

# **VIKI: Inteligência virtual com interação por voz aplicada à internet das coisas**

Aluno: Luís Henrique Goll  
Orientadora: Joyce Martins

# Roteiro

- Motivação
- Objetivos
- Fundamentação teórica
- Trabalhos correlatos
- Requisitos
- Ferramentas utilizadas
- Especificação
- Implementação
- Resultados e discussões
- Conclusões
- Demonstração

# Motivação

Por que as pessoas têm que se adaptar à tecnologia? Não deveria ser ao contrário, a tecnologia se adaptar aos indivíduos?

# Objetivo Geral

Desenvolver uma ferramenta de inteligência virtual com conversação por voz aplicada à Internet das Coisas.

# Objetivos Específicos

- Reconhecer e sintetizar voz, permitindo toda interação do usuário ser por voz, em português brasileiro
- Ser capaz de tomar ações a partir de um diálogo com o usuário, como por exemplo processar sentenças que façam sentido diante do contexto da última sentença reconhecida, dando a impressão de uma continuidade no diálogo
- Ser integrada com automação de alguns eletrodomésticos, tais como televisão e luzes, em um determinado ambiente

# Reconhecimento e Síntese de Voz

- Reconhecimento de voz é a capacidade do computador entender a linguagem natural do ser humano
- Síntese de voz é a capacidade do computador gerar um áudio, ou uma fala, da linguagem natural do ser humano

# Internet das Coisas

- Internet das Coisas consiste em integrar dispositivos do mundo físico com o mundo virtual
- Hoje em dia é possível ver diversos exemplos de dispositivos ligados a internet, tais como:
  - Ar condicionados
  - Televisões
  - Luzes
  - Portões elétricos
  - Entre outros

# Inteligência Virtual

- Inteligência Virtual é a capacidade de uma aplicação parecer inteligente diante do usuário, embora não seja, pois é apenas programada para compreender determinado ambiente ou situação
- Uma inteligência virtual, ou artificial, é muito comum em assistentes virtuais ou bots de ajuda e conversação



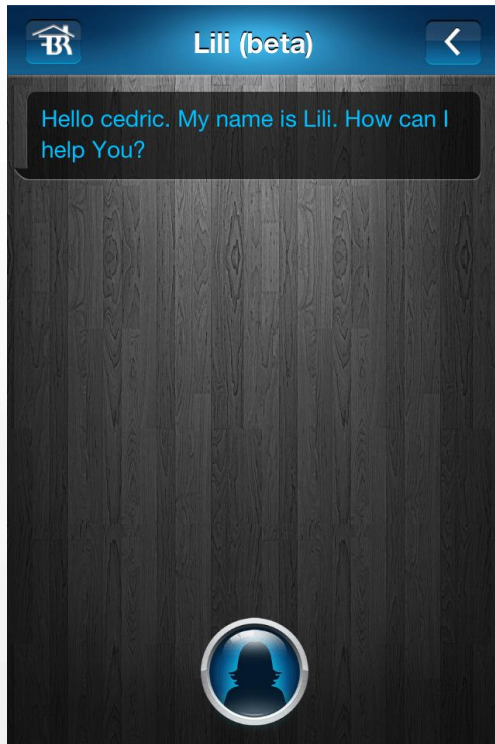
# Trabalhos Correlatos

## Dragon Naturally Speaking

Característica	
Reconhecimento de voz por comandos e sentenças específicas	Sim
Reconhecimento de voz livre	Sim
Reconhecimento e síntese do idioma português brasileiro	Não
Síntese de voz	Limitada
Capacidade de diálogo básico	Não
Integração com Internet das Coisas	Não

# Trabalhos Correlatos

## Fibaro Home System



Característica	
Reconhecimento de voz por comandos e sentenças específicas	Sim
Reconhecimento de voz livre	Não
Reconhecimento e síntese do idioma português brasileiro	Não
Síntese de voz	Limitada
Capacidade de diálogo básico	Não
Integração com Internet das Coisas	Sim

# Trabalhos Correlatos

## Cortana



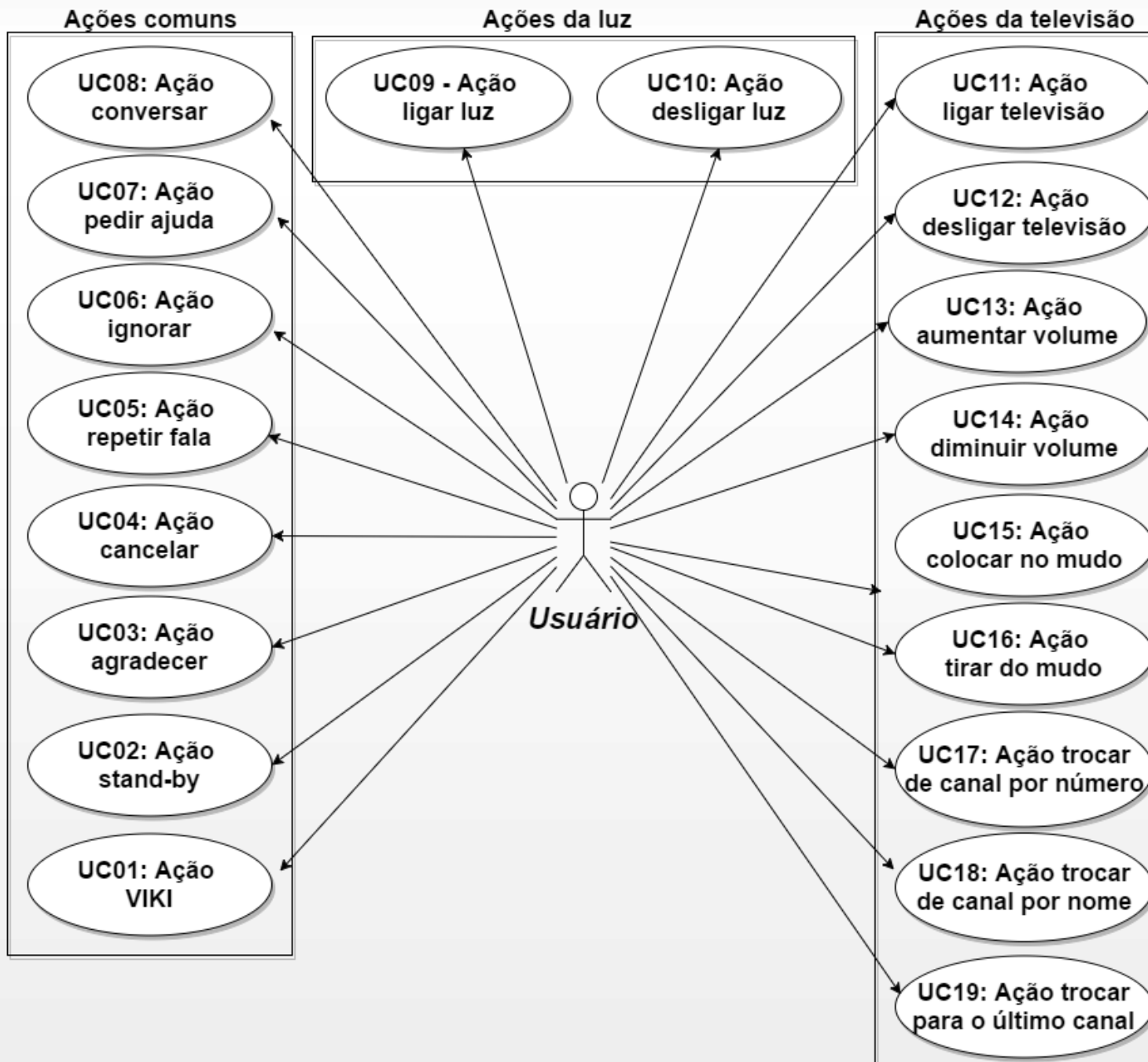
Característica	
Reconhecimento de voz por comandos e sentenças específicas	Sim
Reconhecimento de voz livre	Sim
Reconhecimento e síntese do idioma português brasileiro	Sim
Síntese de voz	Sim
Capacidade de diálogo básico	Não
Integração com Internet das Coisas	Não

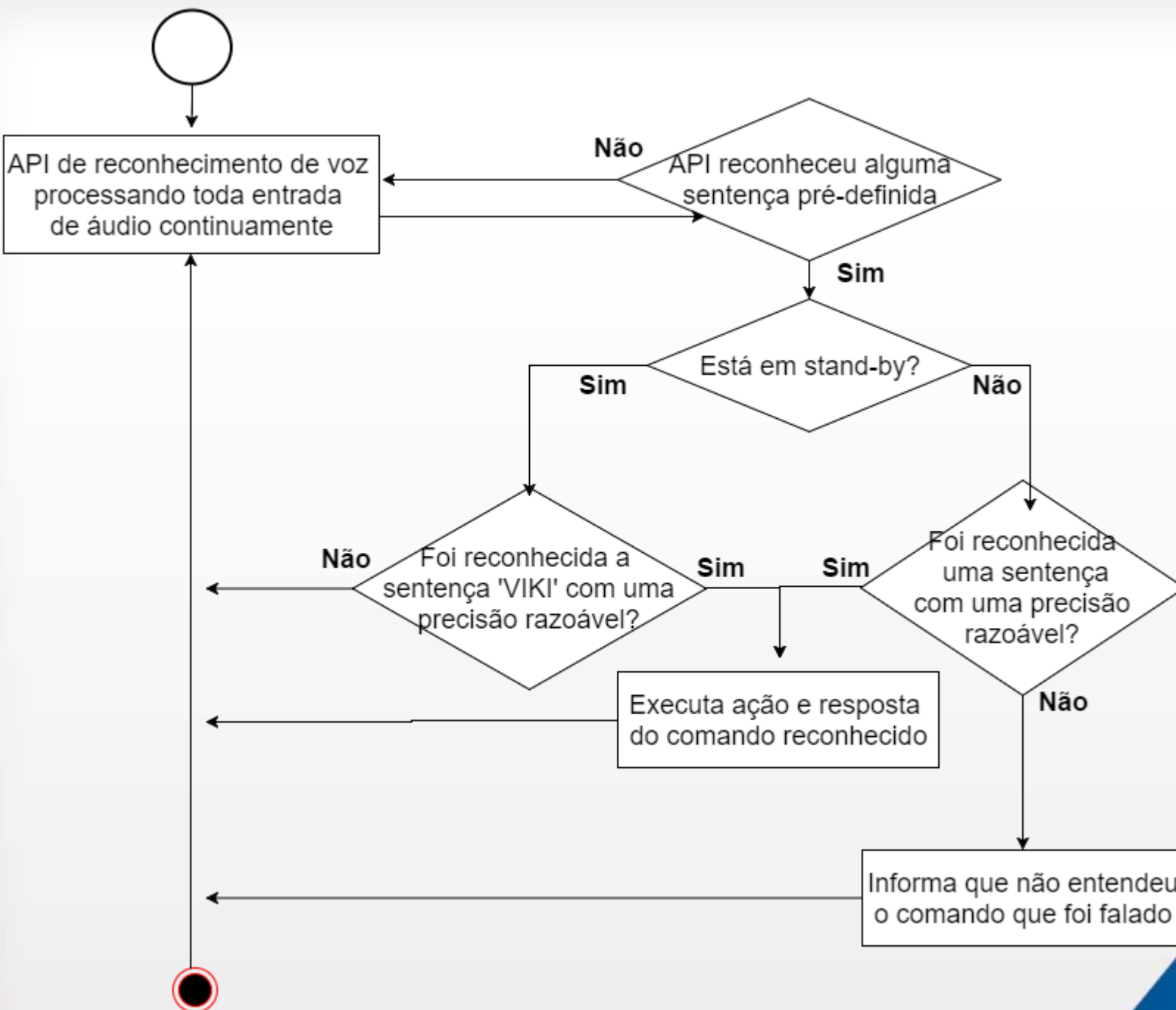
# Requisitos Funcionais

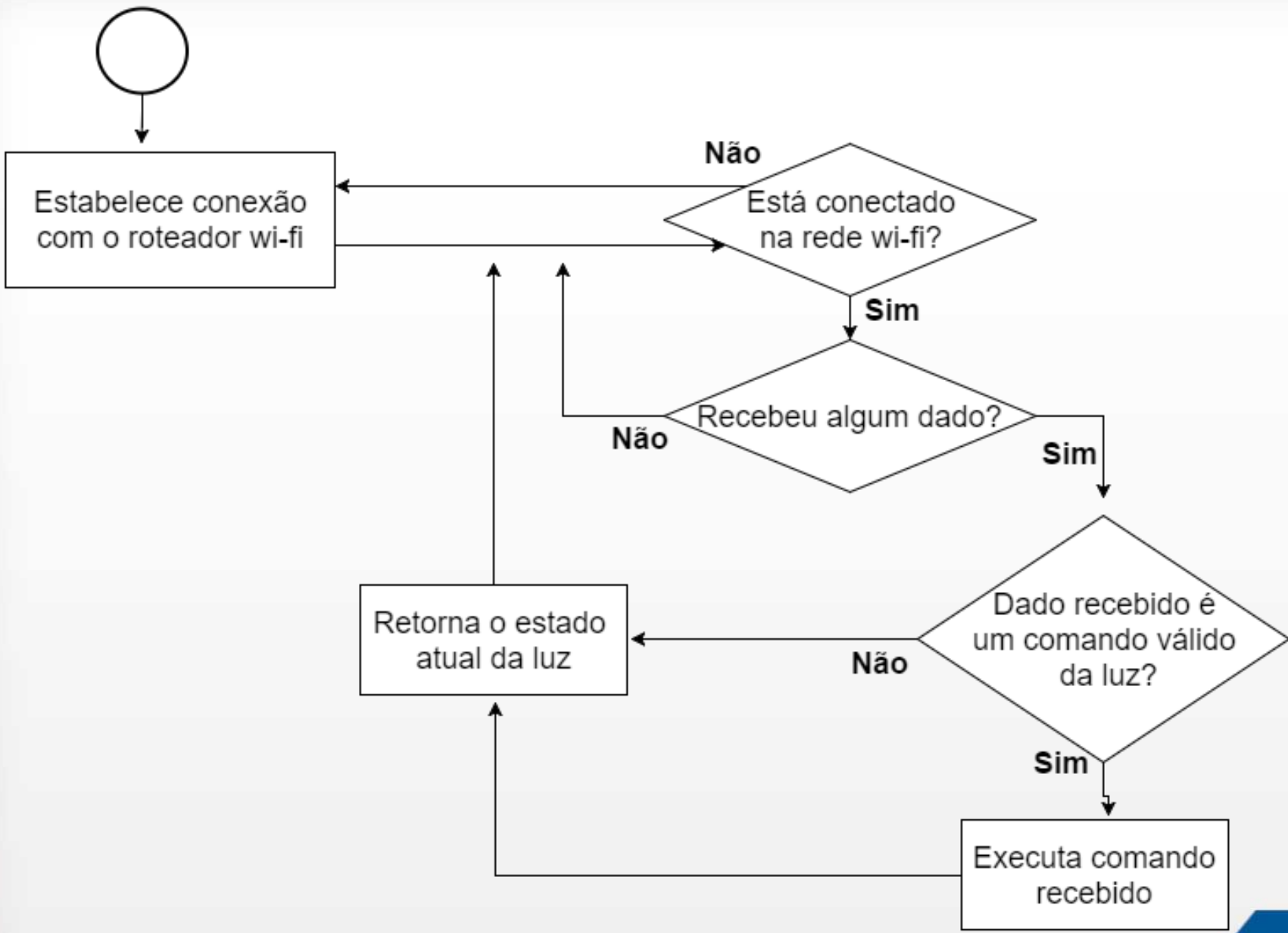
- Permitir o controle da luz a partir da aplicação
- Permitir o controle da televisão a partir da aplicação
- Ser capaz de questionar o usuário em caso de dúvidas e processar a resposta do usuário
- Ser capaz de reconhecer por voz sentenças pré-definidas
- Ser capaz de sintetizar respostas diferentes

# Requisitos Não-Funcionais

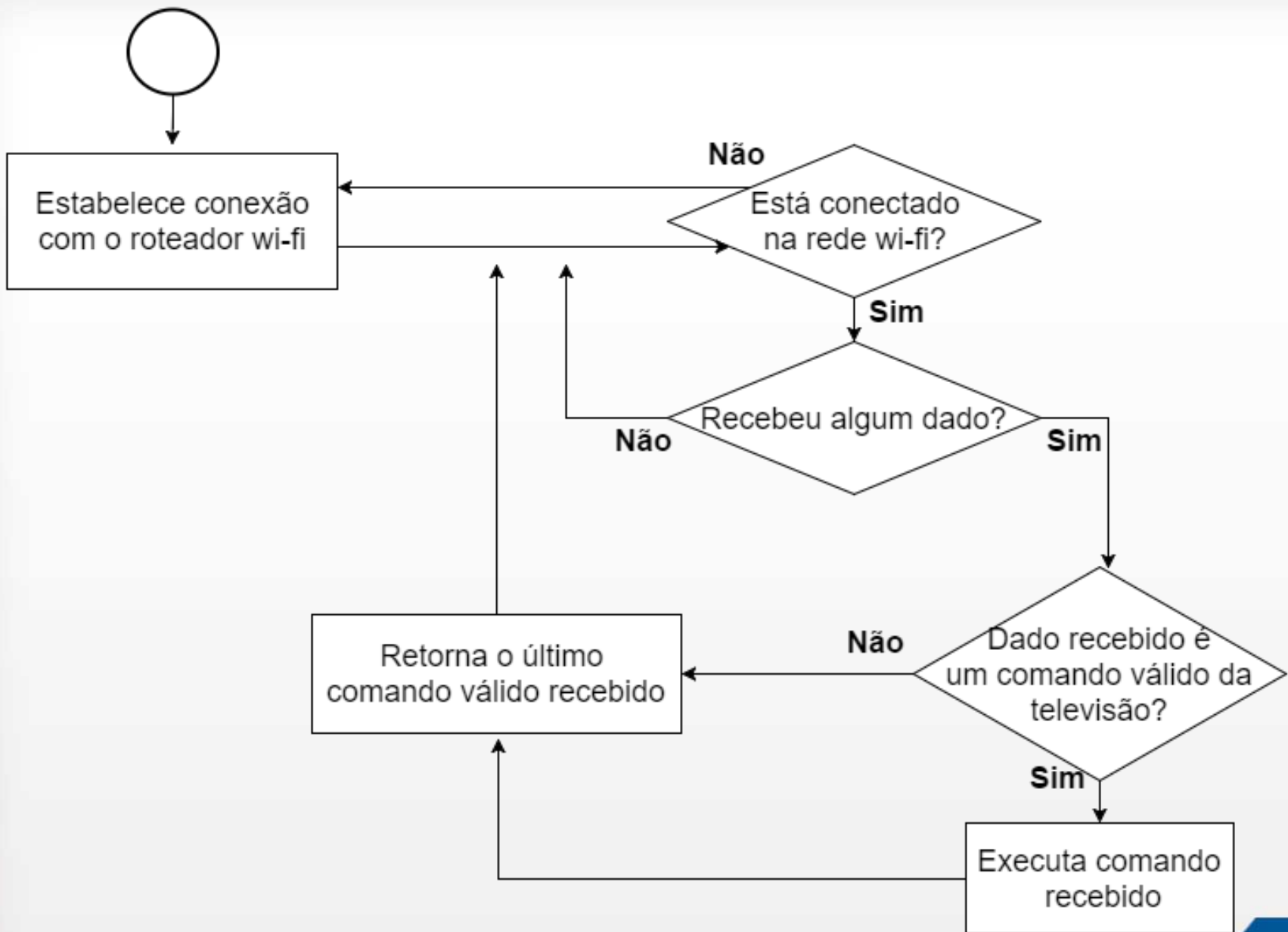
- A aplicação deve ser totalmente funcional por conversação por voz
- A aplicação deve ser de fácil utilização
- A aplicação deve se comunicar com microcontroladores via wi-fi para efetuar o controle dos dispositivos físicos











# Ferramentas Utilizadas - Softwares

- Visual Studio 2015 e C#
- API Microsoft Speech Platform
- Arduino IDE e C++ para Arduino
- Biblioteca ESP8266 para IDE do Arduino
- Biblioteca IRremote

# Ferramentas Utilizadas - Hardwares

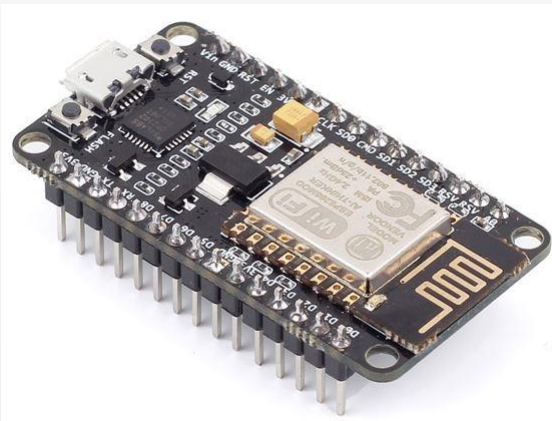
Placa AFSmart 031604 ESP8266-01



Fonte HLK-PM01



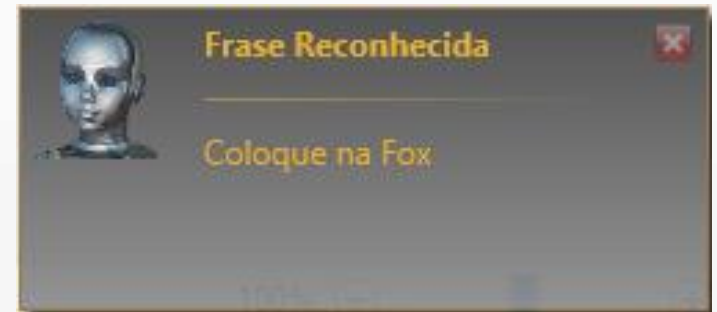
ESP8266 NodeMCU



Circuito Emissor Infravermelho



# Operacionalidade da VIKI



# Resultados e Discussões

- **Experimento 01:** uso da VIKI em um ambiente real
- **Experimento 02:** uso da VIKI em um ambiente simulado
- Comparativo em relação aos trabalhos correlatos

## Experimento 01: uso da VIKI em um ambiente real

Pergunta	Resposta
De forma geral, você conseguiu entender e utilizar a VIKI?	100% sim
Você conseguiu entender todas as falas e respostas da VIKI?	100% - 5
A VIKI entendeu tudo que você quis dizer, dentro dos comandos especificados?	25% - 4 75% - 5
A VIKI travou ou apresentou problema alguma vez?	100% - nunca
O quanto você gostou e achou interessante a VIKI?	100% - 5
Em um momento em que a VIKI esteja mais desenvolvida e disponível para o público, qual a chance de você recomendá-la para outra pessoa?	25% - 4 75% - 5

## Experimento 02: uso da VIKI em um ambiente simulado

Pergunta	Resposta
De forma geral, você conseguiu entender e utilizar a VIKI?	100% sim
Você conseguiu entender todas as falas e respostas da VIKI?	100% - 5
A VIKI entendeu tudo que você quis dizer, dentro dos comandos especificados?	62,5% - 4 37,5% - 5
A VIKI travou ou apresentou problema alguma vez?	62,5% - nunca 12,5% - uma vez ou outra 25% - outros
O quanto você gostou e achou interessante a VIKI?	25% - 4 75% - 5
Em um momento em que a VIKI esteja mais desenvolvida e disponível para o público, qual a chance de você recomendá-la para outra pessoa?	25% - 4 75% - 5

# Relatos dos experimentos

- Dificuldades:
  - Sentenças “ligue a televisão” e “desligue a televisão” podem ser identificadas incorretamente se o início da sentença não for falado de forma bem declarada
  - Algumas dificuldades com o instalador, em especial com o instalador da API o qual as vezes necessitava rodar duas vezes
  - Vozes de crianças não são bem reconhecidas pela API



# Comparativo em relação aos trabalhos correlatos

Característica	VIKI	Dragon Naturally Speaking	Fibaro Home System	Cortana
Reconhecimento de voz por comandos e sentenças específicas	Sim	Sim	Sim	Sim
Reconhecimento de voz livre	Não	Sim	Não	Sim
Reconhecimento e síntese do idioma português brasileiro	Sim	Não	Não	Sim
Síntese de voz	Sim	Limitada	Limitada	Sim
Capacidade de diálogo básico	Sim	Não	Não	Não
Integração com Internet das Coisas	Sim	Não	Sim	Não

# Conclusões

- Os objetivos de controlar a televisão e a luz por voz foram alcançados de forma satisfatória
- A usabilidade da VIKI foi bem aprovada nos experimentos pelos usuários de maneira geral

# Limitações

- Incapacidade de sintetizar perguntas de maneira realista com ênfase e entonação apropriada
- API de voz tem dificuldade para reconhecer vozes muito agudas como de crianças

# Extensões

- Desenvolver ou usar um sintetizador capaz de elaborar perguntas, em especial dando ênfase e entonação para as frases sintetizadas
- Processar diálogos e contextos mais aprofundados, utilizando alguma estrutura de dados como árvores
- Adaptar para outros idiomas, como espanhol ou inglês, que já são suportados pela mesma API utilizada
- Ampliar a Internet das Coisas, permitindo que também possa acionar a luz por meio de um interruptor, que se comunique com a aplicação

# Demonstração