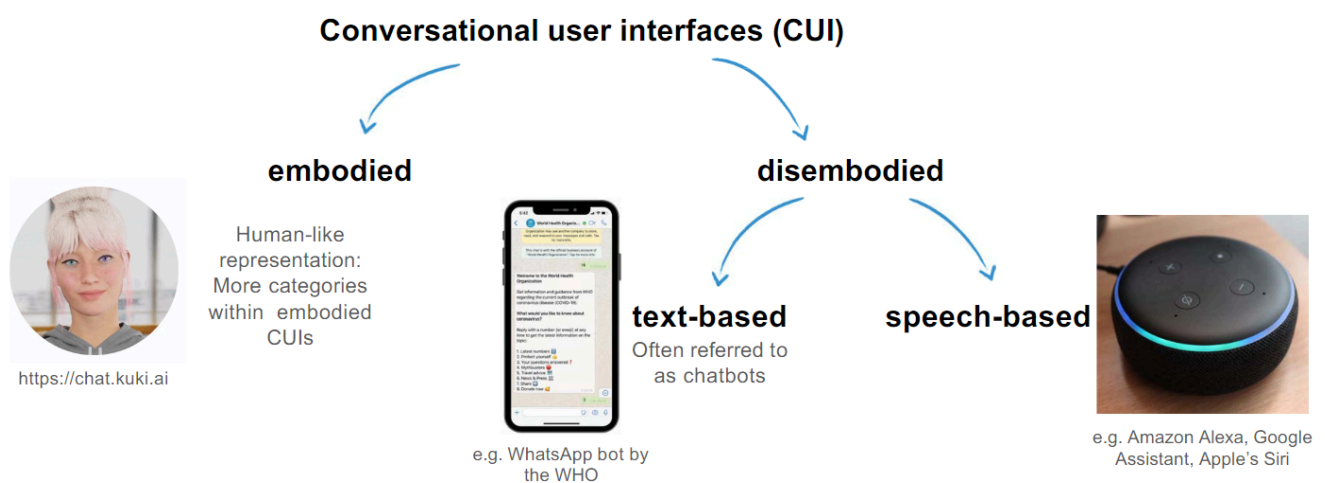# Lecture 4 - Voice User Interfaces (VUI)

What are Voice User Interfaces?

A **Voice User Interface (VUI)** enables users to interact with systems using spoken language. Unlike **Graphical User Interfaces (GUIs)**, VUIs rely entirely on **speech recognition and voice synthesis**.

> 💬 **Definition of VUI - Cohen et al. (2004)**
>
> "A Voice User Interface (VUI) is what a person interacts with when communicating with a spoken language application."

Conversational user interfaces (CUI)

embodied — Human-like representation: More categories within embodied CUIs
https://chat.kuki.ai

disembodied

text-based — Often referred to as chatbots
e.g. WhatsApp bot by the WHO

speech-based
e.g. Amazon Alexa, Google Assistant, Apple's Siri

Key Characteristics:

- Spoken **natural language interaction**.
- Hands-free and eyes-free usability.
- Often integrated with AI-driven **conversational agents**.

> 👌 **VUI**
>
> Language is the most crucial channel of communication between people

Key historical Milestones:

- **ELIZA (1964)** – Early chatbot simulating conversation.
- **IBM Watson (2006)** – Advanced AI-driven speech recognition.
- **Interactive Voice Response (IVR) Systems (~2000s)** – Automated phone-based speech systems.
- **Smart Assistants (2010s–present)** – Siri, Alexa, Google Assistant.

Advantages of VUIs:

- **Easy to learn** – Natural speech-based interaction.
- **Hands-free operation** – Useful in driving, cooking, accessibility.
- **Fast communication** – Speaking is often quicker than typing.

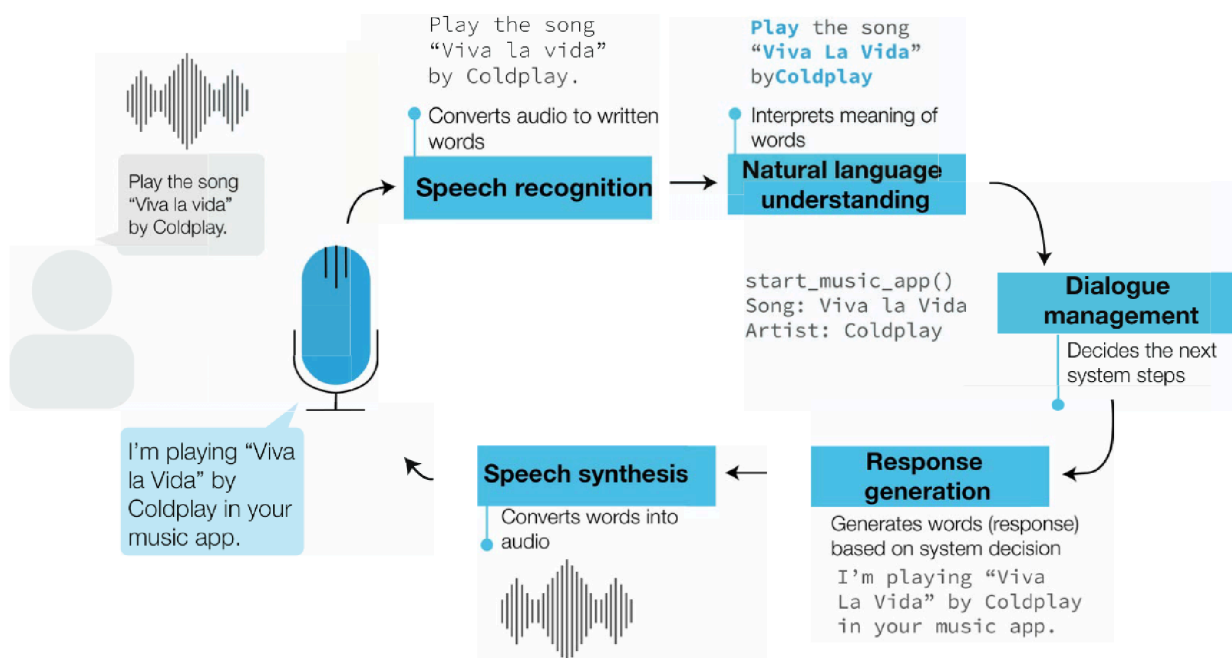- **Accessibility**

Challenges:

- **Speech recognition issues** – Accents, dialects, background noise.
  - **Homophones**: similar sounding words that have different meanings.
    - There vs their vs they're
    - Flower vs flour
  - Accents/dialects: Indian or Singapore English dialect
  - Speech disabilities: e.g., stuttering, Dysarthria (weakened speech muscles)
- **Lack of contextual understanding** – Difficulty with ambiguous phrases.
- **Privacy concerns** – Constant microphone listening raises security risks.
- Public and Group use can be difficult

💬 **Pragmatics**

Language always needs to be understood in the context, like:

- That's a typical example of a job well done
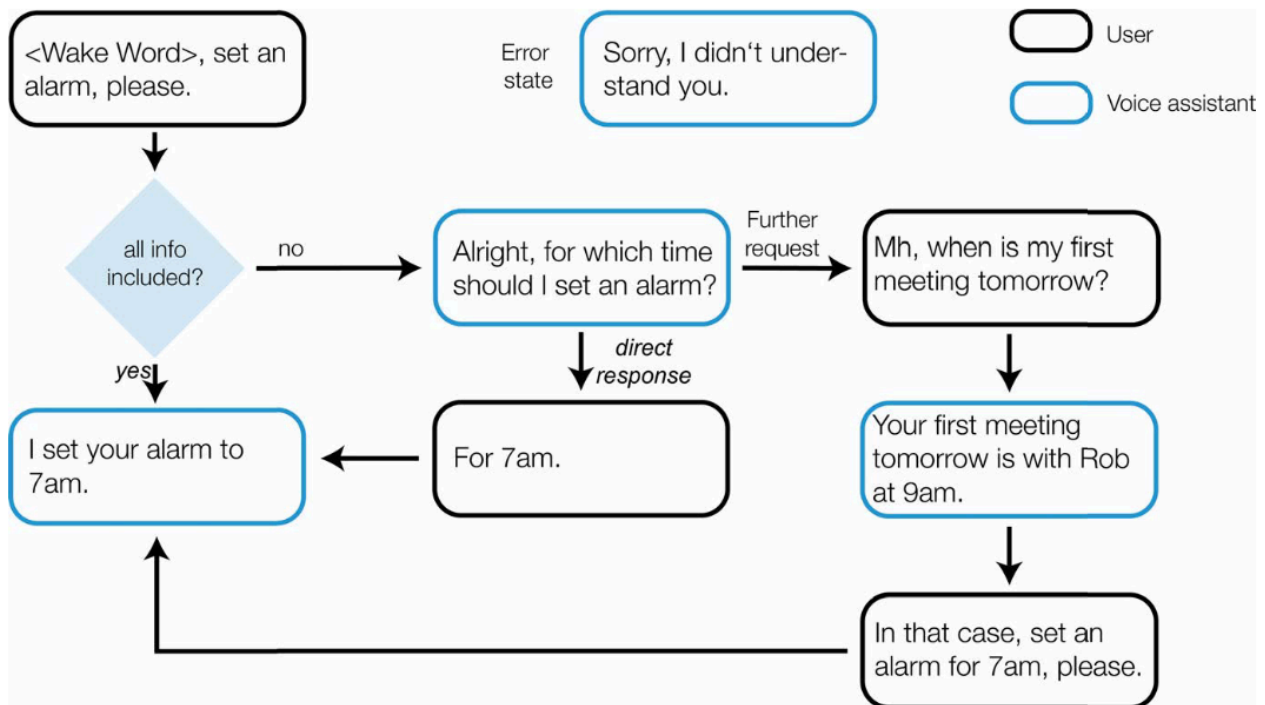- You've a green light

---

Design Principles for VUIs *(Slide 30-35)*

Best Practices:

1. **Guide users** – Indicate available commands clearly.

2. **Limit information overload** – Short responses improve usability.
   - inform users about their options
   - inform users about the current function; eg.: vui should not answer "done" but rather: "the timer has been set to 5min"
   - only list up to four items when giving feedback
3. **Confirm user input** – Reduce errors by repeating key responses.
4. **Use conversational flow** – Avoid robotic or rigid phrasing.
5. **Manage user expectations** – Clearly define what the VUI can and cannot do.

⟳ **Short-Term Memory Limitations**

Users can retain only **4-7 chunks** of spoken information. *(Miller, 1956)*

Humanizing VUI

- **Automatic Personality Attribution:**
  We naturally assign personalities to voice user interfaces (VUIs), even if none was intentionally designed. This subconscious process **affects user trust, likability, self-disclosure, and even purchasing decisions.**
- **Impact on User Behavior:**
  The personality perceived in a VUI can make users more inclined to interact, share personal details, or follow recommendations. Since many VUIs tend to adopt female voices, reinforcing stereotypes.

⟳ **Gender Stereotypes in VUI Design**

Historically, female voices have been perceived as more nurturing, approachable, and submissive, which may influence how users interact with the VUI. For example, users might be more willing to disclose personal information or follow suggestions from a female voice, simply because of these socially ingrained associations. While these design choices can enhance user comfort and engagement, they also

perpetuate gender norms and diminish the diversity of roles that both males and females can embody in automated systems. für einige wenige Sekunden

- **Avatar and Visual Feedback Options:**
  While using avatars can enhance emotional expression, they risk falling into the uncanny valley, where the human-like appearance becomes unsettling. Alternatively, non-anthropomorphic feedback (e.g., a glowing blue ring like the Alexa Dot) can communicate status without mimicking human traits.
- **Design Complexity and Ethical Considerations:**
  Designers face a dilemma: making a VUI as human as possible may boost engagement, but it also risks manipulation and reinforces gender stereotypes. The challenge lies in balancing intuitive, empathetic design with ethical transparency.

> ⓘ *Should we make VUIs as human as possible?*
>
> - **Pro:** A more human-like VUI can foster natural interaction, build trust, and increase user satisfaction by mimicking real human conversation.
> - **Contra:** Over-humanization might create unrealistic expectations, trigger discomfort (uncanny valley), and perpetuate stereotypes, ultimately compromising ethical design and user autonomy.

Challenges and Future Trends

Current Limitations:

- **Context interpretation** – Understanding **pragmatics** and **syntactic ambiguity**.
- **Emotional intelligence** – Recognizing tone and emotions remains difficult.
- **Bias in AI** – Some VUIs struggle with gender and language biases.

Future Directions:

- **Multimodal interaction** – Combining voice with gestures or visuals.
- **More personalized assistants** – Learning user preferences dynamically.
- **Ethical AI considerations** – Improving privacy and reducing biases.

> 🗩 **Social Perception of VUIs - Reeves & Nass (1996)**
>
> "Users apply the same social rules to VUIs as they do to human conversations."