It seems like I'm not able to provide you help on that subject. Would you like me to connect you with a member of our customer support team?

Continuous Learning

Chatbots can improve over time through learning processes. This process generally relies on the ongoing collection and analysis of conversational data, with improvements to the chatbot's performance coming through several key mechanisms:

Supervised Learning

Chatbots, particularly those using ML models, can be retrained on newer data that includes recent interactions.

Feedback Loop

Chatbots can be designed to incorporate user feedback by asking them to rate the helpfulness of its responses.

Active Learning

Chatbots identify cases where it's unsure how to respond which are then reviewed by human operators who provide the correct responses.

Rule-based Chatbots

Definition

A **rule-based chatbot** is a type of chatbot that operates based on predefined rules on which it was initially programmed. These bots can only respond to specific commands and use a tree-like model of knowledge, meaning they follow a "if user says A, then respond with B" logic.

Advantages

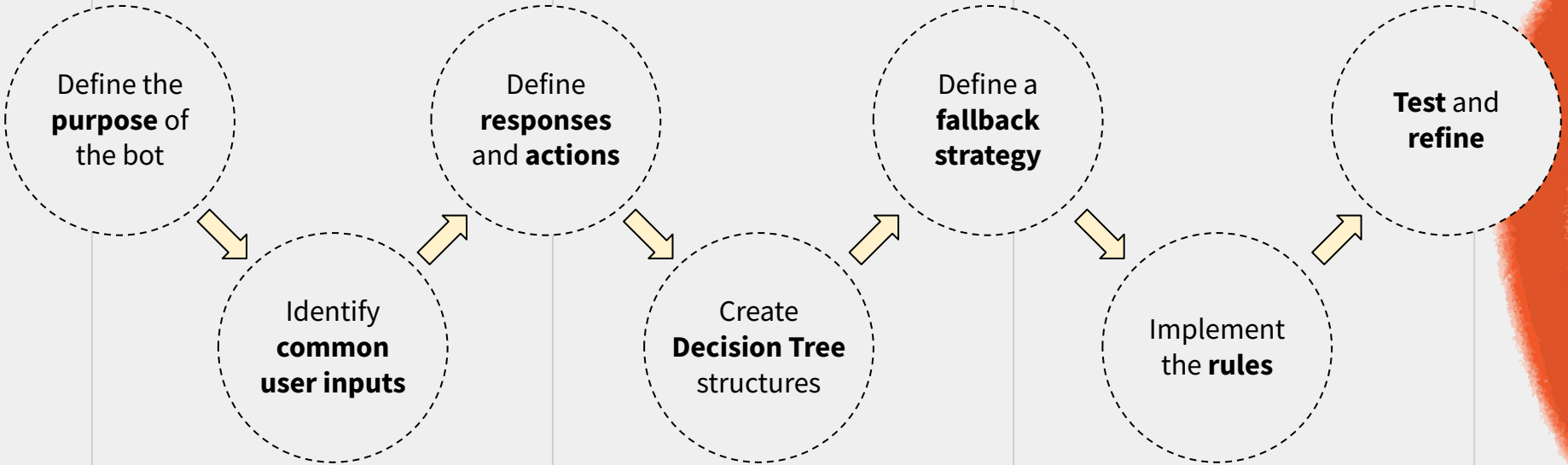
The main advantage of rule-based bots is their predictability, as their behavior is completely under the control of their designers. They are a good fit when the task or the domain is well-defined and doesn't need a lot of improvisation.

Drawbacks

Rule-based chatbots are unable to process and generate responses to queries beyond their programming. They can't handle complex queries and don't have the capability to learn from past interactions.

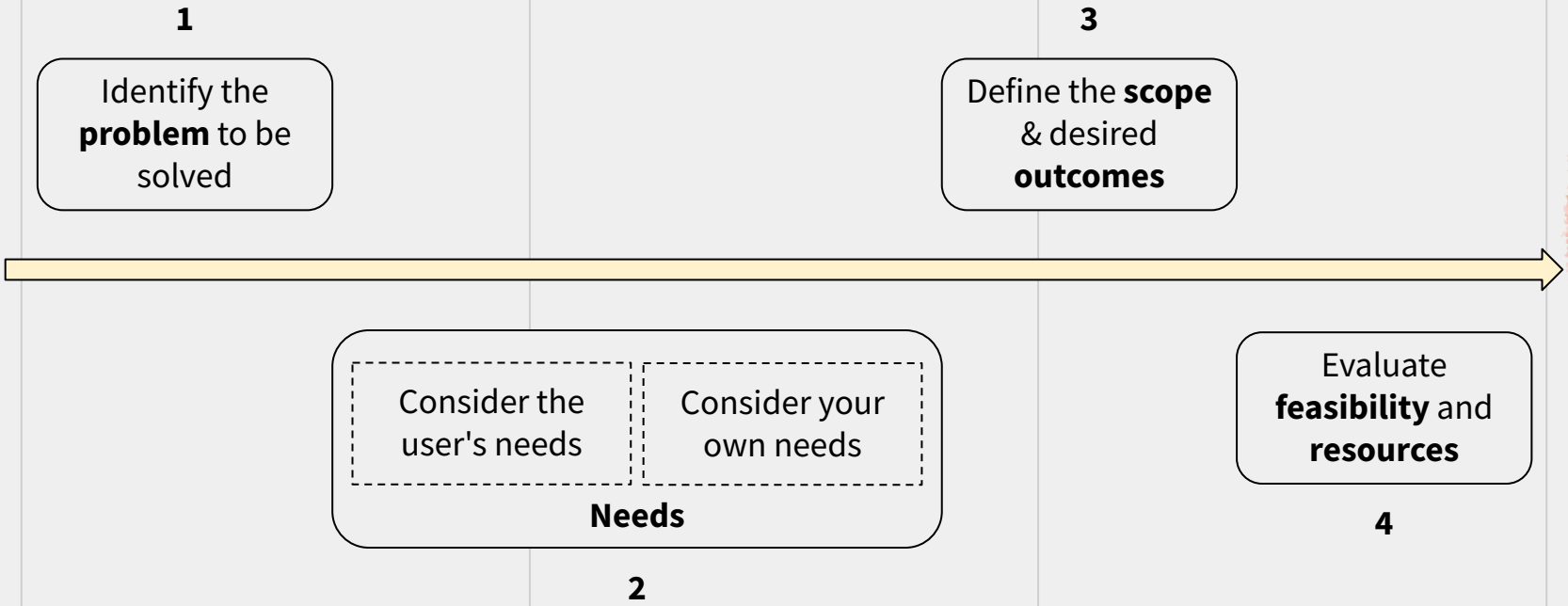
Building a chatbot

Building a good **rule-based chatbot** efficiently involves a blend of defining the bot's purpose, understanding your users, drafting an interactive dialog flow, implementing, testing, and refining the bot.



Defining the purpose

The purpose of the chatbot is its fundamental reason for being, and this purpose will guide all subsequent decisions regarding its design and implementation.



Common User Inputs

The purpose of the chatbot is its fundamental reason for being, and this purpose will guide all subsequent decisions regarding its design and implementation.

Analyze existing interactions

emails, live chat logs, social media messages …

Brainstorm and role play scenarios

think from the user's perspective

Categorize by Intent

classify every user inputs by intent

Exercise

You are expected to collaborate in pairs to develop the initial structure for a rule-based chatbot using Python. The primary focus of this task is the conceptualization stage by **defining the purpose of your chatbot** and **finding the common user inputs** as seen in the course.

> *Go to the Common User Input 🔗 section to know more about the methodology.*

Tips : Ask yourself, what service or information should the chatbot provide? How will it be helpful to its users? The clarity of this purpose will guide your design and functionality choices. Then, Think about the typical inquiries or statements your chatbot might receive in the course of fulfilling its defined purpose.

Antoine PALISSON

Responses & Actions

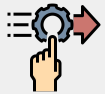
Defining responses and actions for a chatbot involves deciding how the chatbot will react to user inputs. Based on the user's input, which have been already identified in previous steps, the chatbot should provide an appropriate response or perform an action.

Map Inputs to Responses

For each common user input, decide what the bot's response should be.

Define Actions

Actions might include looking up information, storing user data, initiating a process, or guiding a user through a sequence of steps.



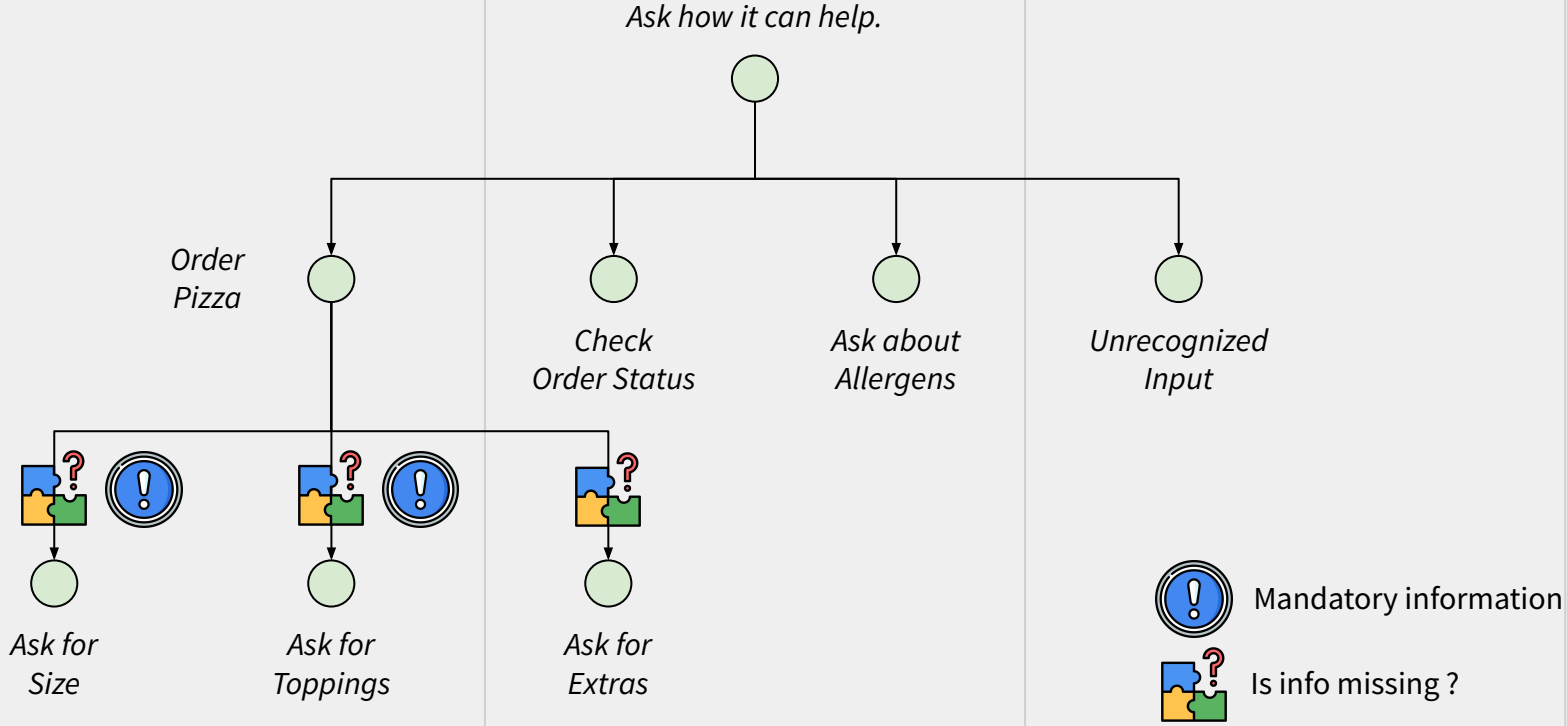
Create a Response Bank

Create varied responses for the same intent to avoid sounding repetitive.

Antoine PALISSON

Decision Tree Structures

Creating **decision tree structures** involves planning out how the chatbot will navigate through different scenarios based on the user's input. Here is a small example:



Exercise

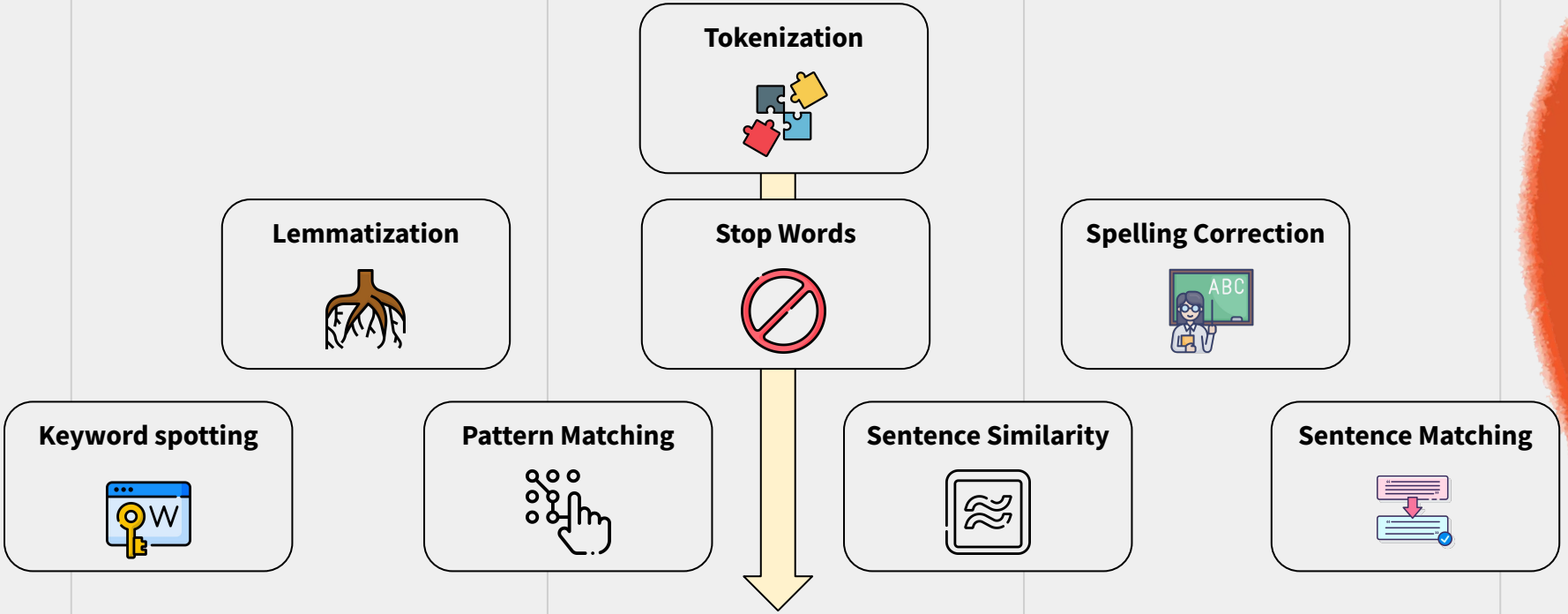
You are expected to **define the responses and actions** of your chatbot, including its fallback strategies. To accomplish this, you will design and **build a detailed decision tree structure** that effectively maps the user inputs to the chatbot's corresponding responses.

> *Go to the Decision Tree Structures 🔗 section to see an example of a decision tree structure.*

Tips : Start simple by sketching a basic flow for common interactions your chatbot might encounter. You can then add complexity, such as additional branches to the tree or more sophisticated responses. Then, design your decision tree to account for the different ways a user might express their intent. Of course, feel free to add some personality to your chatbot!

Intent Recognition

In a **rule-based chatbot**, understanding user utterances and answering correctly primarily involves mapping user inputs to predefined responses using Natural Language Processing. The overall process could look like:



Exercise

Exercise I

For this exercise, you are to **implement a rule-based chatbot using Python without the aid of any external libraries**. Your chatbot should be capable of delivering predefined responses based on specific user inputs, in line with the decision tree structure you've already created.

Exercise II

For this exercise, you are to **test your chatbot's functionalities** by asking another group to play with it. In the meantime, you are to try the chatbot of another group in order to **find its weaknesses !**