

# Developing Chatbots in an Evolving AI Landscape

by Guillaume Slevan-Tremblay



# Guillaume Slevan-Tremblay

Senior Conversation Bot Engineer

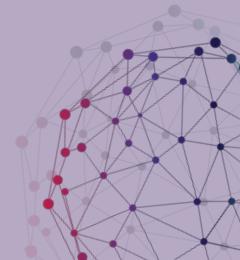
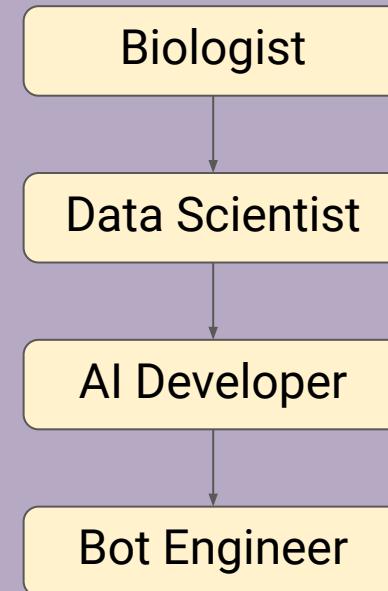
*also titled internally as...*

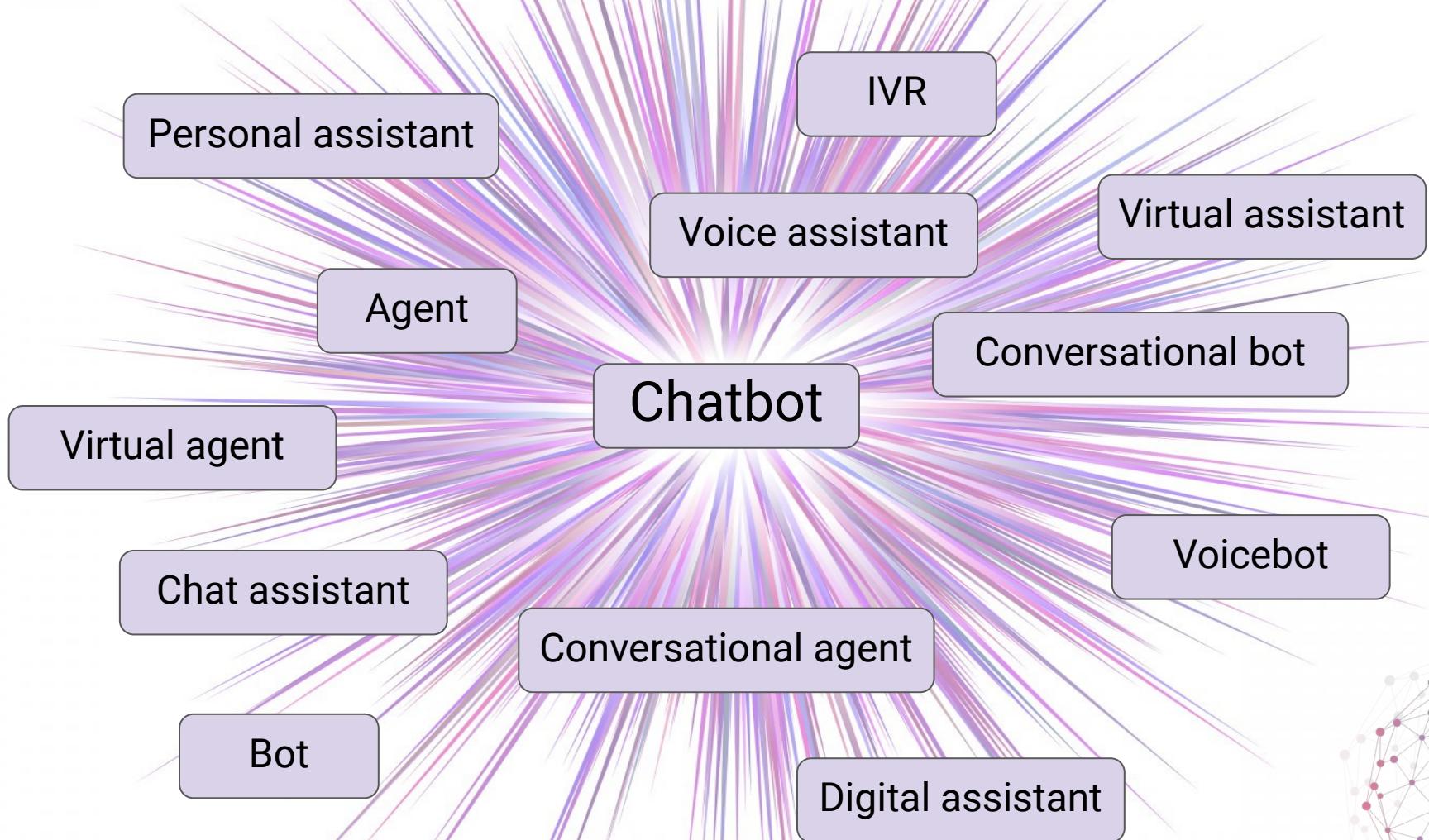
“Conversation Architect”

“Conversation Designer”

“Bot developer”

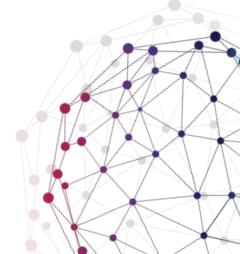
## Journey to Conversational AI

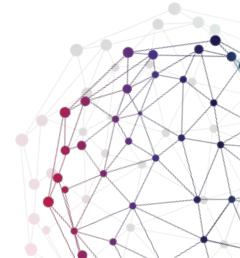




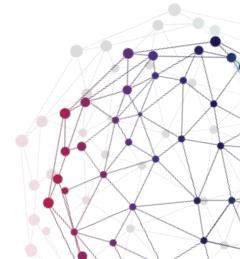
recognition Through technologies  
sorcerers discourse **interactive** provide  
understand engage **digital** intent  
knowledge information intricately **assistants**  
information **virtual** symphony emissaries  
prowess natural **linguistic** adorned  
virtuosity **agents** Chatbots decipher leveraging  
employ realms NLP systems dialogues  
queries crafted more lexical bots sphere  
language simulate core art speech  
tirelessly realms **conversational** nuances devoted  
**human** into array **chatbots** even responses  
interactions learning realm marvels sentient ML  
sophisticated realm wonders **chatbots** respond  
remarkable communication chatterbots companions technological  
support conversation context  
domain-specific

# Do you like interacting with chatbots?

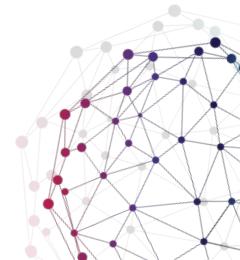
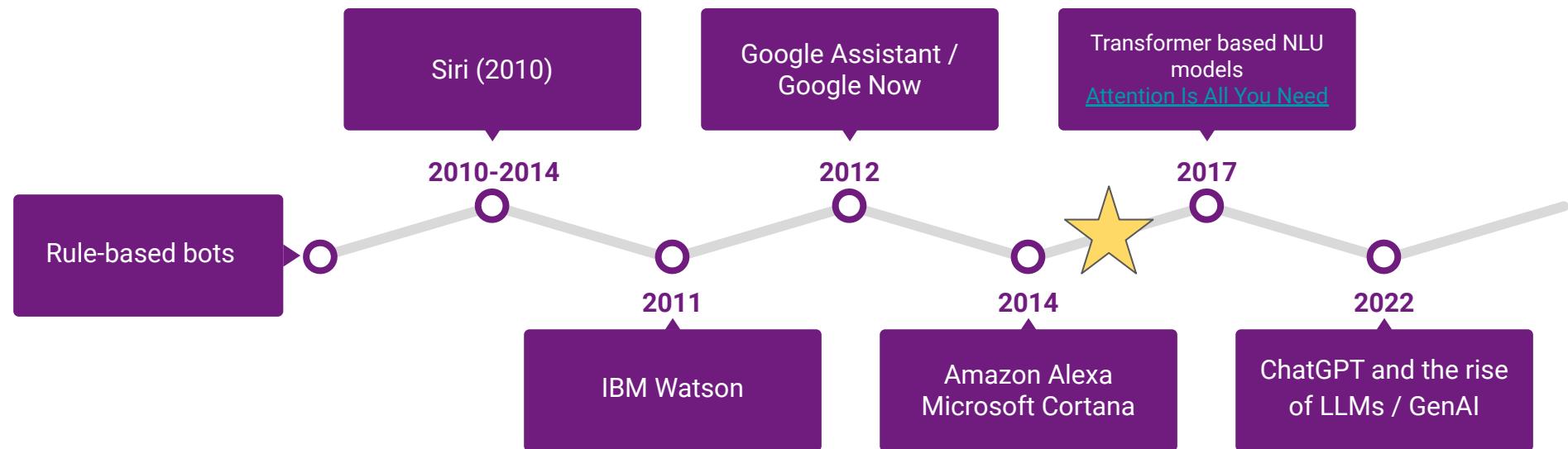




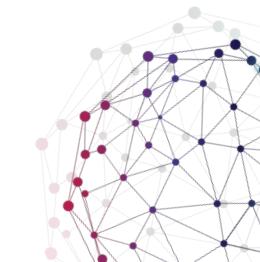
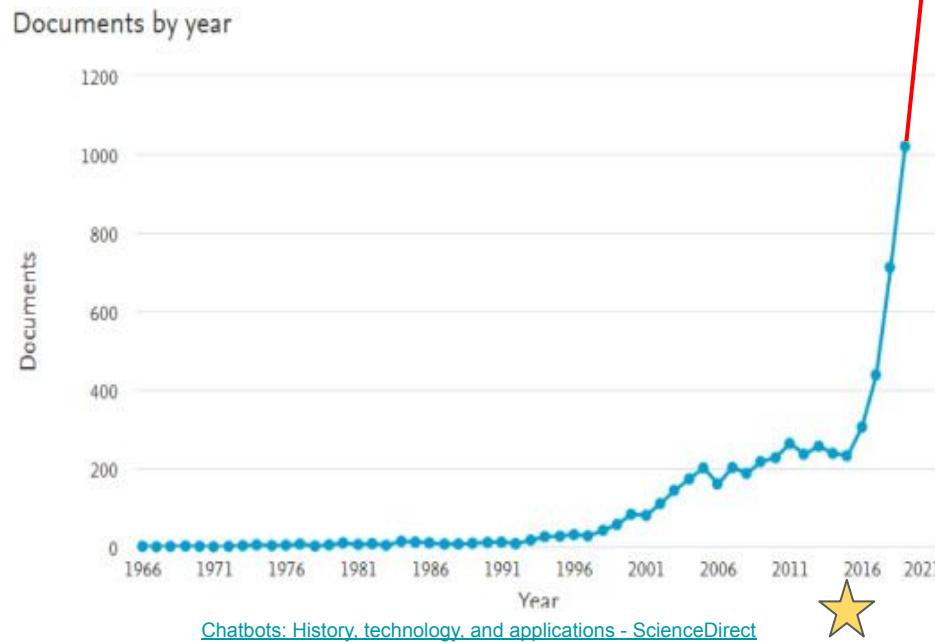
# History of Chatbots



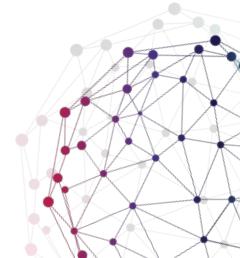
# History of chatbots since 2010

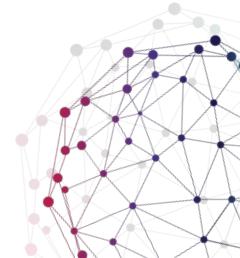
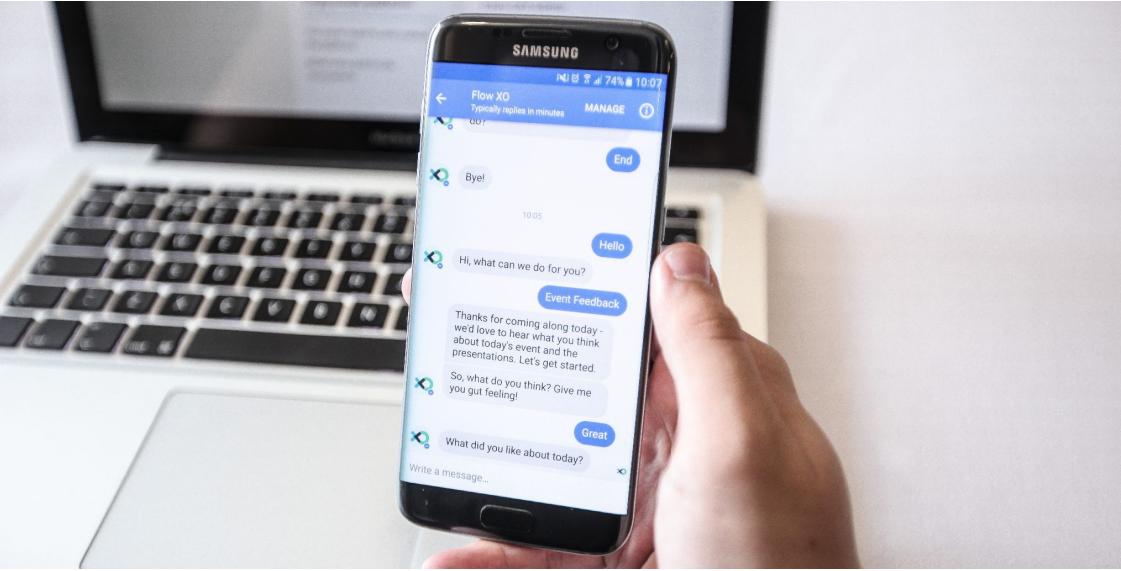


# Rise in recent publications



# What are chatbots?





 Reset Thread

 Dark Mode

 OpenAI Discord

 Learn More

 Log out

# ChatGPT



## Examples

"Explain quantum computing in simple terms"



## Capabilities

Remembers what user said earlier in the conversation



## Limitations

May occasionally generate incorrect information

"Got any creative ideas for a 10 year old's birthday?"

Allows user to provide follow-up corrections

May occasionally produce harmful instructions or biased content

"How do I make an HTTP request in Javascript?"

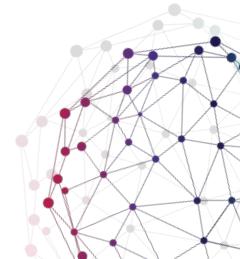
Trained to decline inappropriate requests

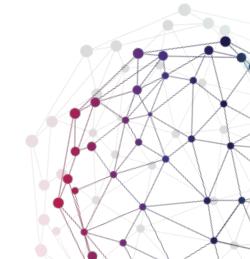
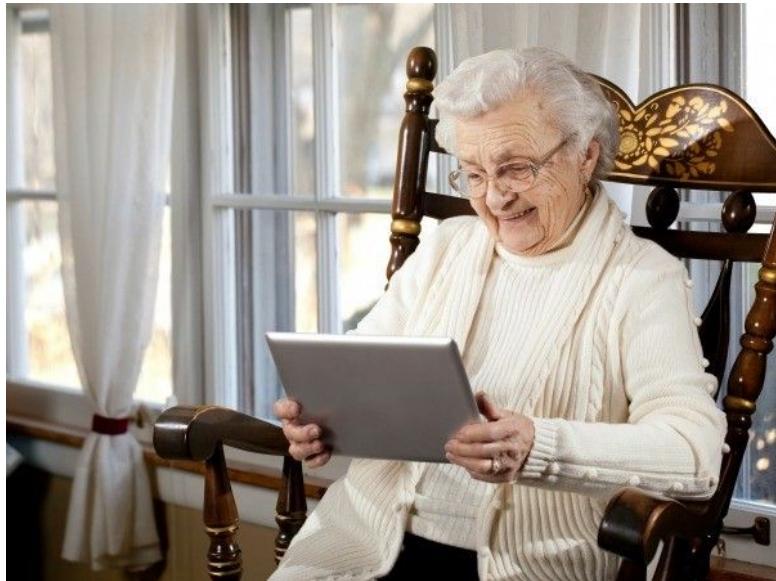
Limited knowledge of world and events after 2021

 >

Free Research Preview: ChatGPT is optimized for dialogue. Our goal is to make AI systems more natural to interact with, and your feedback will help us improve our systems and make them safer.

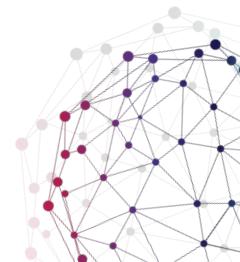






# Categories of bots

- Rule-based bots
- Frequently asked questions (FAQ-bot)
- Machine learning and NLP chatbots
- NLU (transformer)
- LLM / generative
- Knowledge-based



# Chat vs Voice Bots



## Chatbots

- Text-based
  - Messaging platforms
  - Website integrations
  - Mobile apps
- Enhanced visual UI
  - Images, carousels, carts, maps
  - URL links
  - Lists of items
- Option selection ability
  - Buttons, chips
  - Dropdown
- Integration of web components



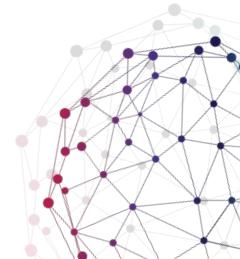
## Voicebot / Voice Assistant

- Voice-based
  - Phone calls
  - Home assistant
  - Car assistant
- Concise / shorter conversations
- DTMF (dual tone multi frequency)
- SSML markup
- Automatic speech recognition (ASR)
- Speech to text (STT)
- Text to speech (TTS)
- SMS to send URL links or images



# Chatbot Terminology

- Intent
- Entity
- Utterance
- Prompt
- Agent
- Flow
- Parameter
- User journey
- Happy / sad path



# Chatbot Terminology

- Intent
- Entity
- Utterance
- Prompt
- **Agent**
- Flow
- Parameter
- User journey
- Happy / sad path
- Fallback



An agent refers to the chatbot or virtual assistant that interacts with users.

It is the program or system responsible for understanding user inputs, processing them, and generating appropriate responses based on predefined intents and entities.



# Chatbot Terminology

- Intent
- Entity
- Utterance
- Prompt
- Agent
- **Flow**
- Parameter
- User journey
- Happy / sad path
- Fallback



A flow refers to the sequence or structure of a conversation between the user and the chatbot (design).

It defines the order in which prompts, utterances, and responses are presented to the user, ensuring a logical and coherent conversation.



# Chatbot Terminology

- Intent
- Entity
- Utterance**
- Prompt
- Agent
- Flow
- Parameter
- User journey
- Happy / sad path
- Fallback



An utterance is a single statement or input made by the user. It represents what the user says or types while interacting with the chatbot.

Each utterance is associated with an intent and may contain one or more entities.



# Chatbot Terminology

- Intent
- Entity
- Utterance
- **Prompt**
- Agent
- Flow
- Parameter
- User journey
- Happy / sad path
- Fallback



A prompt is a message or question presented by the chatbot to solicit a response from the user. It helps guide the conversation and elicit specific information.

Prompts can be used to ask for missing entities or clarify user intent.



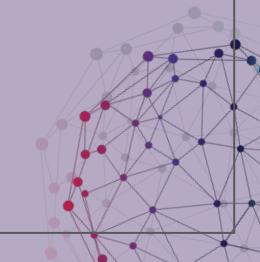
# Chatbot Terminology

- Intent
- Entity
- Utterance
- Prompt
- Agent
- Flow
- Parameter
- User journey
- Happy / sad path
- Fallback



An intent represents the goal or purpose behind a user's input. It refers to what the user wants to accomplish or the action they want the chatbot to perform.

For example, if a user asks, "What's the weather today?", the intent could be "weather\_forecast."



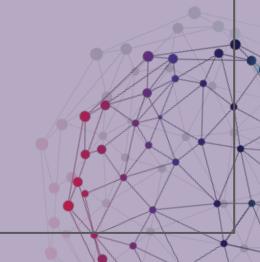
# Chatbot Terminology

- Intent
- **Entity**
- Utterance
- Prompt
- Agent
- Flow
- Parameter
- Variable
- User journey
- Happy / sad path
- Fallback



An entity refers to a specific piece of information within a user's input that is relevant to fulfilling the intent. Entities provide context or details about the user's request.

In the previous example, the entity could be "today," indicating the specific time frame for the weather forecast.



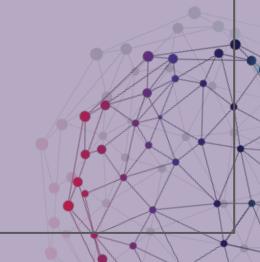
# Chatbot Terminology

- Intent
- Entity
- Utterance
- Prompt
- Agent
- Flow
- **Parameter**
- User journey
- Happy / sad path
- Fallback



A parameter is a variable that holds a value and is used to pass information within a system or program.

In the context of chatbots, parameters are often associated with intents and entities and are used to capture and store user-provided information for further processing or response generation.



# Chatbot Terminology

- Intent
- Entity
- Utterance
- Prompt
- Agent
- Flow
- Parameter
- **User journey**
- Happy / sad path
- Fallback



User journeys refer to the paths or routes a user can take during a conversation with a chatbot.

They encompass various interactions, intents, and prompts that guide the user towards achieving their desired outcome.

User journeys can vary depending on the user's input, choices, and the chatbot's capabilities.



# Chatbot Terminology

- Intent
- Entity
- Utterance
- Prompt
- Agent
- Flow
- Parameter
- User journey
- **Happy / sad path**
- Fallback



The happy path refers to the ideal or successful sequence of interactions between a user and a chatbot to accomplish their goal.

It represents the smoothest and most straightforward user journey, where the chatbot understands the user's intent accurately and provides the desired information or performs the requested action without any issues.



# Chatbot Terminology

- Intent
- Entity
- Utterance
- Prompt
- Agent
- Flow
- Parameter
- User journey
- Happy / **sad path**
- Fallback



The sad path, also known as an alternative or error path, refers to the sequence of interactions when a user encounters an issue or the chatbot fails to understand the user's intent properly.

It represents deviations from the happy path, such as when the chatbot asks for clarification, fails to provide a response, or encounters an error.

The sad path usually involves error handling or fallback mechanisms to address user issues or misunderstandings.



# Chatbot Terminology

- Intent
- Entity
- Utterance
- Prompt
- Agent
- Flow
- Parameter
- User journey
- Happy / sad path
- **Fallback**

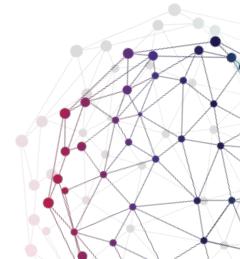


A fallback is a mechanism or strategy used when the chatbot encounters a user input or situation that it cannot handle or understand.

It serves as a backup or alternative approach to provide a meaningful response or assistance when the chatbot's regular processing or understanding fails



# Chatbot Development



# Transcript labelling

Agent: Thank you for calling the telecom support center. My name is Mark. How may I assist you today?

Client: Hi Mark, I need to know the amount due on my bill and the payment deadline.

Intent

Agent: Of course, I'll be happy to help you with that. May I have your account number or phone number, please?

Client: Sure, my account number is **123456789**

Entity

Agent: Thank you for providing that information. Let me check your account for you. It seems that your current bill amount is \$75.42.

Client: Great, thanks for the update. Could you also let me know by which date I should pay the bill?

Agent: Absolutely. Your payment due date is **June 15th, 2023**

Parameter

Client: Perfect. I'd like to make the payment right away. Can I pay it with my credit card directly through this chat?

Agent: Yes, certainly. We offer secure credit card payments for your convenience. Could you please provide me with the credit card details for the transaction?

Client: Sure, my credit card number is 1234 5678 9012 3456, expiration date is 05/25, and the CVV is 123.

Agent: Thank you for providing the details. To ensure the security of your information, I will process the payment through our encrypted payment gateway. One moment, please.

Client: Take your time. I appreciate your assistance.

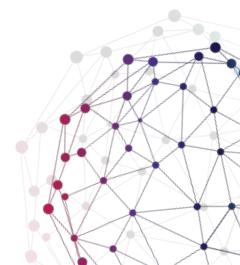
Agent: Thank you for your patience. I have successfully processed your payment of \$75.42 using the credit card ending in 3456. You will receive a confirmation email shortly.

Client: That's great! Thank you for your help, Mark. Is there anything else I need to do?

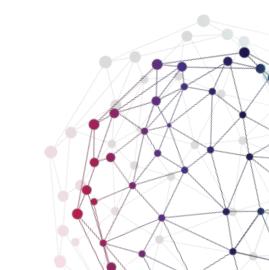
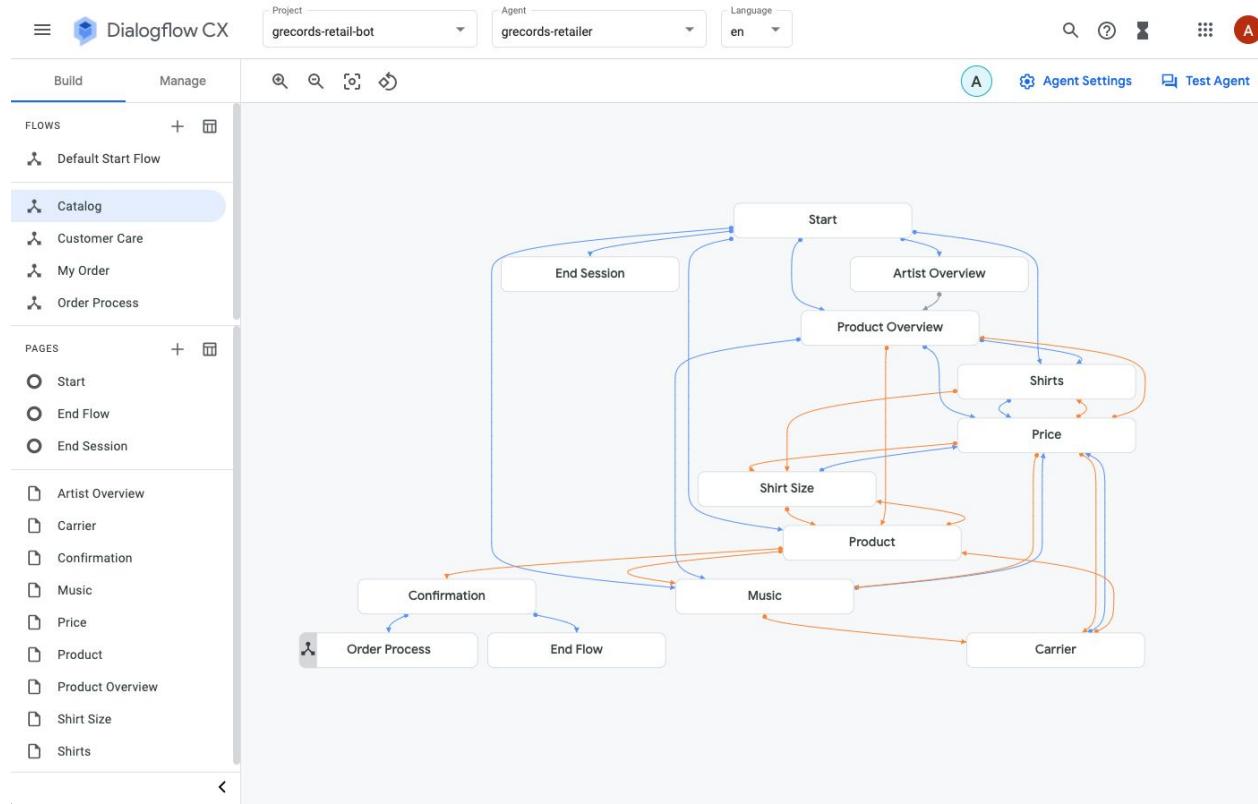
Agent: You're welcome! Your payment has been successfully processed, and you're all set. If you have any other questions or concerns, feel free to ask. Otherwise, you're good to go.

Client: Alright, Mark. Thank you again for your help. Have a great day!

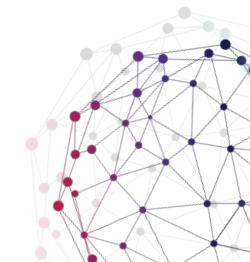
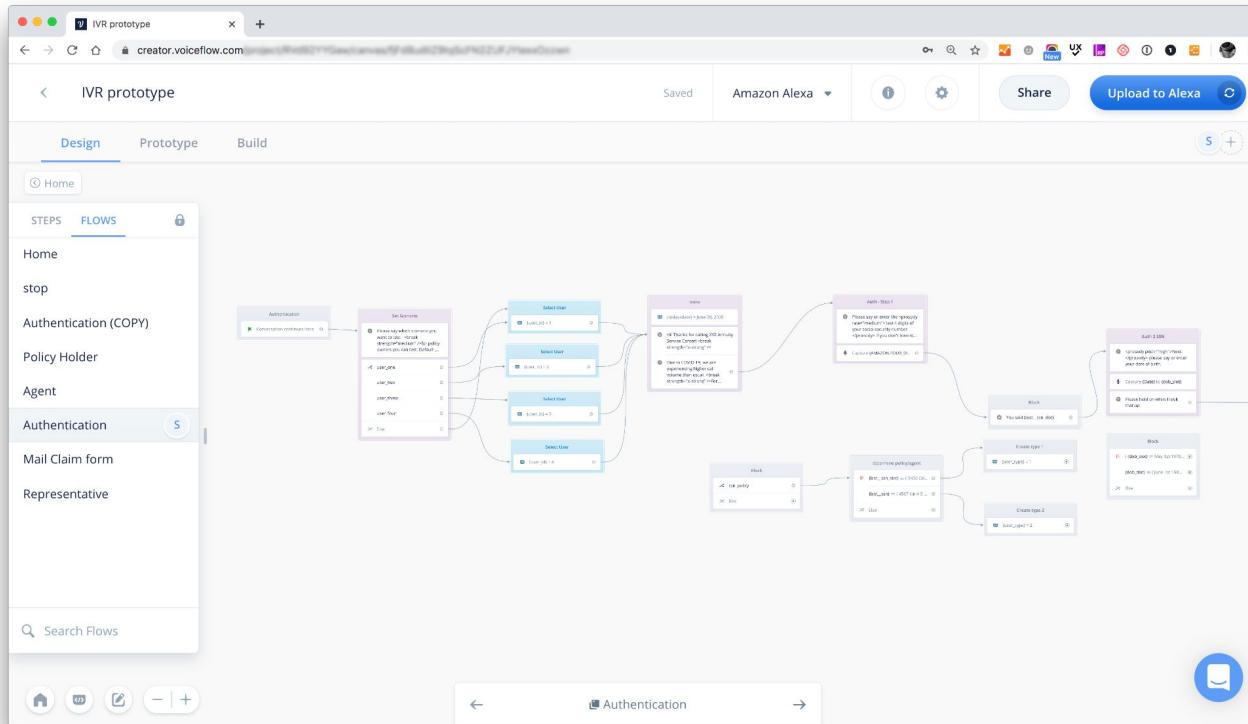
Agent: You're welcome! It was my pleasure assisting you. Have a wonderful day too! Goodbye!



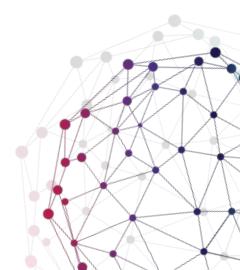
# Chatbot framework - Dialogflow



# Chatbot framework - Voiceflow



# Chatbot framework - RASA



A screenshot of Visual Studio Code showing a RASA project structure and a terminal window.

The Explorer sidebar shows the project structure:

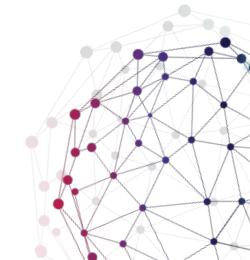
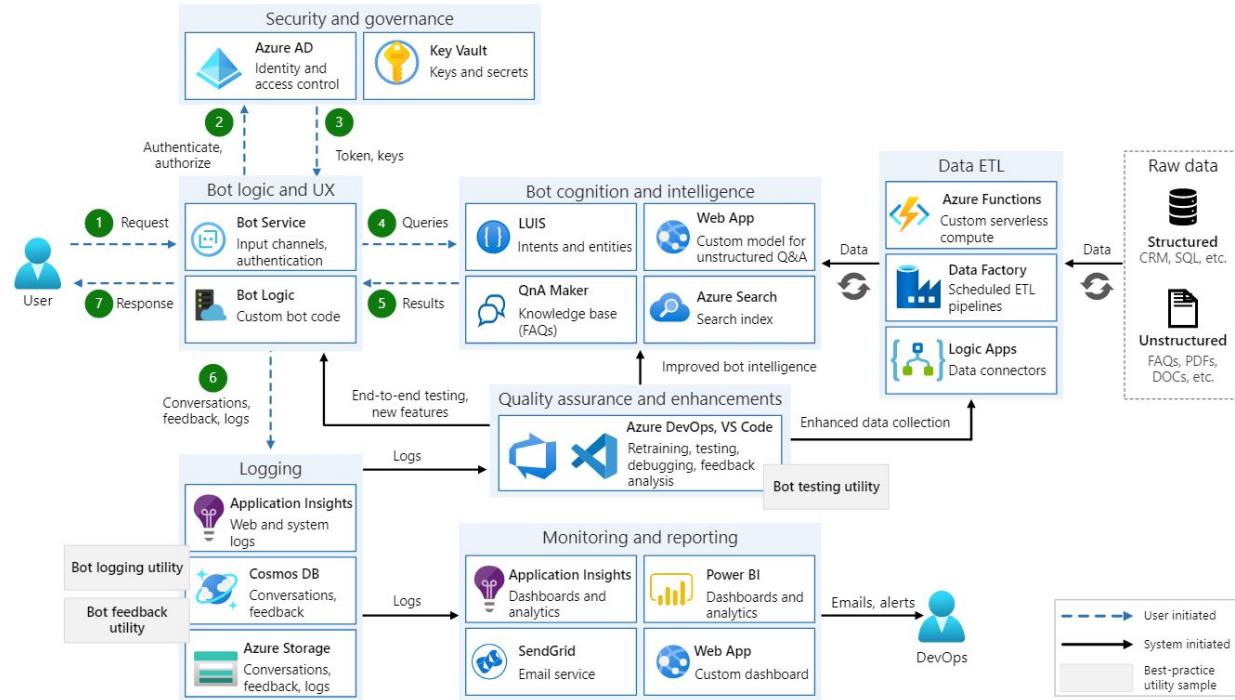
- HELLORASA
- |.pycache\_
- .vscode
- settings.json
- data
  - nlu.md
  - stories.md
- models
- tests
- \_init\_.py
- actions.py
- ! config.yml
- ! credentials.yml
- ! domain.yml
- ! endpoints.yml
- ① README.md

The Terminal tab shows the RASA shell output:

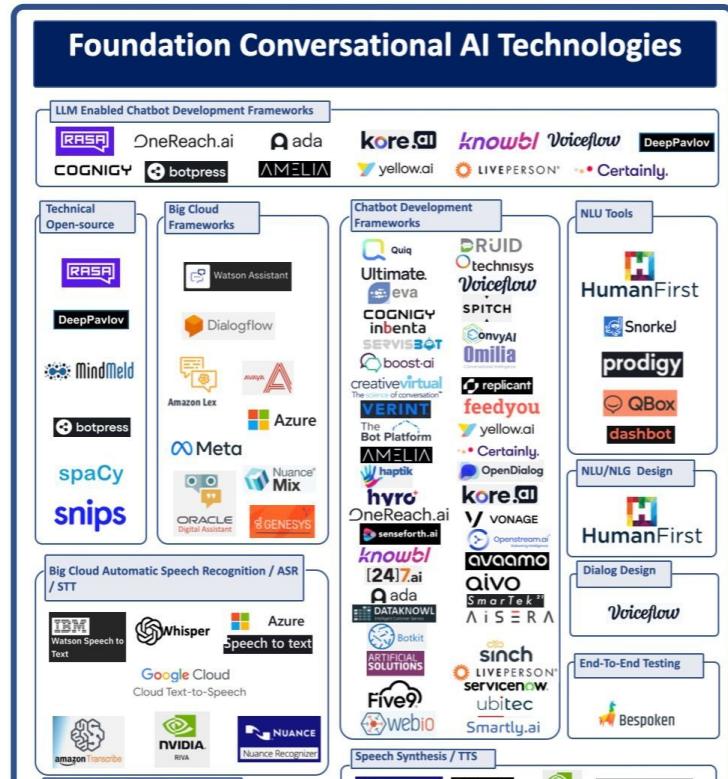
```
(hellorasa) C:\Users\xtago\Alpha\Projects\HelloRasa>rasa shell
2020-04-05 12:13:29 INFO    root - Connecting to channel 'cmdline' which was specified by the '--connector' argument. Any other channels will be ignored. To connect to all given channels, omit the '--connector' argument.
2020-04-05 12:13:29 INFO    root - Starting Rasa server on http://localhost:5005
Bot loaded. Type a message and press enter (use '/stop' to exit):
Your input -> Hello!
Hey! How are you?
Your input -> Great, thanks for asking :)
Great, carry on!
Your input -> bye now
Bye
Your input -> 
```

Bottom status bar: Ln 1, Col 1 | Spaces: 2 | UTF-8 | LF | Markdown | Prettier | 🔍 | ⚙️

# Bot Solution Architecture



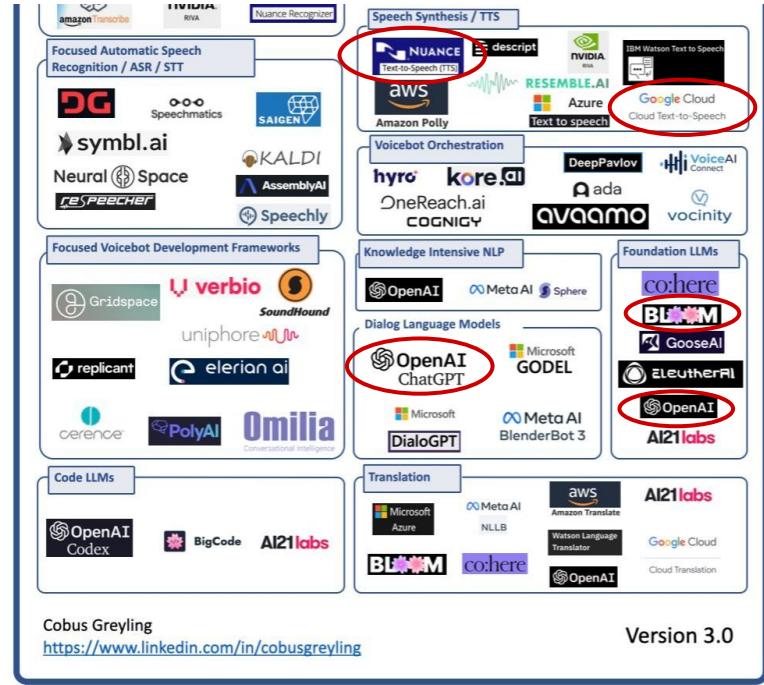
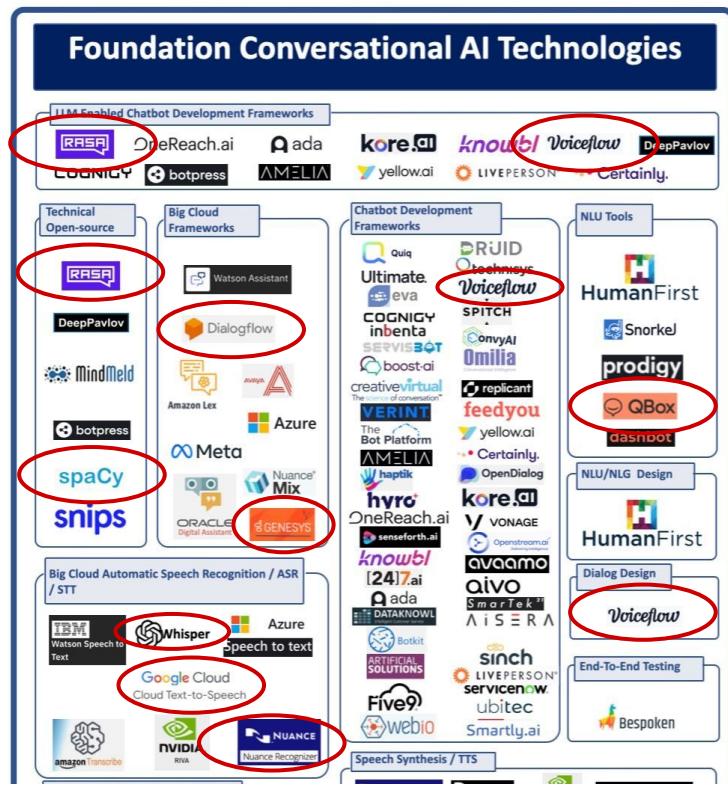
# Bot building technologies in 2023



Saturated landscape!



# Bot building technologies in 2023



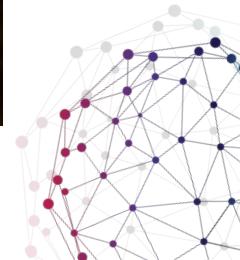
Saturated landscape!





**OH, YOU HAVE A CHATBOT**

**YOU MUST HAVE A BIG TECH TEAM**



# Realities of the chatbot ecosystem

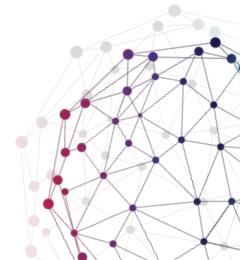
Biggest industries using bots are consumer focused

- Ex: banking, telecom, retail, healthcare, insurance, hospitality

Chatbot development is often outsourced to consulting firms

Top (big) organizations are behind

- Can't move fast enough
- Slow change management
- Technologies are 3-4 years behind today's standards
- Legal and marketing limitations



# Who works on bots?

## TECHNICAL

Bot Engineer (chat/voice)

AI Engineer / Developer

Data Scientist

Software Engineer

ML Engineer

NLP/NLU Engineer

Cloud Engineer / Architect

Devops Engineer

Front End Engineer

Web Developer

## NON-TECHNICAL

Dialog Designer

Copywriter

AI Trainer (labeler)

Linguist

UX Designer

Content Designer

Translator

Conversation(al) Designer

Conversation(al) Architect

## SUPPORTING

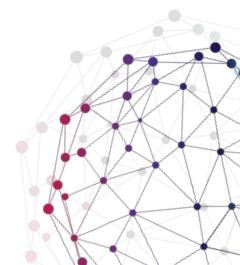
Business Analyst

Test Automation Engineer

Quality Analyst

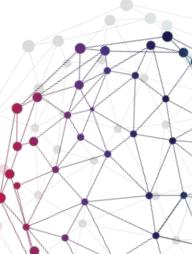
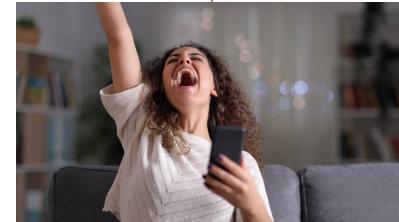
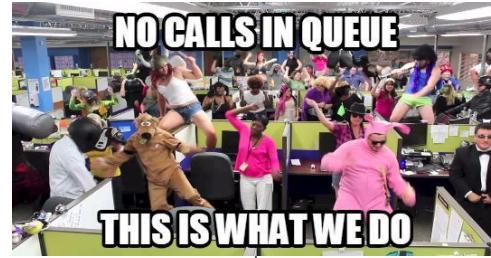
User Acceptance Tester

+ Agile team



# Why building a chatbot?

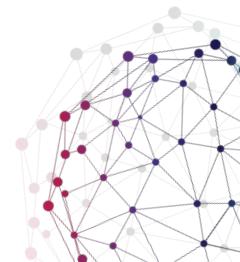
1. Customer support
  - o Improve and scale customer service
  - o Provide 24/7 availability
  - o Streamline information retrieval
2. User experience (UX)
  - o Increase task completion efficiency
  - o Add self-serve modules for customers to leverage
  - o Offer personalized recommendations
  - o Centralized customer interactions
3. Business
  - o Reduce cost associated to running call centers
  - o Gather customer insights
  - o Capture leads and generate sales
  - o Reduce customer churn



# Scoping criterias

How to scope the development of a chatbot?

- Company size
- FAQ or Transactional
- Problem to solve
- Time to market
- Development budget
- Number of users
- Backend services / databases
- Deployment framework / UI
- ...

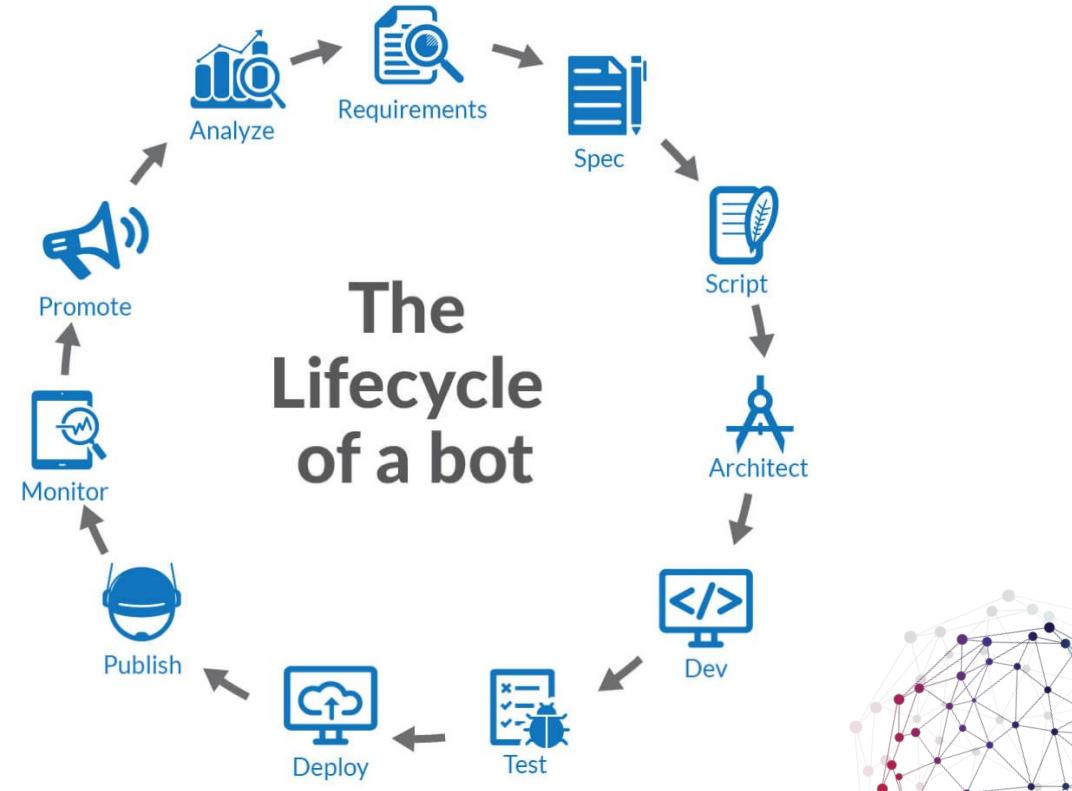


# Chatbot development life cycle

Many variations of this cycle

Depends on:

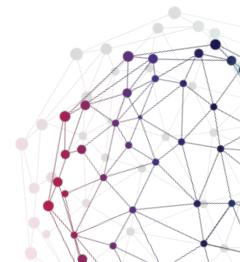
- Size of the agent
- Team composition
- Resources
- Integrations
- Connected services
- Frameworks
- etc.



# Linear steps to chatbot development



1. Collect data / transcripts
2. Build taxonomy / label intents and entities
3. Train and fine-tuned NLU model
4. Design the conversations (“user experiences”)
5. Integrate with backend databases and services
6. Perform quality analysis at design level
7. Complete user acceptance testing
8. Release to production
9. Monitor performance and metrics
10. Identify and pr areas of improvement
11. Provide solutions as development features

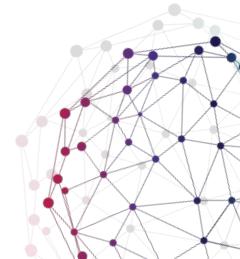


# UX Challenges

Bot users are historically conditioned to have poor and frustrating experiences

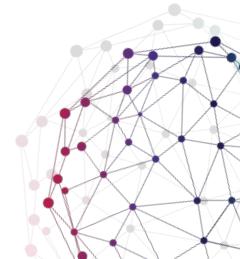


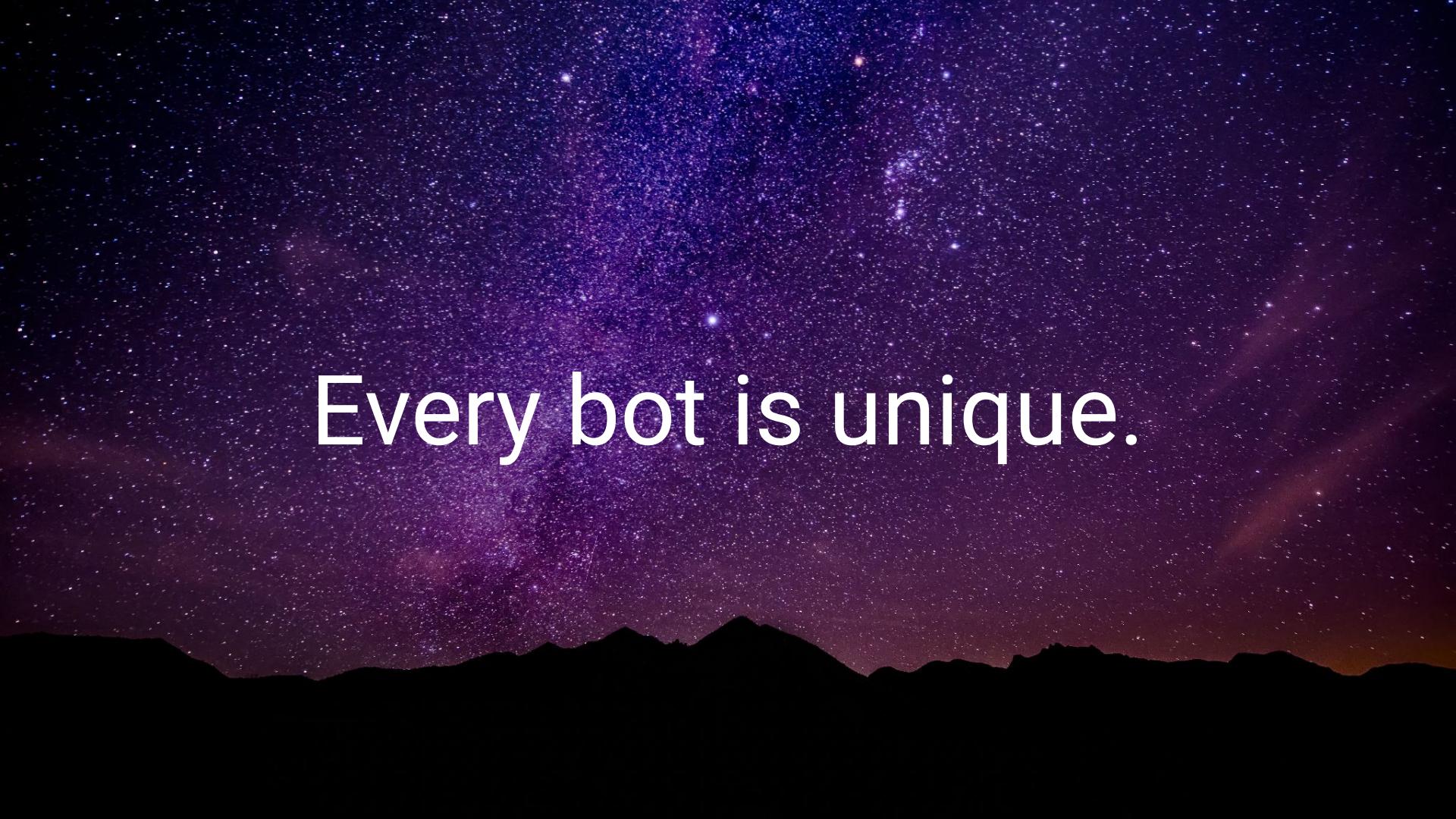
- Build a pleasing and natural conversation
- Provide a short and concise experience
- Keep user engaged without falling in lengthy interactions
- Handle angry and frustrated users
- ...



# Performance Metrics

- Customer Satisfaction (Csat - thumbs up/down)
- Escalation (transfers to live agents in a call center)
  - User requested agent from the start = **HARD to overcome**
  - User requested agent after interacting = **WORK ZONE**
  - Intentional by design = **OK**
- Deflection (bot not able to “understand” the user)
- Goal completion rate (user reach the end of an engaged interaction)
- Average conversation length
- Post-hoc user surveys
- Time between 2 interactions
- ...





Every bot is unique.