

# Conversational interaction and web accessibility

# What is conversational interaction?

We call **conversational interaction**, an interaction paradigm where users interact with a machine through **natural language** instructions.

There are two main channels: **text** and **voice**

The underlying **technologies** that enable machines to understand and respond to human language is what we call **conversational AI**

# Characteristics of conversational interaction

## Personality

Agents are able to simulate human attitudes

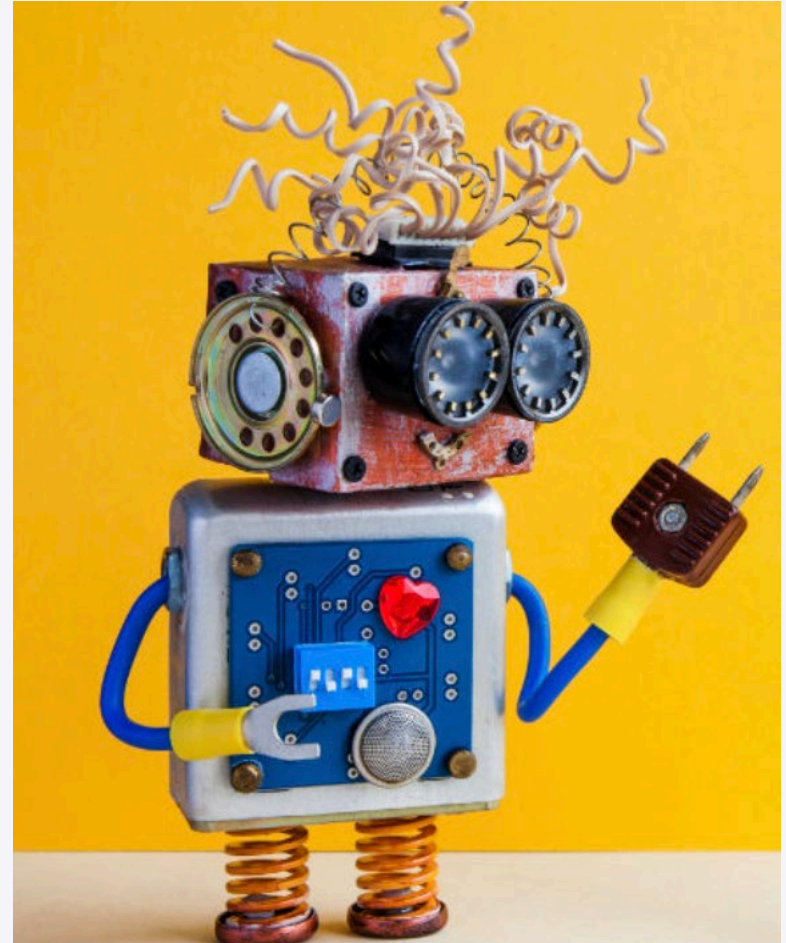
## Dialogue

Agents are able to simulate natural language

## Empathy

Agents are to appropriately express human emotions

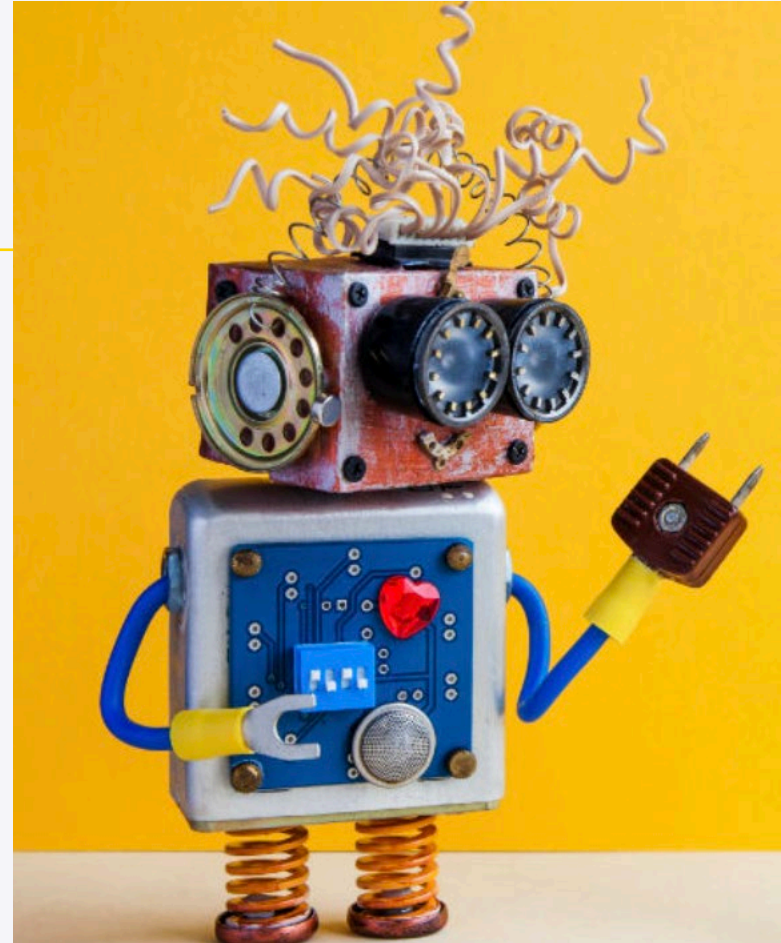
# Anatomy of conversational AI



# Anatomy of conversational AI

Comprehending what the user says

---



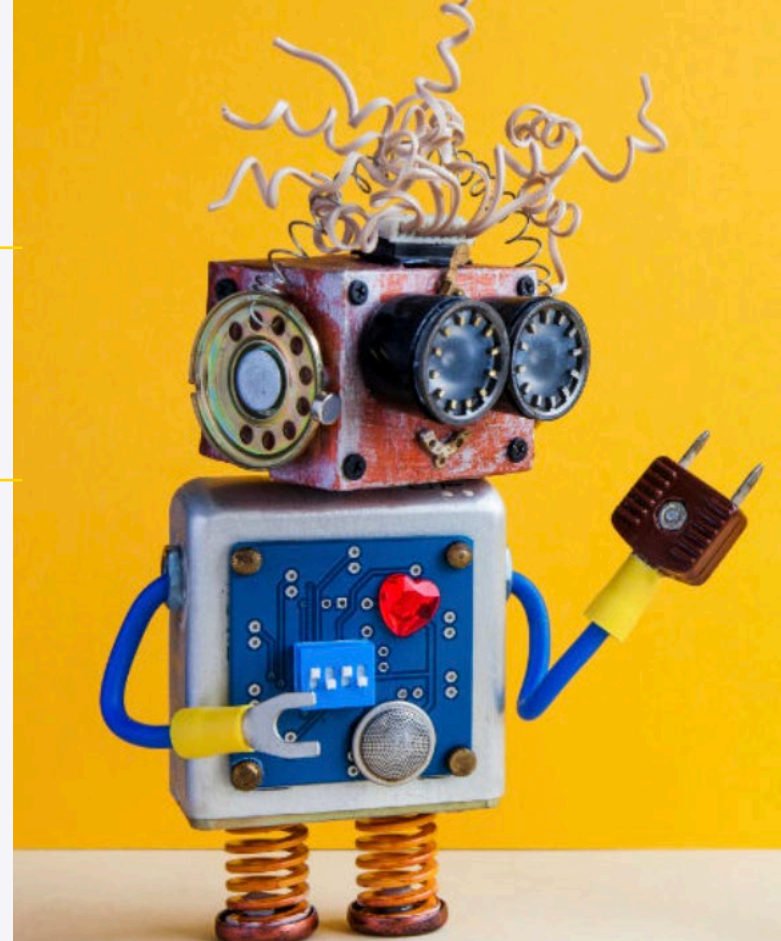
# Anatomy of conversational AI

Comprehending what the user says

---

Answering user's questions

---



# Anatomy of conversational AI

Comprehending what the user says

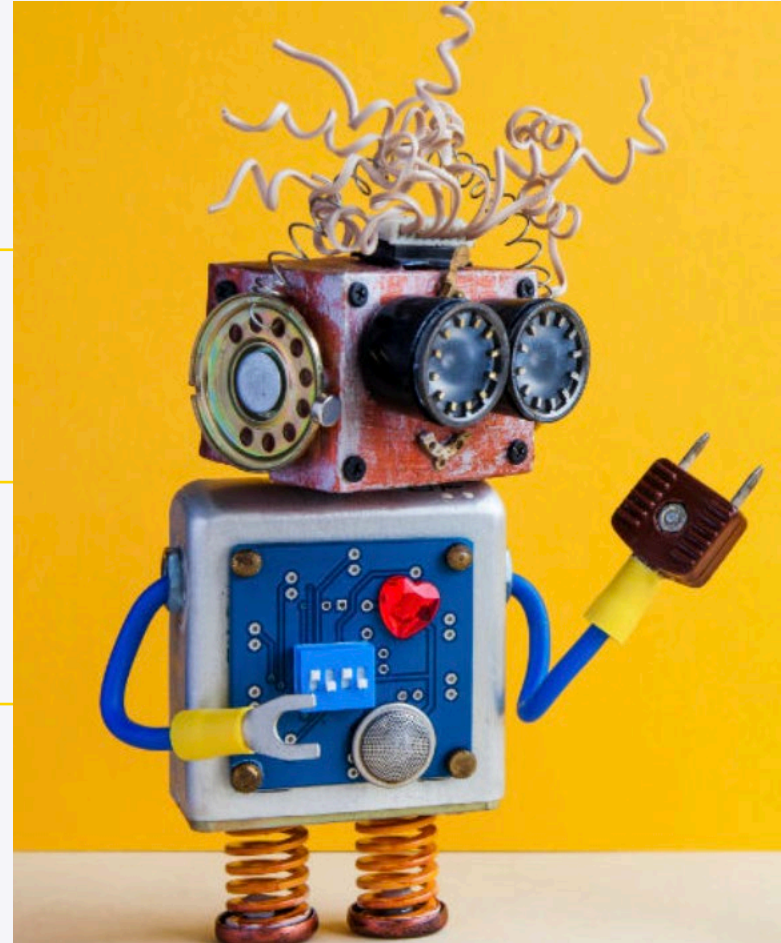
---

Answering user's questions

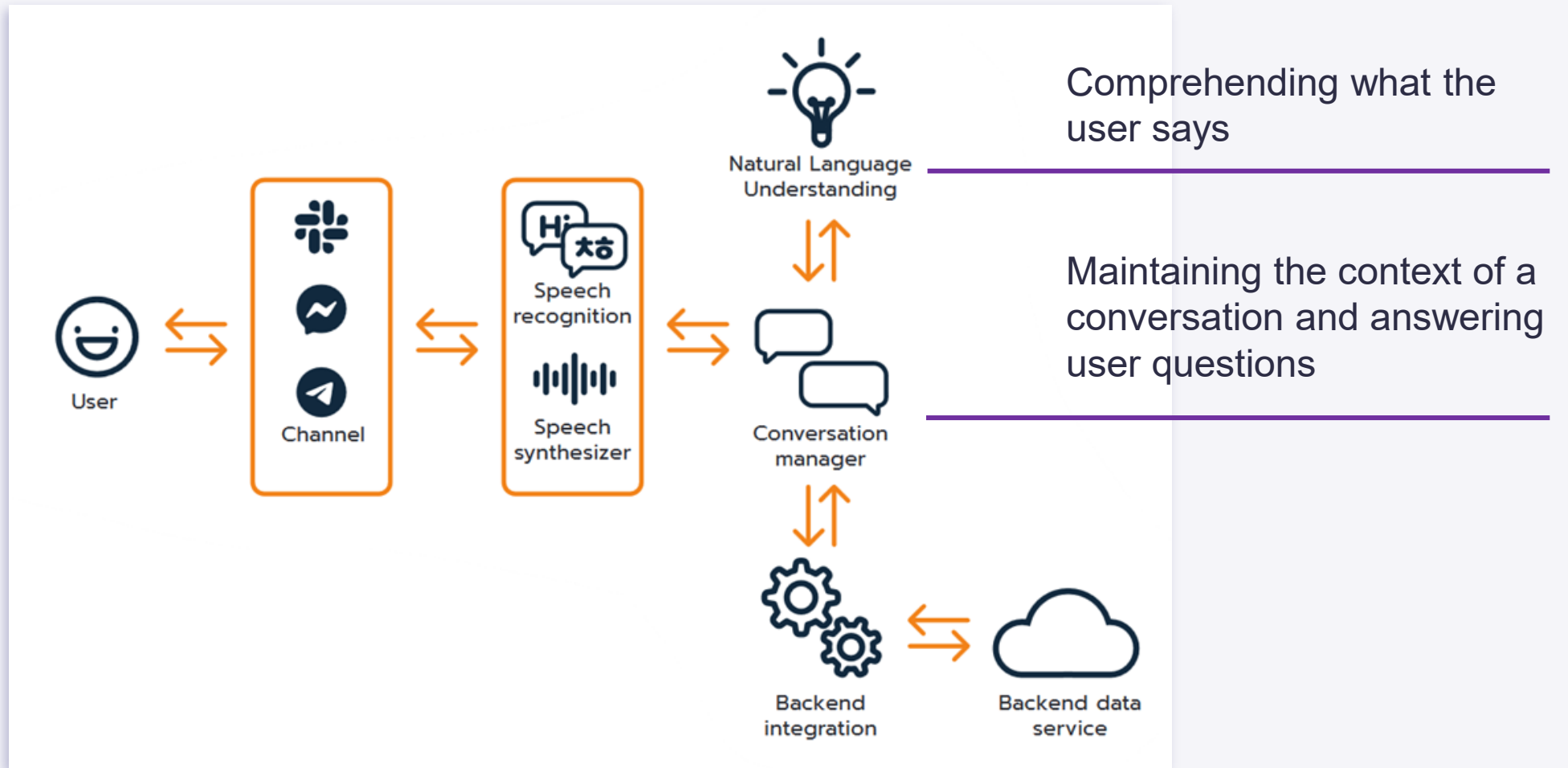
---

Maintaining the context of a conversation

---



# How?



# Natural Language Understanding 101

**Intents:** the goal of the user

**Entities:** relevant parts of a sentence that characterize the user intent

User says:



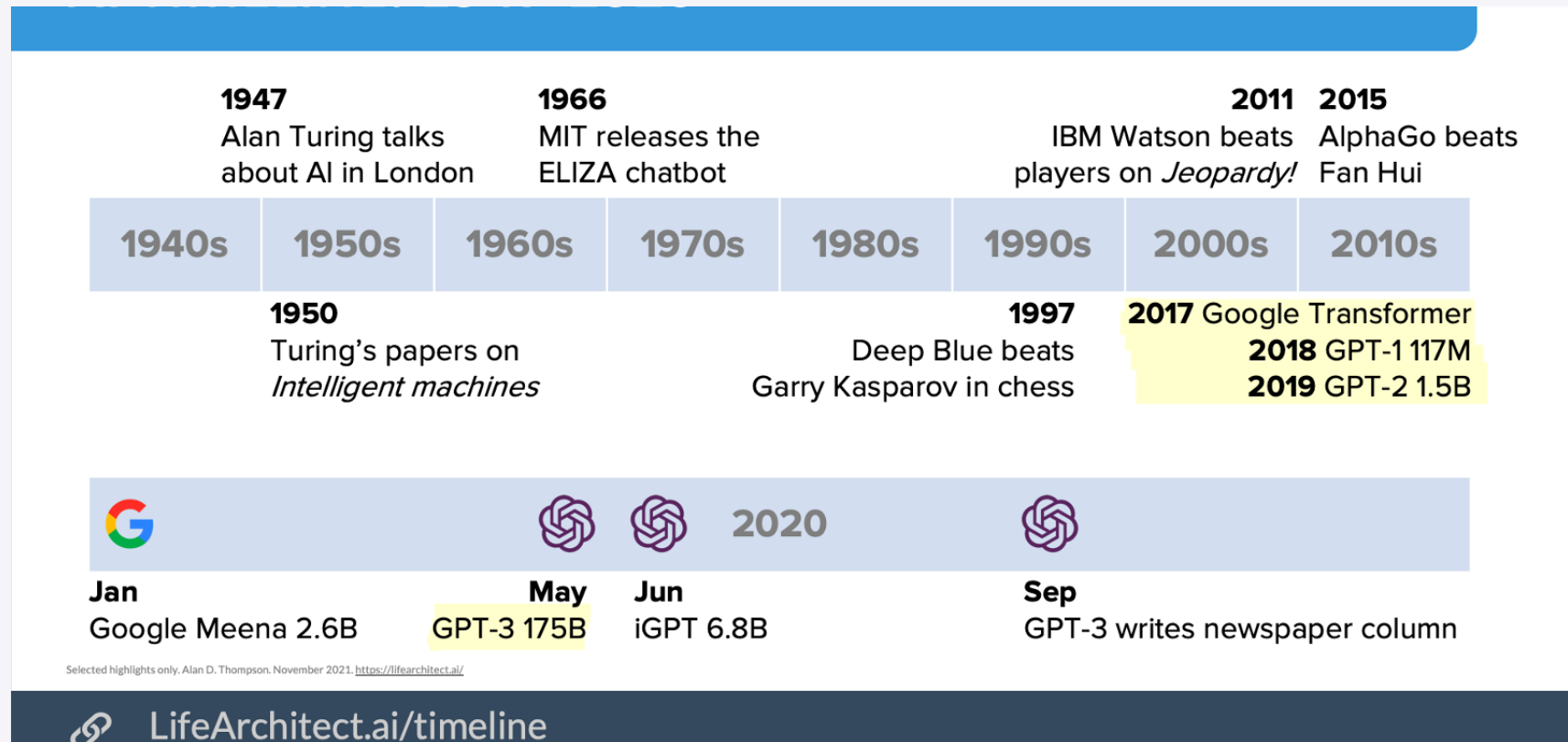
Can I buy an orange shirt?

Chatbot understands

Check if there is an entity clothing item= shirt with the entity colour= orange



# Not really a new technology



# From Eliza...

Eliza was a rule based chatbot. ELIZA simulated conversation by using a pattern matching and substitution methodology

Men are all alike.

They're always bugging us about something or other.

Well, my **boyfriend** made me come here.

He says I'm **depressed** much of the time.

[...]

My **father**

**Bullies.**

*In what way*

*Can you think of a specific example*

*Your **boyfriend** made you come here*

*I am sorry to hear you are **depressed***

*Who else in your family takes care of you?*

*What else comes to mind when you think of your **father***

*Does that have anything to do with the fact that your **boyfriend** made you come here*

# ...To ChatGPT

OpenAI's chatbot based on their LLM model GPT-4



Hi ChatGPT, can you present yourself to a class of web engineering students?



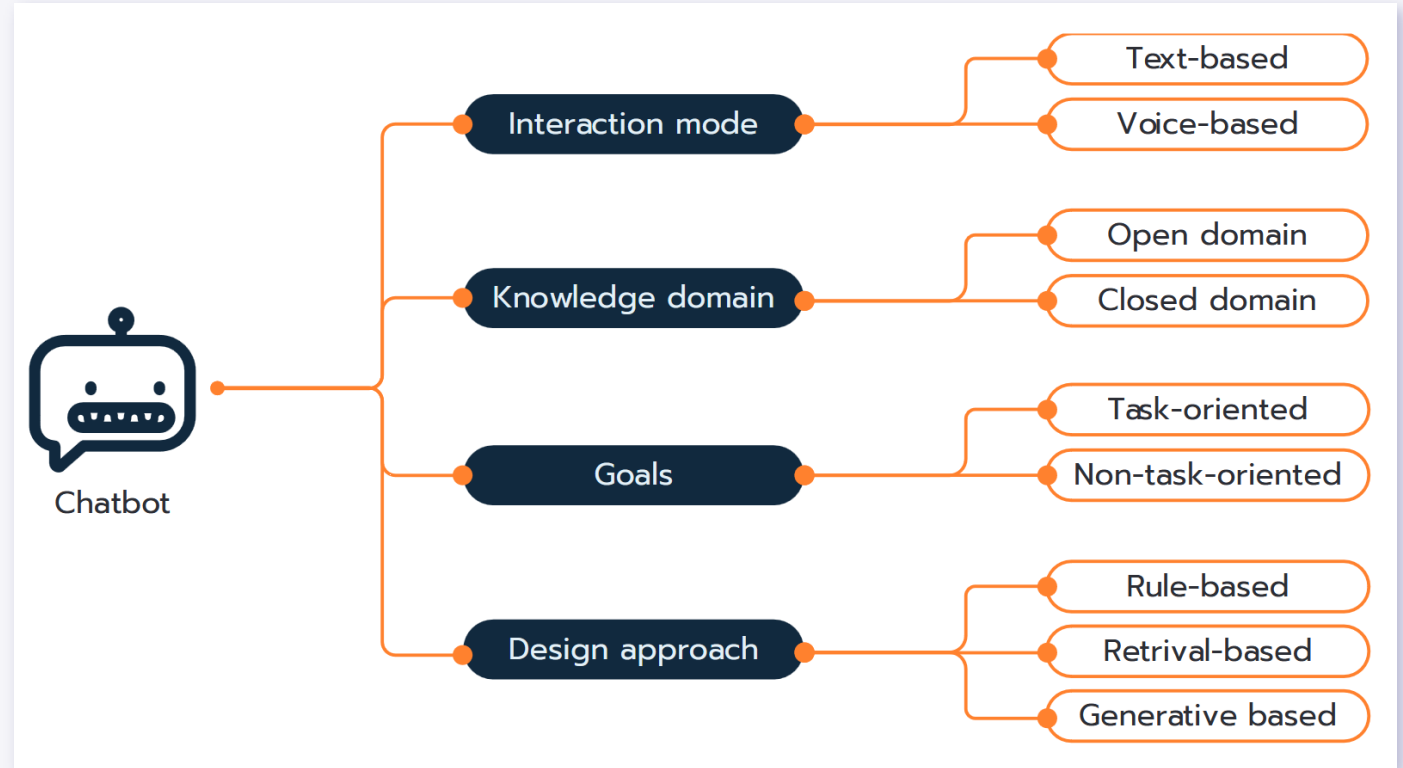
Hello everyone,

I'm ChatGPT, an AI developed by OpenAI. Think of me as your versatile assistant, collaborator, and coding buddy for all things web engineering! I've been trained on a massive amount of text, which means I can help with tasks like:

- **Explaining Concepts:** Whether it's REST APIs, web frameworks, or the nuances of HTTP, I can break down complex ideas into understandable chunks.
- **Coding Assistance:** Need help debugging, optimizing, or generating code? I can write snippets in various programming languages like JavaScript, Python, HTML, CSS, and more.
- **Best Practices:** I can guide you on industry standards for accessibility, responsive design, or web performance optimization.

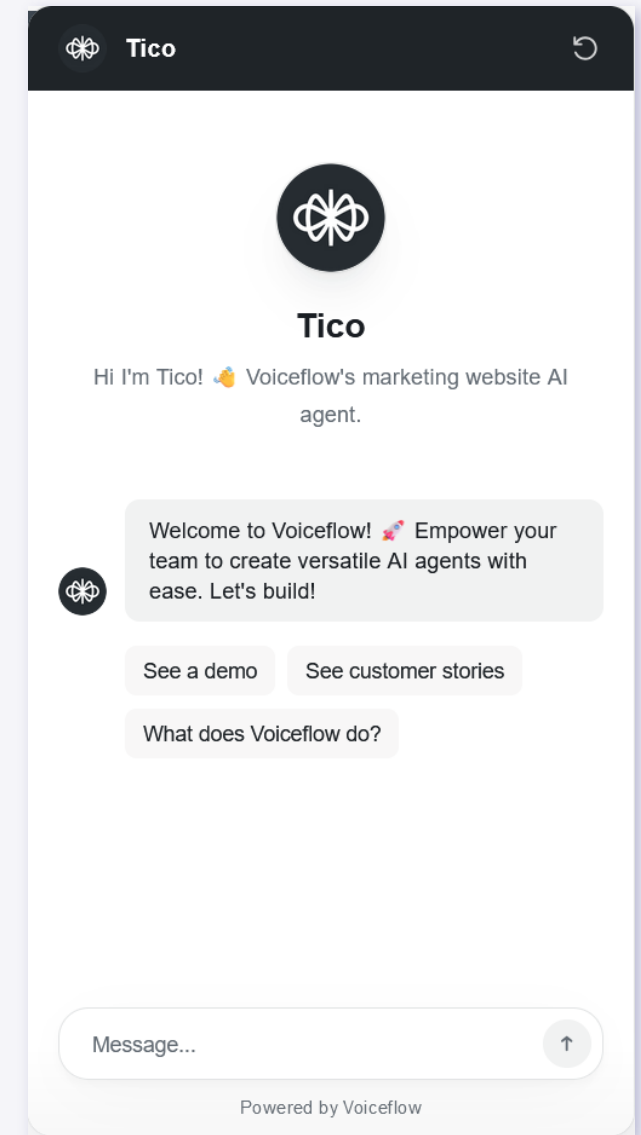
# 1 - Types of conversational user interfaces

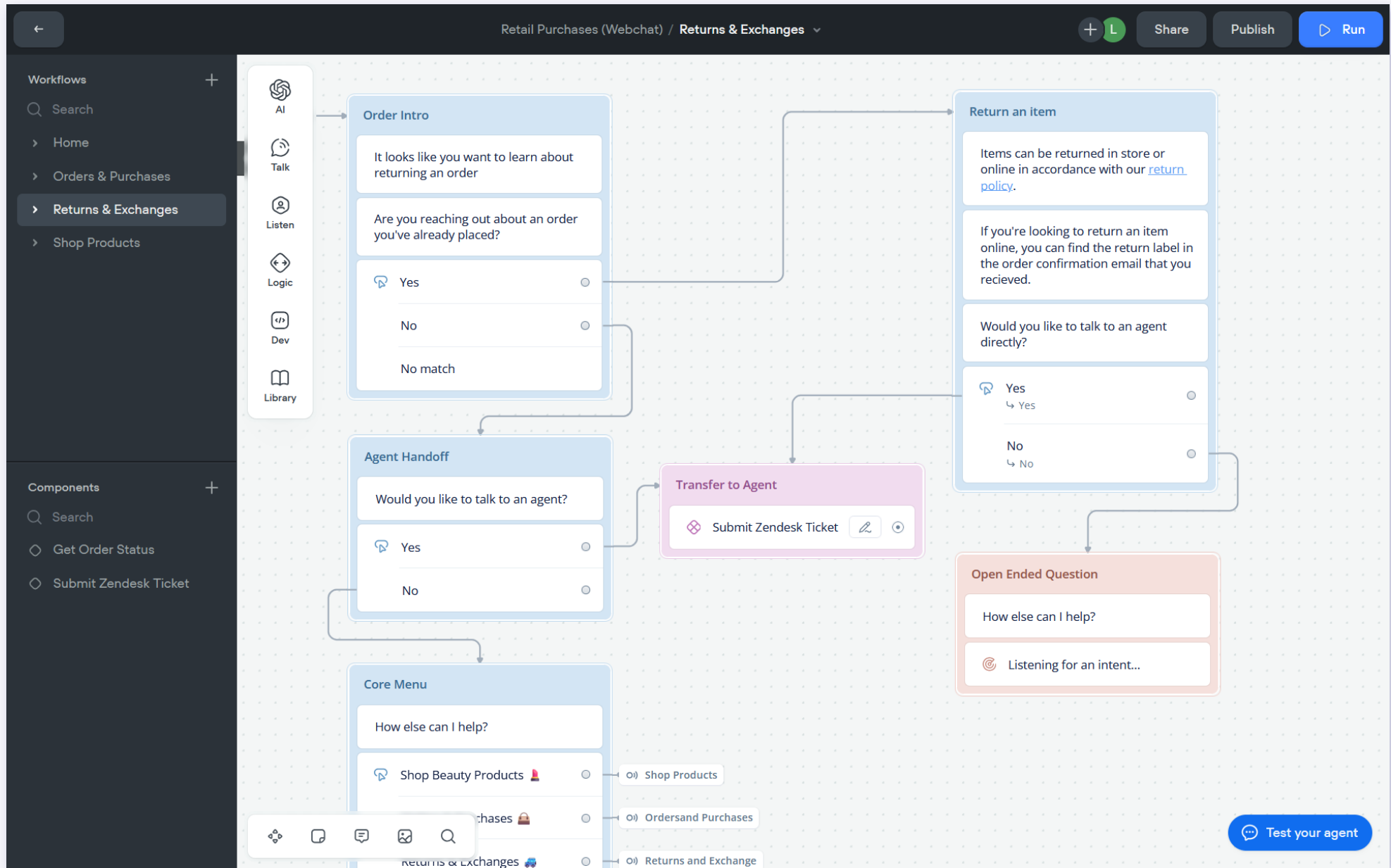
- **Chatbot:** all those applications that adopt *natural language interaction* as their interaction paradigm. Usually, the term refers to applications that use **text** as their primary channel



# Rule-based chatbots

Also referred to as decision-tree bots, they are built on manually constructed rules. The rules can be based on an “if-then” logic or involve pattern matching. The bot contains a knowledge base with documents





# Generative AI-based chatbots

Currently the most popular and advanced architecture for chatbot relies on Large Language Models (LLMs) fine-tuned on conversational interaction tasks

The Gemini logo features the word "Gemini" in a sans-serif font. The "G" is blue, and the remaining letters "emini" are purple. A four-pointed starburst is positioned above the "i".

Gemini

The Claude logo consists of an orange sunburst icon to the left of the word "Claude" in a black serif font.

Claude



# Retrieval augmented generation chatbots

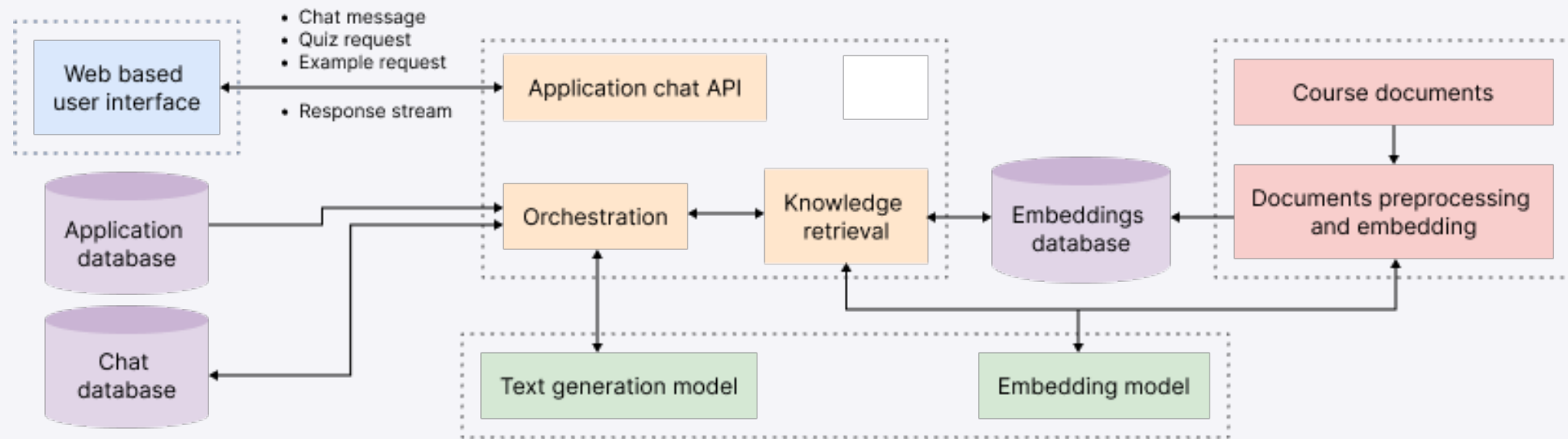
RAG chatbots combine **retrieval-based techniques** (fetching information from external sources) with generative AI to craft natural responses

The screenshot shows the MyLearningTalk chatbot interface. On the left, there is a 'Library' section with a list of chat history items, including 'New chat 28/01/2025 11:35...' and 'New chat 20/12/2024 16:59...'. The main chat area on the right displays the user's query 'What is smart learning design?' and the chatbot's response. The response includes a definition of Smart Learning Design, its benefits, and an example of its application in a design course. Below the response, there are buttons for 'Resource', 'Example', and 'Quiz'. At the bottom, there is a text input field with the placeholder 'Ask MyLearningTalk' and a send button.

[MyLearningTalk: An LLM-Based Intelligent Tutoring System](#)

# MyLearningTalk

MLT exploits retrieval augmented generation to offer interactive features to provide users with grounded answers and a tailored experience to enhance and facilitate the learning process



# 2 - Types of conversational user interfaces

- **Embodied agent:** interfaces that use conversational interaction using either text or voice that are also **personified** or have an avatar, for example:
  - **Virtual avatars**
  - **Social robots**



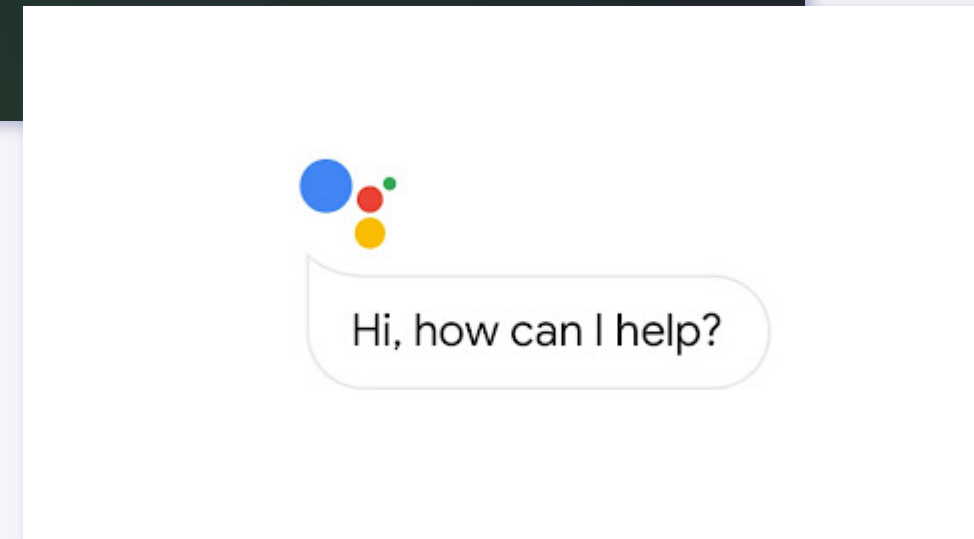
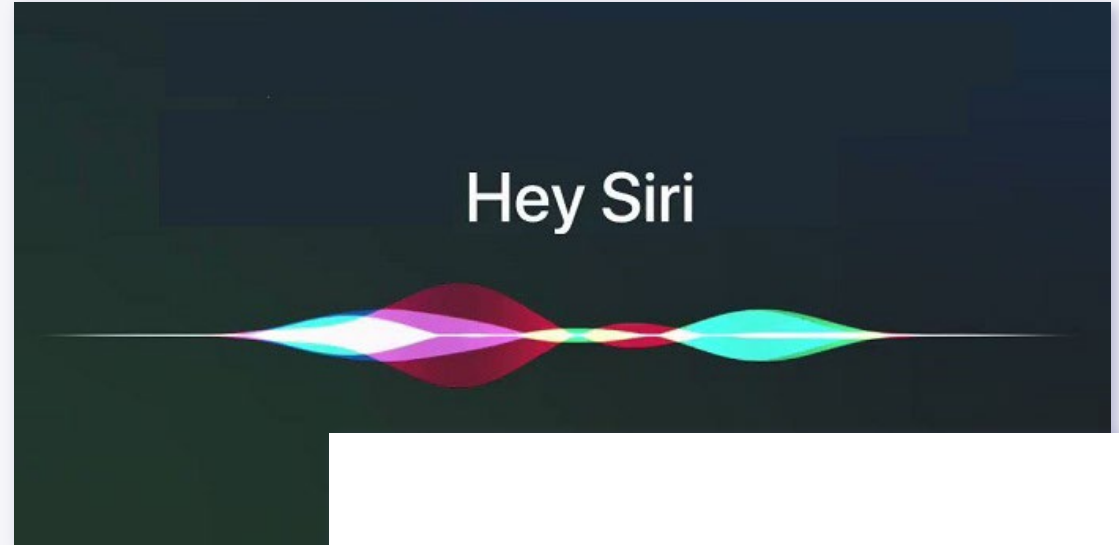
# Kuki.ai



[https://youtu.be/5EflAOI\\_itY?si=WtJ9KEHNqFILu2E0](https://youtu.be/5EflAOI_itY?si=WtJ9KEHNqFILu2E0)

# 3 - Types of conversational user interfaces

- **Voice User Interface:** conversational interface that use as its primary interaction channel **voice**
  - **Personal Virtual Assistant:** all those voice interfaces that also have an **element of sociality**. (Alexa, Siri...)

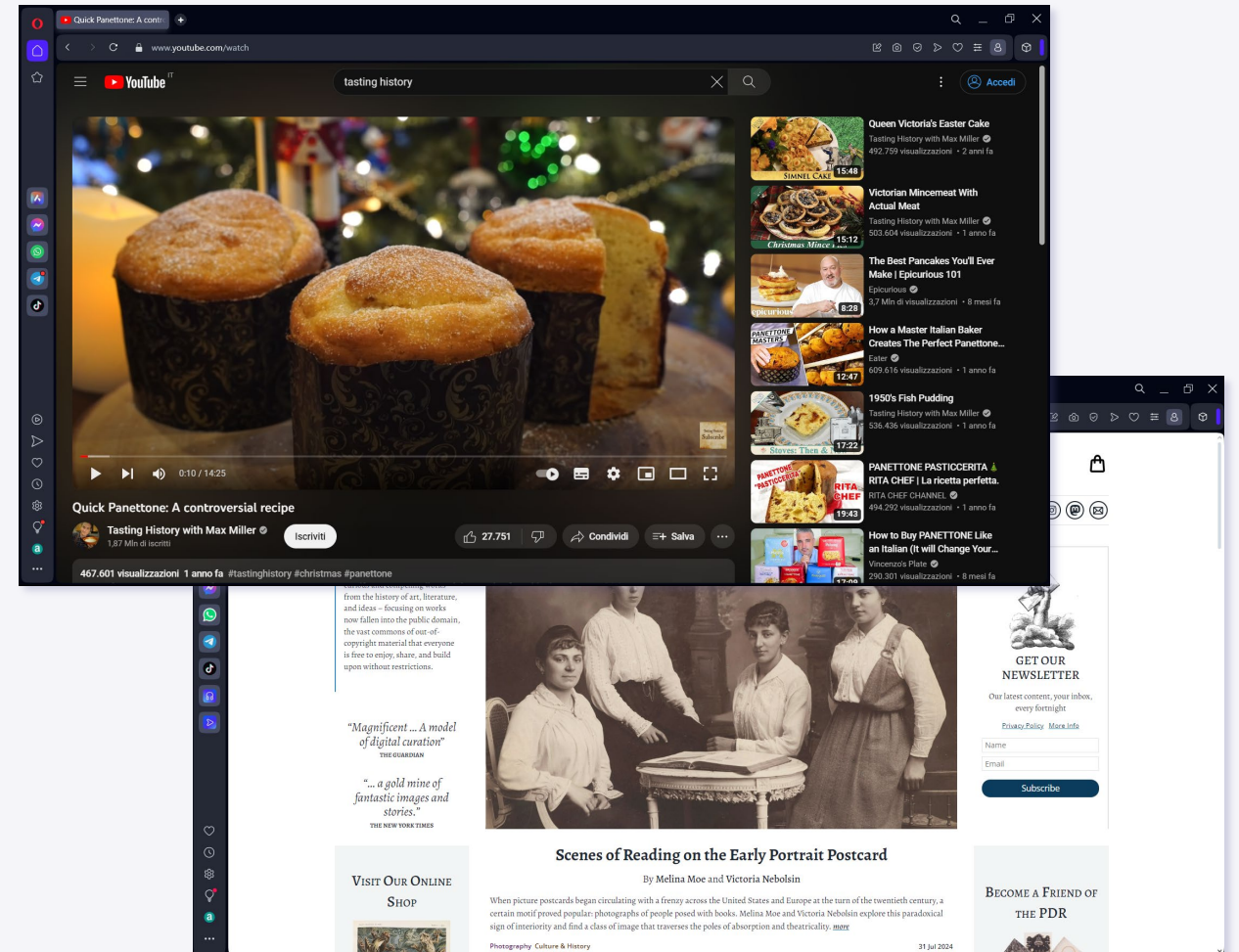


# Opportunities for accessibility

# Web browsing and accessibility

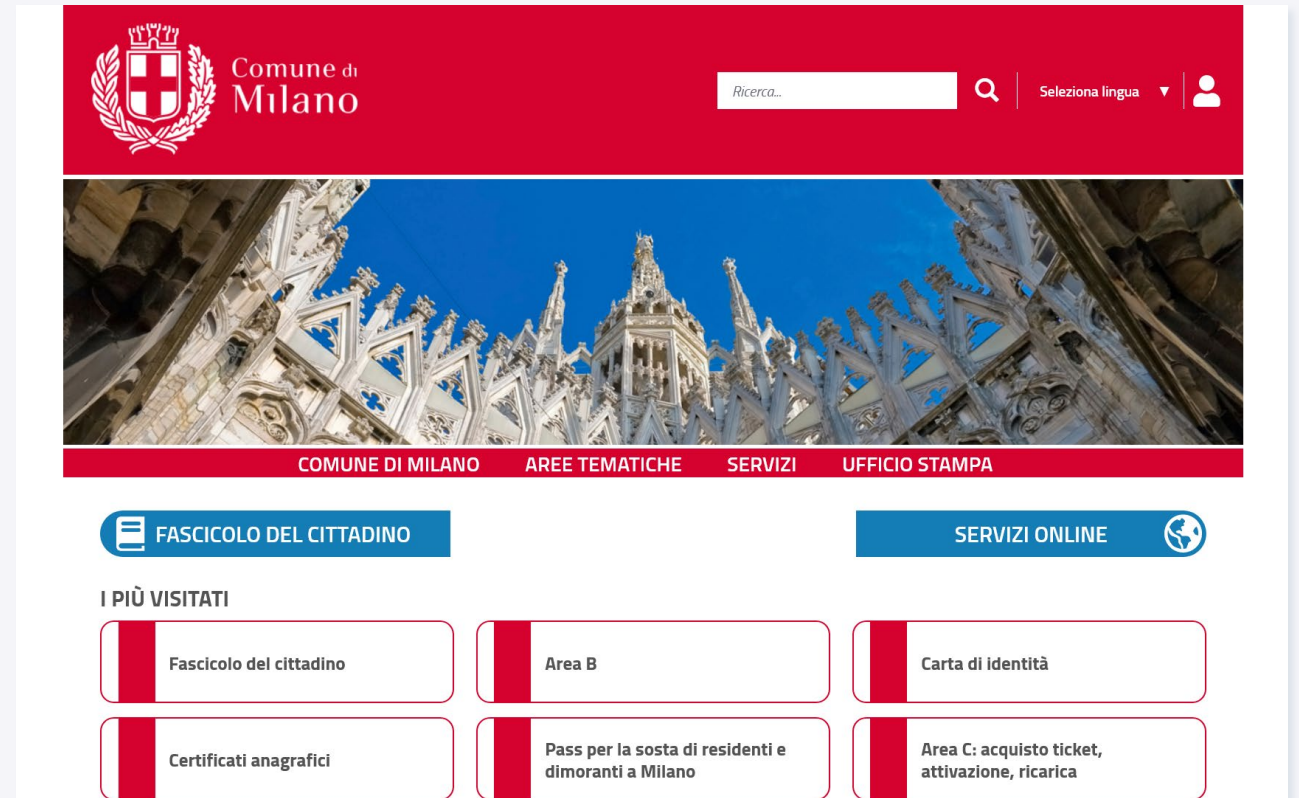
The standard interaction paradigm for web browsing presents two main barriers for people living with certain disabilities:

- Mouse and keyboard interaction
- Visual content



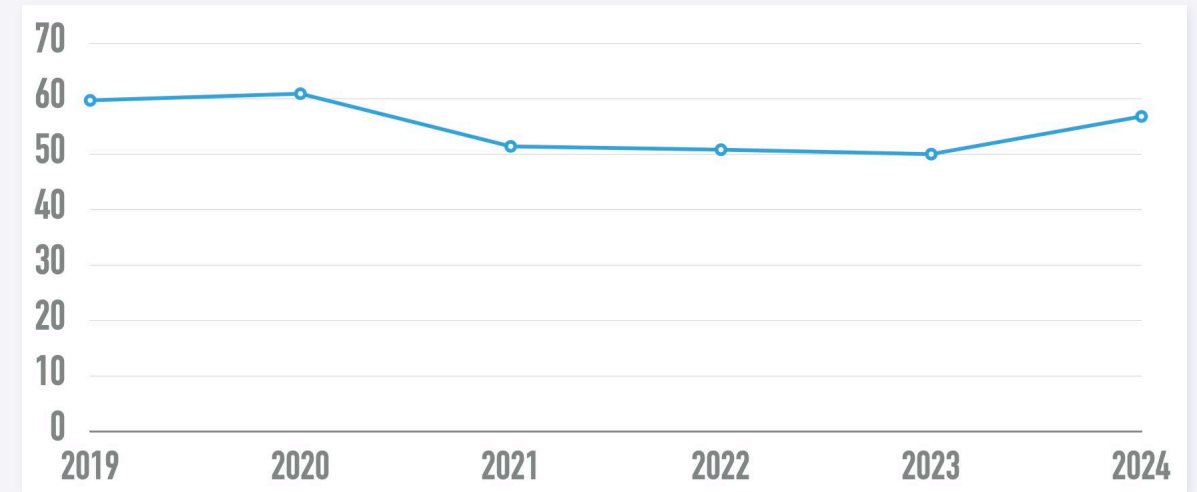
# Web browsing and digital inclusion

- By EU directives<sup>1</sup>, by 2025 all digital platforms should be accessible.
- This is even more important for the Public Administration since it means creating opportunities for inclusion.



# Assistive technology is often not enough

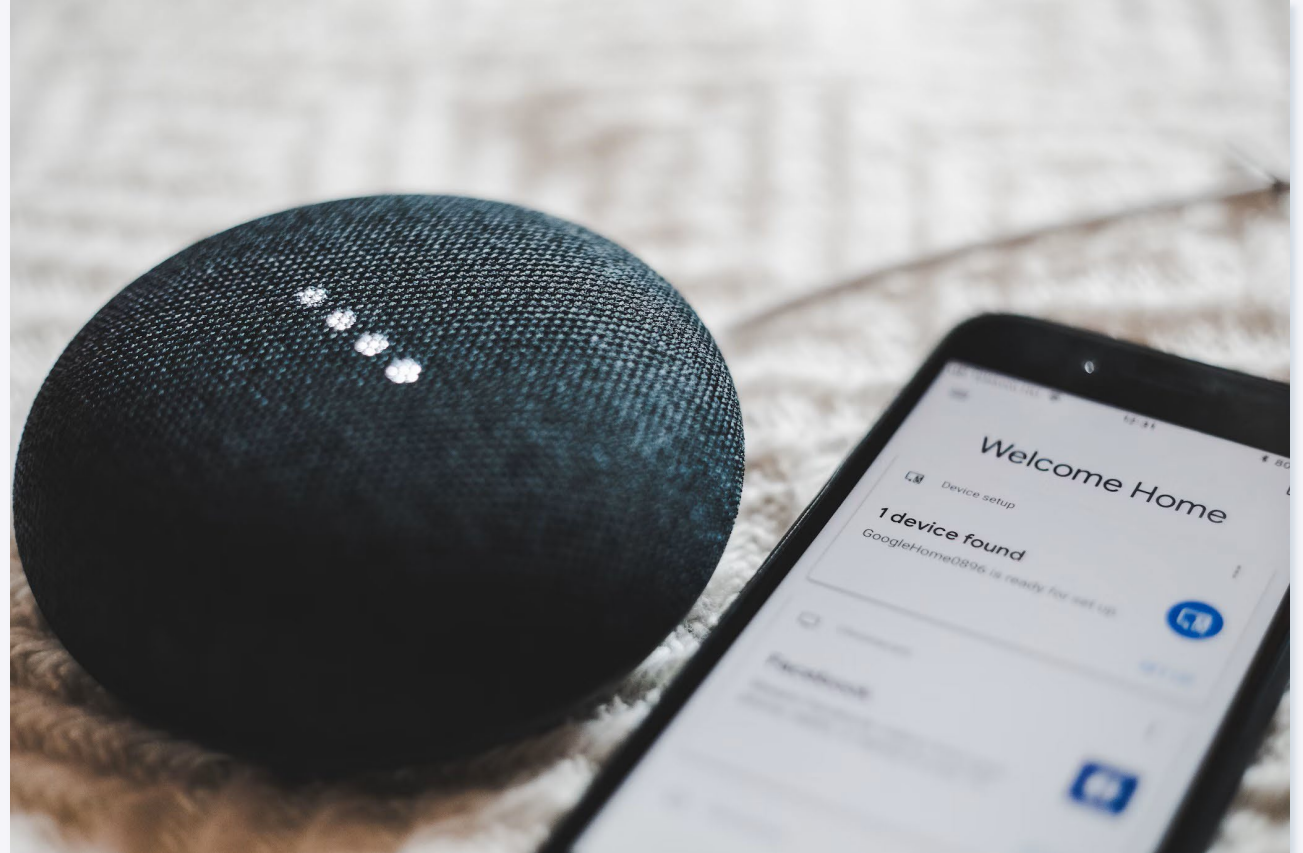
- Screen readers are the most common assistive technology in use. However, their efficacy is **dependant on code quality**.
- The WebAIM 2024 report shows that **95.9% of home pages** had detected WCAG 2 failures.



Average of detected errors per page. In 2024 the average is 56.8 errors per page. The report sampled the most visited 1 million websites.

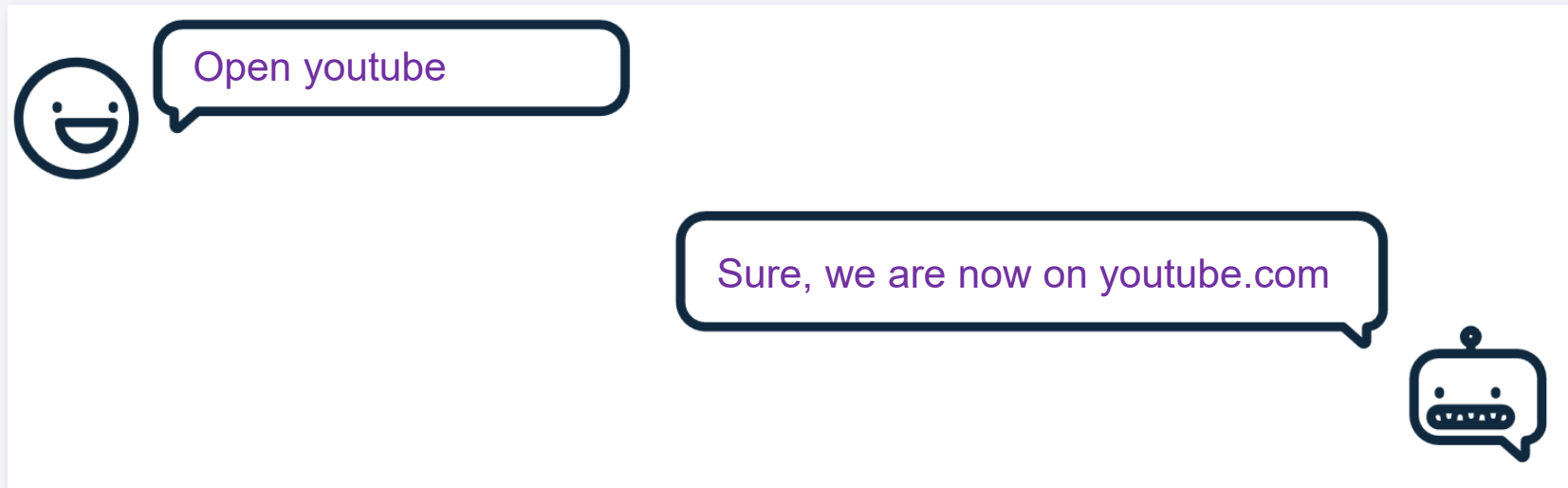
# Conversational AI as an opportunity

- Voice assistants like Amazon Alexa or Google Assistant show how Natural Language interaction has the potential to be a more accessible interaction paradigm.
- **At the moment these assistants cannot browse the Web**



# Conversational web browsing

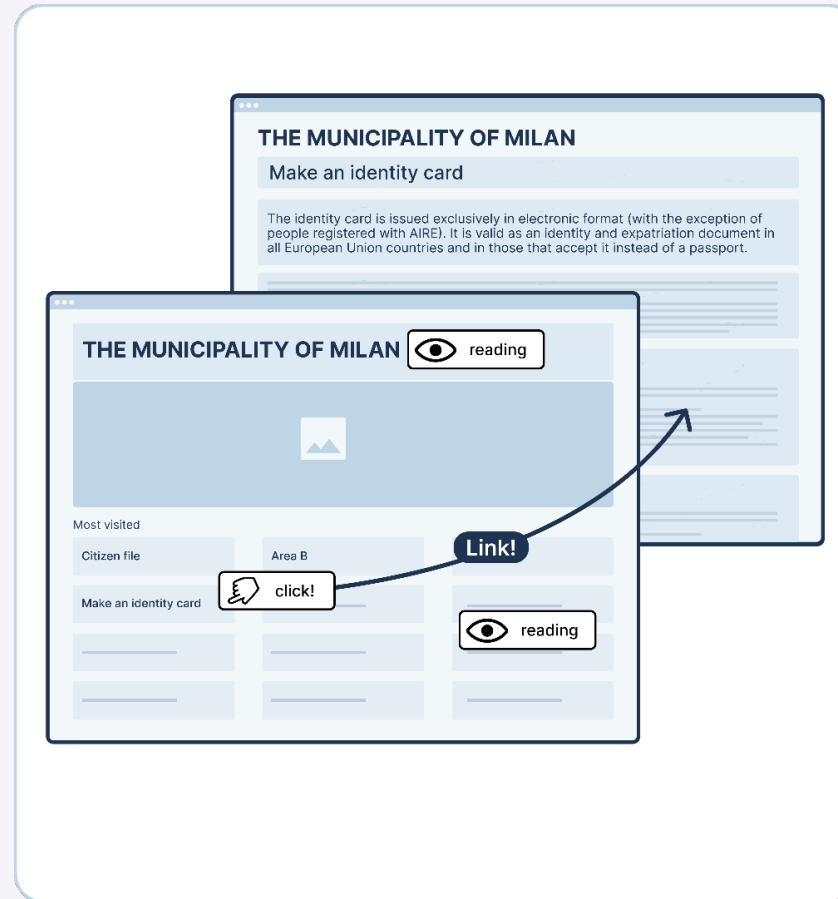
The act of performing common **browsing tasks**, such as skimming through content, reading it, or navigating pages **using only natural language commands**



# ConWeb

Browser extension that enables users to browse the web through voice commands, such as “read” and “navigate”

It is able to describe pages' structure and retrieve content based on users' requests



(a) Automated Web browsing actions

CONWEB MIDDLEWARE

- CW** You are now on the Municipality of Milan home page. You can browse the [main menu](#), navigate to the [Most visited services](#) or to the [news](#).
- Is there anything about how to make an identity card? **U**
- CW** Yes, you can find Make an identity card in the [services](#), do you want to navigate to the page?
- I want to navigate to the page **U**
- CW** No problem! Here it is: [The identity card is issued exclusively in electronic format \[...\]](#). Do you want to know more or to read something else?
- Yes read more **U**
- CW** It is valid as an identity and expatriation document in all European Union countries and in those that accept it instead of a passport.
- Save this page with a bookmark **U**
- CW** You saved the page [Make an identity card](#) from the [Municipality of Milan](#) website as a bookmark. Do you want to access all your bookmarks, or do you want to continue reading the page?

(b) Natural language dialogue

# Research with blind and visually impaired users

ConWeb was developed through a User centred design process.

Blind and visually impaired people were involved in co-design sessions to collect their challenges with browsing the web





Comune di  
Milano

Ricerca...



COMUNE DI MILANO

AREE TEMATICHE

SERVIZI

UFFICIO STAMPA

FASCICOLO DEL CITTADINO

SERVIZI ONLINE



### I PIÙ VISITATI

Fascicolo del cittadino

Dichiarazione Tari: occupazione di  
appartamenti e immobili

Carta di identità

Certificati anagrafici

Area B | Registra il tuo veicolo

Area C | Acquisto ticket,  
attivazione, ricarica

Multe: guarda le foto delle

Multe e sanzioni: pagare un

Prenota il tuo appuntamento in

elixForms - Richiesta pass per ...

formshd2.comune.milano.it/rwe2/forms/form.jsp?IUQOID=164706468&JURLGY=schemadata&JUXSID=TX\_I4227\_COMUNEMILANOHD\_1732181655687\_R-448048141&ELAN...

Research Video altro PhD allenamenti Google Docs Google Sheets Karate Assistants Playgrou... All Bookmarks

Comune di Milano

hai bisogno di aiuto?

Il modulo è in modalità COMPILAZIONE DI TEST

Richiesta pass per la sosta e la circolazione di persone con disabilità (CUDE - art. 381 DPR 495/1992) - Simulazione IA

1. Informativa Privacy 2. Dati richiedente 3. Dati intestatario del pass disabili 4. Richiesta Pass 5. Dichiarazioni integrative 6. Modalità di ritiro del pass 7. Targa 8. Riepilogo 9. Convalida 10. Inoltra

2. Dati richiedente

**on this module you can request a special parking permit**

Il diretto interessato;

Form filling

# Designing for conversational web browsing

Based on the interaction channel (voice or text) and the use case (automotive, healthcare...), there exist **specific interaction guidelines** to design conversational interfaces

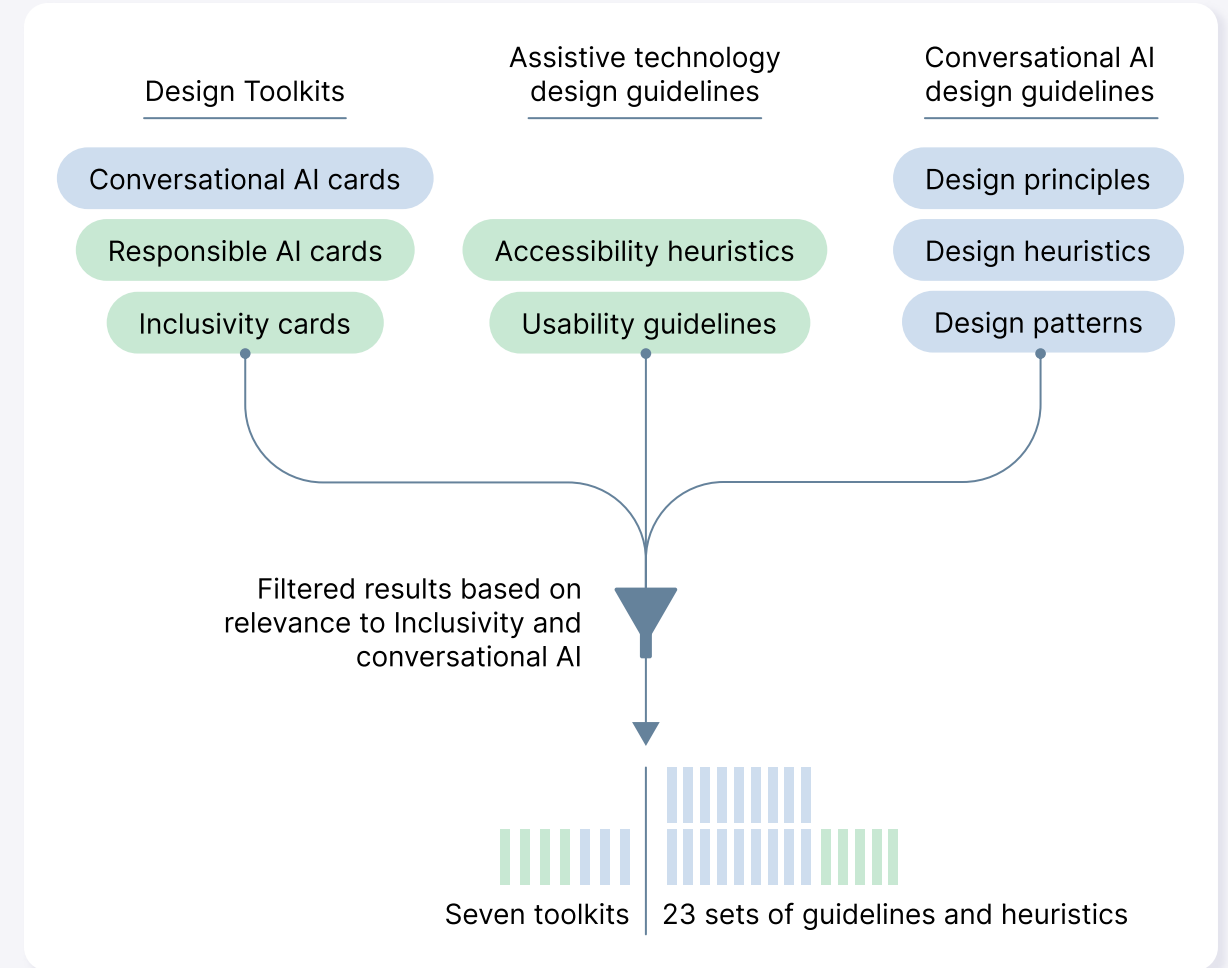
## Research questions

- 1) What are the effects of designing for voice on traditional web design?
- 2) Which of the current guidelines support voice web browsing?

# Surveying the literature

## Topics analysed:

- Accessible and Inclusive design
- Conversational AI



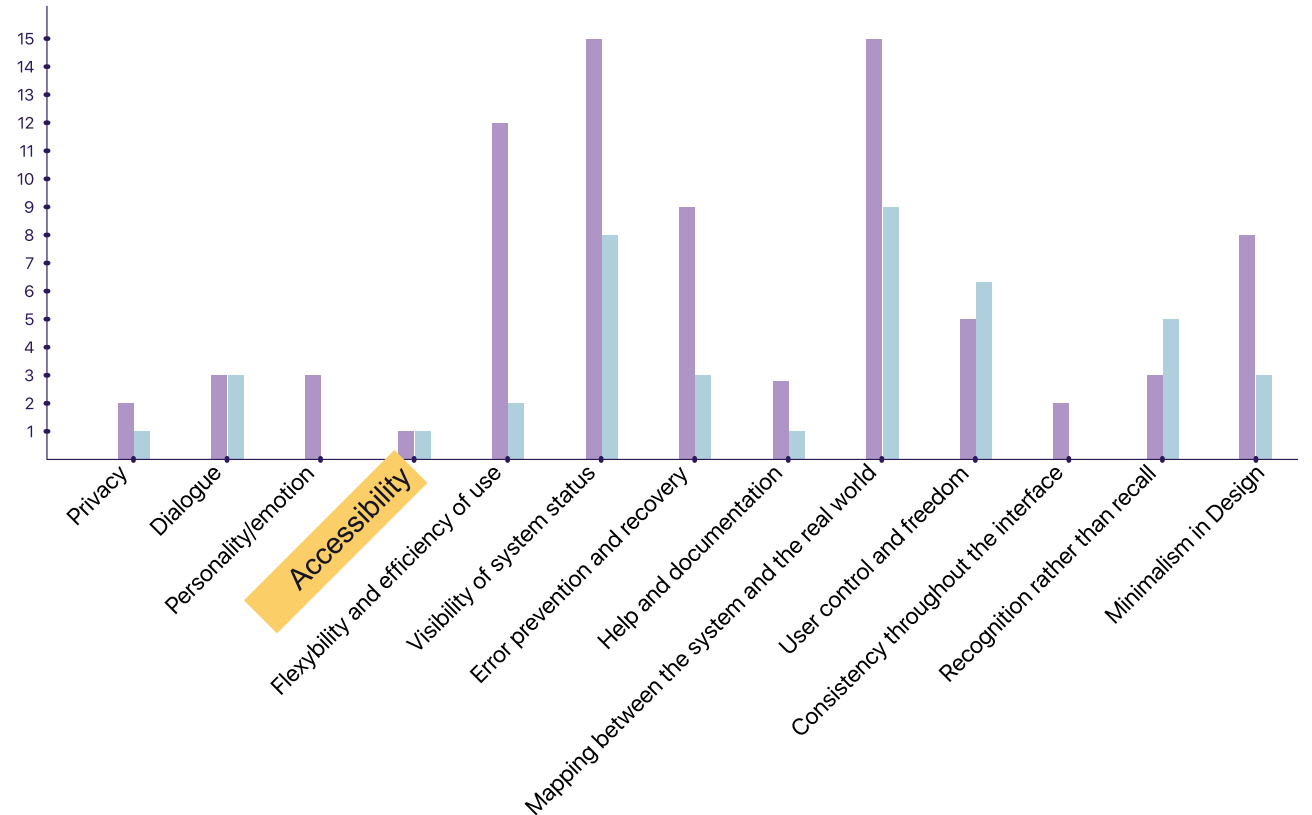
# Findings

Current guidelines do not guide practitioners and researchers in developing and designing voice agents that support interaction with web components.

Number of guidelines per topic in conversational interaction literature

**Legend**

- Design guidelines
- Evaluation heuristics



# Evaluating guidelines

- Conversational interaction designers and developers have been involved in workshop sessions to validate the synthesized guidelines

- 7 sessions: 2-3 participants per session
- 20 participants: 7 CUI researchers, 4 HCI researchers, 5 web developers, 4 interaction designers, 1 accessibility expert

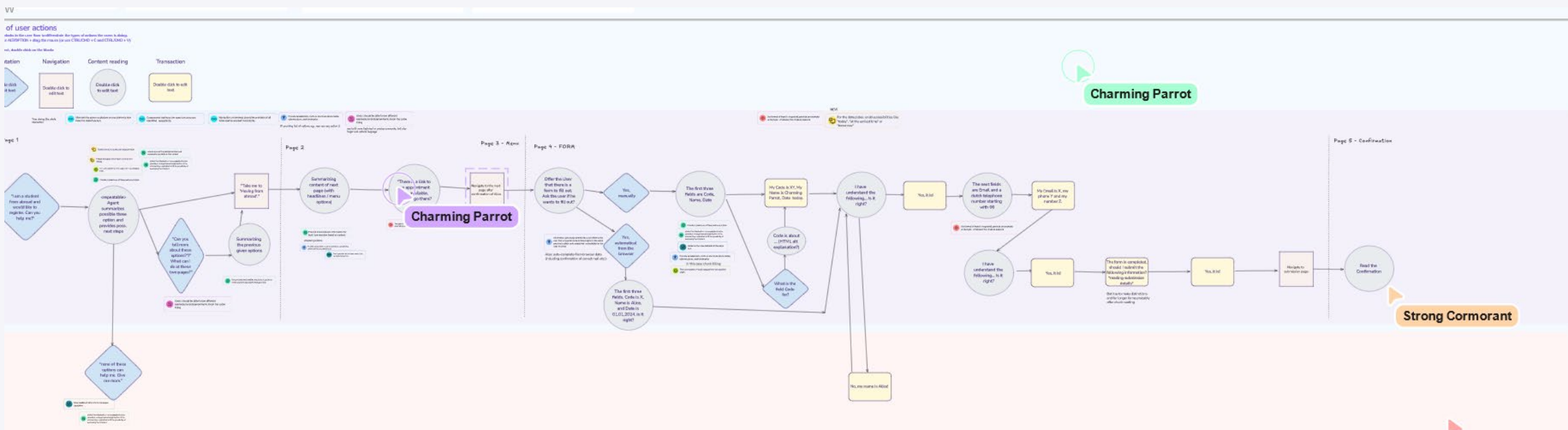
The image displays ten guidelines for conversational user interfaces, organized into a grid. Each guideline is represented by a colored card with a title, an icon, and a list of specific recommendations.

- User control and freedom** (Pink card):
  - Users should be able to override, change or delay preferences and inputs at any point of the conversation.
  - Users should be able to use different expressions and synonyms to mean the same thing.
  - Users should be able to deference where they are in the conversation. The design should provide users information about where they are in the website.
  - A conversation should support direct questions to ask whether the website is supplying a specific feature.
- Error prevention** (Red card):
  - If a format of input is suggested, provide an example of the type of answer the chatbot expects.
  - Feedback events are present to handle unknown requests.
  - Error messages are in plain language and suggest how to overcome the error.
- Flexible, efficient and inclusive use** (Blue card):
  - Provide accelerators, such as shortcuts, bookmarks, abbreviations, and tab/dropdowns.
  - Information previously entered by or provided to the user needs to remain in the system again in the same process to either auto-populated, or available for the user to select.
  - Providing direct and informative questions and sentences.
  - Assessing a short round to different elements used of these kinds of a multi-media environment can make the user more "productive".
  - The system is able to start tasks and be used by users independently from their health conditions, work being, age, etc.
- Consistency throughout the interaction** (Cyan card):
  - Maintain the same secondary access elements that have the same function.
  - Components that have the same functions are identified consistently.
  - Navigation commands should be available at all times and be needed consistently.
- Help and documentation:** (Light green card):
  - The conversation should suggest the next possible steps.
- Recognition rather than recall** (Teal card):
  - Present content hierarchically.
  - Provide a view in the large of how the content is organized.
  - Only present information relevant to the task.
  - Ensure link predictability providing a preview of the content reachable through a link.
  - Divide the information in manageable chunks: provide a concept based organization of the information, associated with the possibility of accessing that content.
  - Provide summaries of long textual information to be accessed before delivering it in full.
  - Inform users of the possible actions and commands available in that context.
- Mapping between system and language conventions:** (Yellow card):
  - Follow dialogue conventions such as turn taking.
  - G1.2 Express emotions: Apologising, encouraging, or enjoying, etc.
  - Name functions as the user expects them.
- System status and conversational context** (Light blue card):
  - The interaction should keep track of the dialogical context.
  - Focus on the new elements in the page first.
  - Inform users when they need to take action and provide instructions.
  - Ask confirmation for user input.
  - Give feedback when there are page updates.

# Let's try together

You have design the conversational interaction with a short form available at the link provided in chat today.

Following the voice interaction guidelines provided, define a user flow describing how a user should interact with your form



# Tasks

1. Analyse the form to understand the user goal
2. Identify the optimal path of interaction
3. Identify which are the elements within the pages that you want to deliver through voice and how.
4. Define how the agent should present the content
  - Wording
  - Help messages
  - ...
5. Think if any changes in the current information architecture would be needed

# Thank you!

## Contact information

Ludovica Piro, PhD candidate, Politecnico di Milano

Ludovica.piro@polimi.it