

# Xây Dựng Ứng Dụng Trợ Lý Ảo Cho Máy Tính Sử Dụng Google Speech API

Khóa luận tốt nghiệp Cử nhân Công nghệ thông tin

Bùi Trung Hải    Phạm Ngọc Tuấn

GVHD: TS. Ngô Minh Nhựt

Trường Đại học Khoa học tự nhiên TP. Hồ Chí Minh  
Khoa Công nghệ thông tin  
Hệ Cử nhân tài năng

Ngày 14 tháng 7 năm 2017

# Nội dung

- 1 Giới thiệu đề tài
  - Trợ lý ảo
  - Google Speech API
  - Mục tiêu đề tài
- 2 Ứng dụng Alexa
  - Mô hình hoạt động
  - Các module chính
  - Các chức năng
  - Giao diện ứng dụng
- 3 Kết quả - Hướng phát triển
  - Kết quả
  - Hạn chế
  - Hướng phát triển

# Giới thiệu đề tài

# Trợ lý ảo là gì?

- Ứng dụng trên máy tính hoặc thiết bị di động
- Hỗ trợ người dùng thực hiện nhiều loại công việc
- Nhận lệnh dưới dạng ngôn ngữ tự nhiên
- Tăng tính hiệu quả và tạo ra sự tự nhiên trong giao tiếp giữa người và máy
- Tạo cho người dùng những trải nghiệm mới và thú vị hơn khi sử dụng các thiết bị công nghệ

# Một số ứng dụng trợ lý ảo nổi tiếng



**Frequency with Which US Smartphone Users Use  
Smartphone Virtual Assistants for Search, Jan 2017**  
% of respondents



Note: ages 18+

Source: HigherVisibility survey as cited in company blog, Feb 7, 2017

223269

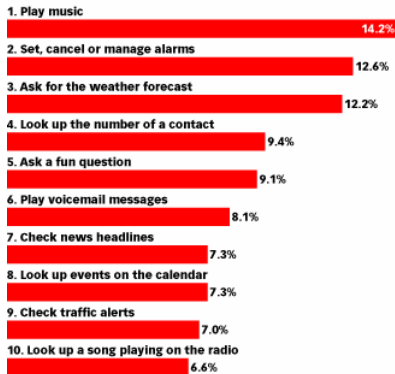
www.eMarketer.com

Tần suất sử dụng trợ lý ảo của người dùng di động tại Mỹ<sup>1</sup>

<sup>1</sup><https://www.emarketer.com/Article/How-Do-People-Use-Virtual-Assistants-on-Their-Smartphones/1015251>

### Top 10 Smartphone Virtual Assistant Search Queries/Requests According to US Smartphone Users, Jan 2017

% of respondents



Note: ages 18+; among respondents who use smartphone virtual assistants at least once a day

Source: HigherVisibility survey as cited in company blog, Feb 7, 2017

223271

www.eMarketer.com

Những chức năng được sử dụng nhiều nhất trên trợ lý ảo<sup>2</sup>

<sup>2</sup><https://www.emarketer.com/Article/How-Do-People-Use-Virtual-Assistants-on-Their-Smartphones/1015251>

# Các vấn đề cần giải quyết

- Phát hiện được keyword của người dùng để kích hoạt hệ thống
- Nhận biết câu lệnh của người dùng với độ chính xác cao
- Hiểu được câu lệnh của người dùng
- Có được những phản hồi chính xác tương ứng với câu lệnh



# Google Speech Recognition

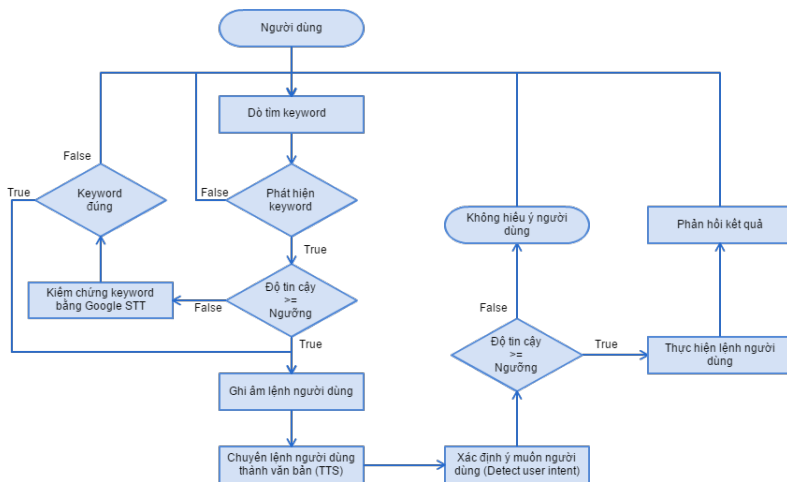
- Hệ thống nhận biết tiếng nói do Google nghiên cứu và phát triển
- Được Google sử dụng trong Google Assistant, Google Home, Google Search,...
- Độ chính xác đạt được hơn 95%
- Google cung cấp Google Speech API cho phép các nhà phát triển sử dụng Google Speech Recognition trong các ứng dụng.

# Mục tiêu đề tài

- Tìm hiểu cấu trúc cơ bản của một ứng dụng trợ lý ảo
- Xây dựng một ứng dụng trợ lý ảo trên máy tính, với các chức năng cơ bản:
  - Hỏi đáp, tra cứu thông tin
  - Hỏi giờ, thời tiết
  - Phát nhạc
  - Trò chuyện đơn giản
- Tìm hiểu và sử dụng những công cụ có sẵn để giải quyết các vấn đề và chức năng đã nêu.

# Ứng dụng Alexa

# Mô hình hoạt động



Mô hình hoạt động của ứng dụng Alexa

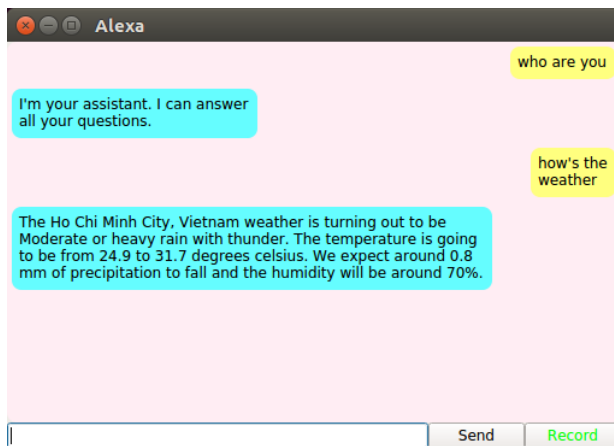
## Các module chính

- **Microphone:** Thực hiện chức năng thu âm, sử dụng thư viện PyAudio.
- **WakeUp:** Thực hiện chức năng dò tìm keyword, sử dụng thư viện pocketsphinx.
- **Recorder:** Ghi âm lệnh của người dùng và nhận biết khi người dùng dừng nói để dừng ghi âm.
- **Speech to Text:** Dùng để chuyển dữ liệu tiếng nói thành văn bản, để đưa lệnh người dùng về văn bản và kiểm chứng lại keyword, sử dụng Google Speech to Text.
- **Text to Speech:** Chuyển đổi văn bản thành giọng nói để phản hồi lại người dùng, sử dụng iSpeech API.
- **Intent Detector:** Xác định ý nghĩa của mệnh lệnh người dùng, sử dụng thư viện Rasa NLU.
- **Intent Processor:** Xử lý và đưa ra kết quả cho mệnh lệnh của người dùng.

# Các chức năng

- Thông báo giờ
- Thông báo thời tiết
- Phát nhạc
- Giao tiếp cơ bản
- Trả lời câu hỏi Wh-question

# Giao diện ứng dụng



Giao diện của ứng dụng Alexa

## Kết quả - Hướng phát triển



# Kết quả

- Tìm hiểu được cấu trúc cơ bản của một ứng dụng trợ lý ảo.
- Xây dựng được ứng dụng trợ lý ảo chạy trên nhiều nền tảng hệ điều hành khác nhau trên máy tính.
- Ứng dụng có các chức năng cơ bản đáp ứng được mục tiêu đề ra.
- Tìm hiểu được các thư viện xử lý âm thanh, speech to text, text to speech, natural language understanding,...
- Áp dụng các kỹ thuật phát triển phần mềm và các webservice API.

# Những điểm hạn chế

- Ứng dụng chỉ hỗ trợ tiếng Anh.
- Việc phản hồi không tạo thành một cuộc hội thoại mà chỉ dựa trên câu nói gần nhất của người dùng.
- Số lượng chức năng ít.
- Chức năng trả lời câu hỏi còn hạn chế.

# Hướng phát triển

- Hỗ trợ các ngôn ngữ khác tiếng Anh.
- Cải thiện, làm cho việc giao tiếp trở nên tự nhiên hơn.
- Tăng số lượng chức năng.
- Cải thiện khả năng trả lời câu hỏi.
- Thêm khả năng tương tác bằng hình ảnh.

Cảm ơn Thầy Cô và các bạn đã lắng nghe!