**BỘ GIÁO DỤC VÀ ĐÀO TẠO**

**TRƯỜNG ĐẠI HỌC ĐẠI NAM**

Logo, company name

Description automatically generated

**BÀI TẬP LỚN**

**HỌC PHẦN CÔNG NGHỆ ĐIỀU KHIỂN TIẾNG NÓI VỚI LOA THÔNG MINH**

**ĐỀ TÀI: Công nghệ điều khiển tiếng nói với loa thông minh.**

|  |  |
| --- | --- |
| Giảng viên hướng dẫn | : Th.S Đỗ Ngọc Điệp |
| Sinh viên thực hiện | : Nguyễn Việt Anh |
| Lớp | : CNTT-1805 |
| Mã sinh viên | : 1871020047 |

Hà Nội, năm 2024

**BỘ GIÁO DỤC VÀ ĐÀO TẠO**

**TRƯỜNG ĐẠI HỌC ĐẠI NAM**

Logo, company name

Description automatically generated

**BÀI TẬP LỚN**

**HỌC PHẦN CÔNG NGHỆ ĐIỀU KHIỂN TIẾNG NÓI VỚI LOA THÔNG MINH**

**ĐỀ TÀI: Công nghệ điều khiển tiếng nói với loa thông minh.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| STT | Mã sinh viên | Họ và tên | Ngày sinh | Điểm | |
| Bằng số | Bằng chữ |
| 1 | 1871020047 | Nguyễn Việt Anh | 28/10/2006 |  |  |

**CÁN BỘ CHẤM THI**

Hà Nội, năm 2024

**CHUYÊN ĐỀ:**

**Công nghệ điều khiển tiếng nói với loa thông minh**

****

**NGUYỄN VIỆT ANH**

**TÓM TẮT NỘI DUNG**

CHƯƠNG 1: GIỚI THIỆU TỔNG QUÁT CÔNG NGHỆ ĐIỀU KHIỂN BẰNG TIẾNG NÓI VỚI LOA THÔNG MINH.

CHƯƠNG 2: PHÂN TÍCH CƠ SỞ LÝ THUYẾT VỀ CÔNG NGHỆ ĐIỀU KHIỂN BẰNG TIẾNG NÓI VỚI LOA THÔNG MINH.

CHƯƠNG 3: TRIỂN KHAI / ỨNG DỤNG VỀ CÔNG NGHỆ ĐIỀU KHIỂN BẰNG TIẾNG NÓI VỚI LOA THÔNG MINH

**MỤC LỤC**

[**CHƯƠNG 1: GIỚI THIỆU TỔNG QUÁT** 6](#_Toc181906124)

[1.1. Lý do lựa chọn đề tài 6](#_Toc181906125)

[1.2. Mục tiêu nghiên cứu 6](#_Toc181906126)

[1.3. Đối tượng và phạm vi nghiên cứu 8](#_Toc181906127)

[1.4. Phương pháp nghiên cứu 8](#_Toc181906128)

[**CHƯƠNG 2: PHÂN TÍCH CƠ SỞ LÝ THUYẾT VỀ CÔNG NGHỆ ĐIỀU KHIỂN BẰNG TIẾNG NÓI VỚI LOA THÔNG MINH.** 9](#_Toc181906129)

[**2.1. Khái niệm công nghệ điều khiển bằng tiếng nói và vai trò của nó trong đời sống** 9](#_Toc181906130)

[***2.1.1. Định nghĩa công nghệ điều khiền bằng tiếng nói*** 11](#_Toc181906131)

[***2.1.2. Lịch sử phát triển và sự phổ biến công nghệ điều khiển bằng tiếng nói*** 11](#_Toc181906132)

[***2.1.3. Vai trò và tác dụng của công nghệ điều khiển tiếng nói*** 12](#_Toc181906133)

[***2.1.4. Tác động đến đời sống*** 13](#_Toc181906134)

[**2.2. Lợi ích của công nghệ điều khiển bằng tiếng nói trong cuộc sống** 14](#_Toc181906135)

[**2.3. Ứng dụng của công nghệ điều khiển bằng tiếng nói** 15](#_Toc181906136)

[**2.4. Ưu điểm của công nghệ điều khiển bằng tiếng nói** 16](#_Toc181906137)

[**2.5. Nhược điểm của công nghệ điều khiển bằng tiếng nói** 17](#_Toc181906138)

[**CHƯƠNG 3: TRIỂN KHAI / ỨNG DỤNG VỀ CÔNG NGHỆ ĐIỀU KHIỂN BẰNG TIẾNG NÓI VỚI LOA THÔNG MINH.** 18](#_Toc181906139)

[**3.1. Triển khai công nghệ điều khiển bằng tiếng nói với loa thông minh** 18](#_Toc181906140)

[***3.1.1. Các thành phần chính*** 18](#_Toc181906141)

[***3.1.2. Quy trình triển khai*** 19](#_Toc181906142)

[**3.2. Ứng dụng công nghệ điều khiển bằng tiếng nói với loa thông minh** 20](#_Toc181906143)

[**3.3. Các nền tảng phát triển ứng dụng công nghệ điều khiển bằng tiếng nói** 21](#_Toc181906144)

[***3.3.1. Google Assistant*** 21](#_Toc181906145)

[***3.3.2. Amazon Alexa*** 21](#_Toc181906146)

[***3.3.3. Microsoft Azure*** 21](#_Toc181906147)

[***3.3.4. Nền Tảng Mở Nguồn*** 21](#_Toc181906148)

[**3.4. Thách thức và giải pháp của công nghệ điều khiển bằng tiếng nói với loa thông minh** 21](#_Toc181906149)

[***3.4.1. Thách thức*** 21](#_Toc181906150)

[***3.4.2. Giải pháp*** 22](#_Toc181906151)

[**3.5. Tương lai của công nghệ điều khiển bằng tiếng nói với loa thông minh** 23](#_Toc181906152)

[KẾT LUẬN 25](#_Toc181906153)

[TÀI LIỆU THAM KHẢO 27](#_Toc181906154)

**DANH MỤC TÓM TẮT**

|  |  |
| --- | --- |
| MFCC | : Mel Frequency Cepstral Coefficients |
| IoT | : Internet Vạn Vật |
| NLP | : Xử lý ngôn ngữ tự nhiên |
| RNN | : Mạng thần kinh tuần hoàn |
| CNN | : Mạng thần kinh hội tụ |

**DANH MỤC HÌNH ẢNH**

[Hình 1:Ứng dụng Kiki Assistant. 12](#_Toc182221879)

[Hình 2:AI và Yolo Bit 13](#_Toc182221880)

[Hình 3:Điều khiển xe lăn bằng Joystick. 15](#_Toc182221881)

[Hình 4:Thư viện nhận dạng giọng nói. 22](#_Toc182221882)

**DANH MỤC SƠ ĐỒ**

[Sơ đồ 1:Điều khiển nhà thông minh bằng giọng nói. 12](#_Toc182221895)

[Sơ đồ 2:Điều khiển bật/ tắt đèn. 18](#_Toc182221896)

[Sơ đồ 3:Hệ điều hành thời gian thực (RTOS). 22](#_Toc182221897)

[Sơ đồ 4:Xử lý ngôn ngữ tự nhiên (NLP). 23](#_Toc182221898)

[Sơ đồ 5:Tương tác Google Assistant. 25](#_Toc182221899)

# **CHƯƠNG 1: GIỚI THIỆU TỔNG QUÁT**

## Lý do lựa chọn đề tài

Ngày nay, xã hội càng hiện đại, khoa học kỹ thuật càng phát triển thì cuộc sống của con người càng có nhu cầu sử dụng đầy đủ các thiết bị thông minh để phục vụ cho sinh hoạt và công việc của mình. Một thực tế rất gần với con người là trong chính căn nhà của mình, mong muốn được dụng công nghệ tự động hóa càng được rộng rãi, tất cả đồ dùng trong nhà từ phòng ngủ, phòng khách đến toilet đều gắn các bộ điều khiển bằng giọng nói kết nối với Internet và điện thoại di động, cho phép chủ nhân điều khiển vật dụng từ xa hoặc lập trình cho thiết bị ở nhà hoạt động theo lịch thời gian đúng mong muốn.

Nhu cầu về kiểm soát hệ thống thiết bị điện và điều khiển bằng tiếng nói ngày càng phổ biến như kiểm tra trạng thái của đèn, quạt, máy lạnh, các thiết bị khác, có thể mở hay tắt và các thiết bị điện trong nhà từ xa bằng thiết bị điện thoại di động, thiết bị máy tính thông qua mạng internet. Hiện nay với nền khoa học phát triển với các trợ lý ảo thông minh hỗ trợ Al như: Google Assistant (Google), Alexa (Amazon), Siri (Apple), Cortana (Microsoft), Các thiết bị trong ngôi nhà thông minh ngoài được điều khiển qua các app trên điện thoại, web,... còn được điều khiển qua trợ lý ảo cũng dần được phát triển.

Vì vậy, em đã tìm hiểu, nghiên cứu, chọn đề tài: “ Công nghệ điều khiển bằng tiếng nói với loa thông minh" để làm bài tập lớn của mình. Kết quả nghiên cứu từ đề tài này sẽ giúp em có nhiều kinh nghiệm để nghiên cứu chế tạo hoàn chỉnh thiết bị điều khiển bằng tiếng nói với loa thông minh đáp ứng được sử dụng yêu cầu trên thị trường với giá thành hợp lý, chất lượng đảm bảo, phù hợp với điều kiện sống tại Việt Nam.

## Mục tiêu nghiên cứu

* Mục tiêu chung :

+ Tăng cường trải nghiệm người dùng: Làm cho việc tương tác với thiết bị điện tử trở nên tự nhiên, trực quan và tiện lợi hơn, giảm thiểu sự phụ thuộc vào các giao diện điều khiển truyền thống.

+ Mở rộng khả năng ứng dụng: Mở rộng phạm vi ứng dụng của công nghệ điều khiển bằng giọng nói vào nhiều lĩnh vực khác nhau trong cuộc sống, từ gia đình, văn phòng đến công nghiệp.

+ Cải thiện hiệu suất: Nâng cao độ chính xác của nhận diện giọng nói, tốc độ phản hồi của hệ thống và khả năng hiểu ngữ cảnh phức tạp.

+ Đảm bảo tính bảo mật: Xây dựng các hệ thống bảo mật an toàn để bảo vệ thông tin cá nhân của người dùng.

* Mục tiêu cụ thể :

+ Phát triển thuật toán

+ Nhận diện giọng nói: Cải thiện độ chính xác của việc nhận diện giọng nói trong các môi trường ồn ào, các giọng nói khác nhau và các ngôn ngữ khác nhau.

+ Xử lý ngôn ngữ tự nhiên: Nâng cao khả năng hiểu ý nghĩa của câu hỏi, yêu cầu và mệnh lệnh của người dùng, đặc biệt là các câu hỏi phức tạp hoặc có tính mơ hồ.

+ Học máy: Áp dụng các kỹ thuật học máy để giúp hệ thống tự học và thích nghi với từng người dùng.

+ Thiết kế giao diện:

Tạo ra các giao diện hội thoại tự nhiên: Thiết kế các cuộc hội thoại giữa người và máy trở nên trôi chảy và giống như giao tiếp giữa người với người.

Xây dựng các kịch bản tương tác đa dạng: Mở rộng phạm vi các tác vụ mà người dùng có thể thực hiện bằng giọng nói.

+ Tích hợp với các thiết bị và dịch vụ:

Kết nối với các thiết bị IoT: Tạo ra các hệ thống nhà thông minh, văn phòng thông minh.

+ Tích hợp với các dịch vụ trực tuyến: Cho phép người dùng truy cập và điều khiển các dịch vụ trực tuyến bằng giọng nói.

* Đánh giá và so sánh:

+ So sánh hiệu suất của các hệ thống khác nhau: Đánh giá độ chính xác, tốc độ và khả năng mở rộng của các hệ thống khác nhau.

+ Đánh giá trải nghiệm người dùng: Thu thập phản hồi từ người dùng để cải thiện hệ thống.

## Đối tượng và phạm vi nghiên cứu

* Đối tượng nghiên cứu

+ Công nghệ điều khiển bằng tiếng nói: Các thuật toán nhận dạng giọng nóichuyển đổi tín hiệu âm thanh thành văn bản ; Phân tích ngữ nghĩahiểu ý nghĩa của câu nói; Tạo ra các câu trả lời phù hợp với câu hỏi của người dùng.

+ Hệ thống loa thông minh: Thiết bị phần cứng để thu âm và phát âm ; Các thiết bị khác trong nhà hoặc văn phòng có thể được điều khiển bằng giọng nói.

* Phạm vi nghiên cứu

+ Thời gian: Nghiên cứu sẽ tập trung vào các công nghệ và ứng dụng hiện tại, trong khoảng thời gian từ 2020 đến 2024, để đảm bảo tính cập nhật và liên quan đến xu hướng mới nhất.

+ Địa lý: Nghiên cứu sẽ giới hạn trong các tổ chức và doanh nghiệp tại Việt Nam, nhằm tìm hiểu về nhu cầu và điều kiện thực tế của thị trường trong nước.

+ Ngành nghề: Tập trung vào các ngành nghề khoa học dữ liệu, kỹ sư phần mềm/ âm thanh/ máy,

+ Khía cạnh nghiên cứu: Đề tài sẽ xem xét cả khía cạnh kỹ thuật (công nghệ và thuật toán) và khía cạnh quản lý (quy trình, lợi ích, và thách thức trong ứng dụng).

## Phương pháp nghiên cứu

* Xây dựng hệ thống mẫu:

+ Lựa chọn nền tảng: Chọn một nền tảng phát triển phù hợp (ví dụ: Google Assistant, Amazon Alexa, Microsoft Azure, hoặc các nền tảng mã nguồn mở).

+ Thiết kế giao diện hội thoại: Xây dựng các câu lệnh, câu hỏi và các kịch bản tương tác mà người dùng có thể sử dụng.

+ Phát triển mô hình nhận dạng giọng nói: Sử dụng các thư viện và công cụ có sẵn hoặc tự xây dựng mô hình dựa trên các thuật toán học máy.

+ Tích hợp với các thiết bị: Kết nối loa thông minh với các thiết bị khác trong hệ thống (ví dụ: đèn, quạt, máy lạnh).

* Thu thập và xử lý dữ liệu:

+ Thu thập dữ liệu âm thanh: Ghi lại các mẫu âm thanh của người dùng nói các câu lệnh khác nhau.

+ Chuyển đổi dữ liệu: Chuyển đổi dữ liệu âm thanh thành các đặc trưng số (ví dụ: MFCC, Mel-spectrogram).

+ Gắn nhãn dữ liệu: Gắn nhãn cho từng đoạn âm thanh tương ứng với câu lệnh mà nó đại diện.

* Phân tích và tối ưu hóa:

+ Phân tích kết quả: Phân tích các lỗi thường gặp và tìm hiểu nguyên nhân.

+ Tối ưu hóa mô hình: Điều chỉnh các tham số của mô hình để cải thiện hiệu suất.

+ Mở rộng hệ thống: Thêm vào các tính năng mới và cải tiến hệ thống.

* Nghiên cứu định lượng: Sử dụng các phương pháp thống kê để phân tích dữ liệu và đánh giá hiệu suất của hệ thống.
* Nghiên cứu định tính: Tiến hành phỏng vấn, quan sát để thu thập ý kiến và đánh giá trải nghiệm người dùng.
* Nghiên cứu so sánh: So sánh hiệu suất của các hệ thống khác nhau.

# **CHƯƠNG 2: PHÂN TÍCH CƠ SỞ LÝ THUYẾT VỀ CÔNG NGHỆ ĐIỀU KHIỂN BẰNG TIẾNG NÓI VỚI LOA THÔNG MINH.**

## **2.1. Khái niệm công nghệ điều khiển bằng tiếng nói và vai trò của nó trong đời sống**

Công nghệ điều khiển bằng tiếng nói là một lĩnh vực của trí tuệ nhân tạo cho phép con người tương tác với các thiết bị điện tử thông qua ngôn ngữ tự nhiên. Thay vì phải sử dụng các nút bấm hay màn hình cảm ứng, người dùng chỉ cần nói ra lệnh và thiết bị sẽ thực hiện theo. Điều này làm cho việc sử dụng các thiết bị trở nên trực quan, tiện lợi và tự nhiên hơn.

Vai trò của Công nghệ Điều Khiển bằng Tiếng Nói trong Đời sống

Công nghệ này đang ngày càng trở nên phổ biến và đóng vai trò quan trọng trong nhiều lĩnh vực của cuộc sống

+ Tăng tính tiện lợi:

+ Giải phóng đôi tay: Người dùng có thể thực hiện nhiều tác vụ khác nhau mà không cần phải chạm vào thiết bị.

+ Tăng tốc độ: Nhiều thao tác có thể được thực hiện nhanh hơn chỉ bằng một câu lệnh.

+ Mở rộng khả năng tiếp cận:

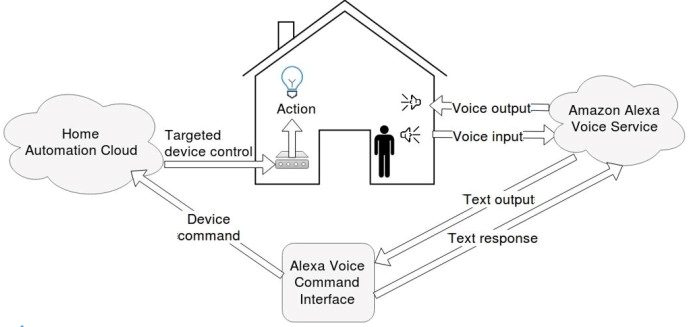
* Người khuyết tật: Giúp người khuyết tật dễ dàng sử dụng các thiết bị điện tử.
* Người già: Tạo điều kiện thuận lợi cho người già tương tác với công nghệ.

+ Tăng cường trải nghiệm người dùng:

* Cá nhân hóa: Các thiết bị có thể được tùy chỉnh để phù hợp với sở thích và thói quen của từng người.
* Tự động hóa: Nhiều tác vụ hàng ngày có thể được tự động hóa, giúp tiết kiệm thời gian và công sức.

+ Mở ra nhiều ứng dụng mới:

* Nhà thông minh: Điều khiển ánh sáng, nhiệt độ, thiết bị gia dụng, thiết bị giải trí, kiểm soát an ninh mở/khóa cửa, bật/tắt hệ thống báo động, tương tác với các thiếtbị IoT khác như điều khiển tủ lạnh, máy giặt, máy hút bụi...



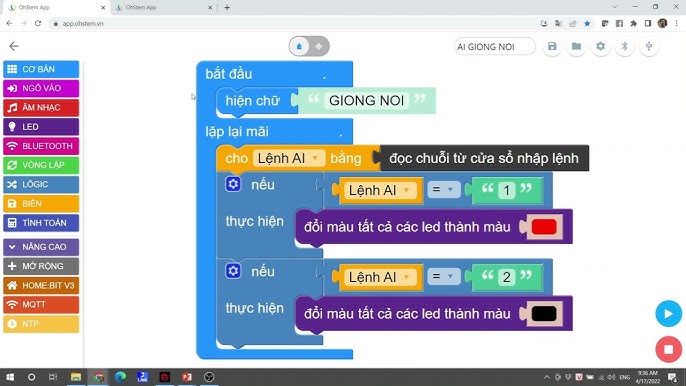
Sơ đồ 1:Điều khiển nhà thông minh bằng giọng nói.

* Ô tô: Điều khiển hệ thống giải trí bật/tắt radio, nhạc, điều chỉnh âm lượng, chuyển đổi giữa các nguồn âm thanh ; Điều khiển hệ thống điều hòa nhiệt độ, chuyển đổi chế độ làm mát/sưởi ; Điều khiển hệ thống định vị tìm kiếm địa điểm, lập trình hành trình ; Thực hiện cuộc gọi gọi điện thoại, trả lời cuộc gọi ; Tương tác với trợ lý ảo đặt câu hỏi, yêu cầu thông tin.



Hình 1:Ứng dụng Kiki Assistant.

* Y tế: Hệ thống ghi chú y tế bằng giọng nói giúp bác sĩ ghi chép nhanh chóng và chính xác thông tin về bệnh nhân ; Trợ lý ảo y tế cung cấp thông tin về các bệnh, thuốc, cách chăm sóc sức khỏe ; Ứng dụng theo dõi sức khỏe cá nhân ; Hệ thống điều khiển thiết bị y tế như điều khiển máy siêu âm, máy chụp X-quang, máy thở... bằng giọng nói.
* Giáo dục: Trợ lý ảo giải đáp thắc mắc trả lời các câu hỏi về bài học, cung cấp định nghĩa, giải thích các khái niệm ; Giáo viên có thể tạo các bài tập trắc nghiệm hoặc tự luận bằng giọng nói, hệ thống có thể tự động chấm điểm các bài làm trắc nghiệm.



Hình 2:AI và Yolo Bit

### 2.1.1. Định nghĩa công nghệ điều khiền bằng tiếng nói

Công nghệ điều khiển bằng tiếng nói được biết đến là một chương trình máy với một hệ thống vi xử lý lớn được tích hợp trong các thiết bị thông minh. Hiện nay công nghệ này đã được tích hợp ngay trên ứng dụng điện thoại thông minh.  
 Mặc dù hiện nay công nghệ nhân tạo Ai với khả năng nhận diện bằng giọng nói không còn quá xa lạ và được sử dụng khá phổ biến trên các sản phẩm thông minh, nhưng hiện nay vẫn giành được rất nhiều sự quan tâm  của giới người dùng ưa thích công nghệ.  
 Công nghệ thông minh, thông qua ứng dụng cho phép robot có thể nhận diện giọng nói, nhận diện hình ảnh và nhận diện được cả văn bản.

### 2.1.2. Lịch sử phát triển và sự phổ biến công nghệ điều khiển bằng tiếng nói

Công nghệ nhận diện giọng nói đã xuất hiện từ khá lâu, từ những năm 1950. Tuy nhiên, trong giai đoạn đầu, nó gặp nhiều hạn chế do sự phức tạp và chi phí cao.

Đến những năm 1970 và 1980, các nghiên cứu và nỗ lực trong lĩnh vực này đã dẫn đến sự ra đời của các hệ thống nhận diện bằng giọng nói đầu tiên. Những hệ thống này thường giới hạn về khả năng và chưa thực sự phổ biến trong ứng dụng thực tế.

Đến những năm 1990 và 2000, với sự phát triển mạnh mẽ của công nghệ máy tính và xử lý tín hiệu số, các hệ thống nhận dạng giọng nói bắt đầu trở nên hiệu quả hơn và được sử dụng rộng rãi trong các ứng dụng thương mại, như các trình ghi âm số và hệ thống điều khiển bằng giọng nói.

Đến nay, Công nghệ điều khiển bằng giọng nói đã trở thành một phần không thể thiếu trong cuộc sống hiện đại của chúng ta. Sự tiện lợi và tính tự nhiên của phương thức tương tác này đã nhanh chóng chiếm được cảm tình của người dùng trên toàn thế giới. Sự tiện lợi, tính tự nhiên và đa dạng ứng dụng đã khiến công nghệ này trở nên phổ biến và không thể thiếu trong cuộc sống hiện đại. Trong tương lai, chúng ta có thể kỳ vọng sẽ chứng kiến sự phát triển vượt bậc của công nghệ này, mang đến nhiều tiện ích và trải nghiệm mới mẻ cho người dùng.

### 2.1.3. Vai trò và tác dụng của công nghệ điều khiển tiếng nói

* Tăng cường trải nghiệm người dùng

**+** Tiện lợi: Cho phép người dùng thực hiện các tác vụ một cách nhanh chóng và dễ dàng chỉ bằng giọng nói, mà không cần phải thao tác bằng tay.

**+** Tự nhiên: Tương tác bằng giọng nói mang lại cảm giác tự nhiên và gần gũi hơn so với các phương thức tương tác khác.

**+** Đa nhiệm: Người dùng có thể thực hiện nhiều tác vụ cùng lúc, giải phóng đôi tay để làm các việc khác.

* Mở rộng khả năng tiếp cận

**+** Người khuyết tật: Công nghệ này giúp người khuyết tật, đặc biệt là người khiếm thị, có thể sử dụng các thiết bị điện tử một cách độc lập.



Hình 3:Điều khiển xe lăn bằng Joystick.

**+** Người già: Việc điều khiển thiết bị bằng giọng nói đơn giản hơn so với các phương thức khác, giúp người già dễ dàng sử dụng các thiết bị công nghệ.

**+** Trẻ em: Trẻ em có thể tương tác với các thiết bị một cách vui nhộn và học hỏi thông qua các lệnh thoại.

* Nâng cao hiệu quả làm việc

**+** Tự động hóa: Nhiều tác vụ lặp đi lặp lại có thể được tự động hóa bằng giọng nói, giúp tiết kiệm thời gian và công sức.

**+** Tăng năng suất: Người dùng có thể tập trung vào các công việc sáng tạo hơn khi các tác vụ đơn giản được thực hiện bằng giọng nói.

**+** Cải thiện quản lý thời gian: Việc tìm kiếm thông tin, đặt lịch hẹn, hoặc gửi email trở nên nhanh chóng và tiện lợi hơn.

* Ứng dụng đa dạng

**+** Gia đình: Điều khiển các thiết bị nhà thông minh, nghe nhạc, tìm kiếm thông tin.

**+** Ô tô: Điều hướng, thực hiện cuộc gọi, điều khiển hệ thống âm thanh.

**+** Văn phòng: Soạn thảo văn bản, gửi email, tham gia cuộc họp trực tuyến.

**+** Y tế: Hỗ trợ chẩn đoán, điều trị, và chăm sóc bệnh nhân.

* Thúc đẩy sự phát triển của các ngành công nghiệp

**+** Trí tuệ nhân tạo: Công nghệ nhận dạng giọng nói là một trong những ứng dụng quan trọng của trí tuệ nhân tạo.

**+** Internet of Things (IoT): Công nghệ này kết nối các thiết bị và cho phép chúng giao tiếp với nhau thông qua giọng nói.

**+** Xử lý ngôn ngữ tự nhiên: Việc hiểu và xử lý ngôn ngữ tự nhiên là yếu tố cốt lõi của công nghệ điều khiển bằng giọng nói.

### 2.1.4. Tác động đến đời sống

* Tăng cường tiện ích

**+** Rảnh tay: Điều khiển thiết bị mà không cần chạm vào, giúp người dùng tập trung vào các công việc khác.

**+** Tốc độ: Thực hiện các tác vụ nhanh chóng chỉ bằng một câu lệnh đơn giản.

**+** Đa nhiệm: Vừa làm việc khác vừa điều khiển thiết bị.

* Nâng cao trải nghiệm người dùng

**+** Tự nhiên: Giao tiếp bằng giọng nói giống như giao tiếp hàng ngày, tạo cảm giác thoải mái và thân thiện.

**+** Cá nhân hóa: Các trợ lý ảo có thể học hỏi và thích nghi với thói quen của người dùng, cung cấp những gợi ý và dịch vụ phù hợp.

* Mở rộng khả năng tiếp cận

**+** Người khuyết tật: Giúp người khuyết tật dễ dàng sử dụng các thiết bị điện tử.

**+** Người già: Đơn giản hóa việc sử dụng các thiết bị công nghệ.

**+** Trẻ em: Giúp trẻ em học hỏi và khám phá thế giới xung quanh một cách thú vị.

## **2.2. Lợi ích của công nghệ điều khiển bằng tiếng nói trong cuộc sống**

**-** Tăng cường trải nghiệm khách hàng: Hệ thống tổng đài định danh khách hàng thông qua giọng nói. Điều này giúp tiết kiệm thời gian và tăng trải nghiệm.

**-** Tăng tính bảo mật: Xác thực danh tính người dùng qua giọng nói. Thay vì nhập mật khẩu hoặc mã PIN. Điều này giúp ngăn chặn các cuộc tấn công và đảm bảo tính bảo mật cho hệ thống.

**-** Tăng năng suất: Giảm thiểu thời gian và công sức để thực hiện các nhiệm vụ quan trọng. Ví dụ, người dùng có thể sử dụng giọng nói để điều khiển các thiết bị hỗ trợ công việc, giúp tiết kiệm thời gian và nâng cao hiệu suất.

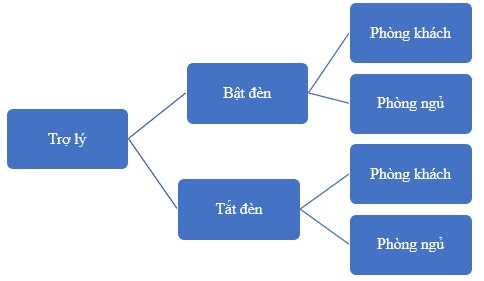
**-** Cải thiện khả năng truy cập: Giúp đỡ cho người dùng bị khiếm khuyết có thể truy cập các thiết bị và dịch vụ điện tử nhanh chóng và dễ dàng hơn.

**-** Cải thiện tính khả dụng: Cho phép người dùng truy cập vào các tài nguyên hoặc dịch vụ bằng cách nói chuyện với hệ thống. Đặc biệt là trong các tình huống mà sử dụng các giao diện truyền thống khó khăn hoặc bất tiện.

## **2.3. Ứng dụng của công nghệ điều khiển bằng tiếng nói**

* Điều khiển bằng tiếng nói

Hỗ trợ người dùng là một giải pháp khác từ công nghệ nhận diện tiếng nói. Các thiết bị điện tử cài đặt công nghệ này cho phép người dùng ra lệnh bằng tiếng nói thay vì bấm điều khiển. Ví dụ, đèn, điều hòa không khí, máy lạnh, máy giặt, máy sấy, và nhiều hơn nữa. Điều này giúp tăng tính tiện lợi và giảm thiểu việc phải sử dụng các phím bấm trên thiết bị.



Sơ đồ 2:Điều khiển bật/ tắt đèn.

* Ứng dụng trợ lý ảo cho ô tô

Trong thời đại số hóa, công nghệ điều khiển bằng tiếng nói còn được sử dụng trong các ứng dụng như trợ lý ảo, định hướng bản đồ, đặt lịch hẹn và điều khiển thiết bị. Nó cũng đang được sử dụng rộng rãi trong ngành công nghiệp như công nghiệp ô tô để kiểm soát giọng nói khi lái xe.

* Xác định danh tính người dùng

Phổ biến nhất là xác định danh tính của người dùng khi họ đăng nhập vào các trang web hoặc ứng dụng. Các hệ thống này sử dụng các kỹ thuật nhận diện giọng nói để so sánh giọng nói của người dùng với dữ liệu giọng nói được lưu trữ trước đó để xác định xem người dùng có phải là chủ sở hữu tài khoản hay không.

* Ứng dụng trong y tế

Chẩn đoán các bệnh lý liên quan đến giọng nói như thoái hóa giọng nói và bệnh Parkinson.

* Ứng dụng trong kinh doanh

Công nghệ nhận diện giọng nói có thể ứng dụng đa dạng ở nhiều nghiệp vụ và khía cạnh kinh doanh khác nhau.

+ Chăm sóc khách hàng: Các giải pháp trợ lý ảo AI hỗ trợ khách hàng giải quyết vấn đề khác nhau. Điều này giúp giảm thời gian chờ đợi và tăng tính tự động hóa trong chăm sóc khách hàng.

+ Quản lý tài khoản: Quản lý tài khoản bằng cách đăng nhập và thực hiện các giao dịch chỉ bằng giọng nói. Điều này có thể giúp giảm thời gian và tăng tính tiện lợi cho người dùng.

+ Quản lý kho: Công nghệ này được sử dụng để quản lý kho bằng cách cho phép nhân viên sử dụng giọng nói để đọc mã vạch hoặc nói tên sản phẩm khi lưu trữ hoặc xuất kho hàng hóa. Điều này có thể giúp giảm thời gian và tăng tính chính xác trong quản lý kho.

## **2.4. Ưu điểm của công nghệ điều khiển bằng tiếng nói**

* Tối ưu hiệu suất làm việc: có thể được sử dụng trong khi lái xe, hút bụi nhà, làm bữa tối hoặc bởi những người có khả năng di chuyển hạn chế;
* Tiết kiệm thời gian: mọi người đều biết rằng thực hiện cuộc gọi nhanh sẽ tốn ít thời gian hơn so với việc nhập email dài và chờ phản hồi. Điều này cũng tương tự với điều khiển giọng nói – bạn cắt giảm tất cả các lần nhấp bằng cách sử dụng nhận dạng giọng nói;
* Dễ dàng sử dụng:khả năng tiếp cận điều khiển bằng giọng nói thật khó tin, bất cứ ai cũng có thể sử dụng nó mà không cần bất kỳ kỹ năng kỹ thuật nào;
* Đa nền tảng và đa ngôn ngữ: điều khiển bằng giọng nói có thể được sử dụng từ bất kỳ thiết bị nào, cho dù đó là iOS hay Android và bằng bất kỳ ngôn ngữ nào. Năm 2018, thị trường điều khiển giọng nói của Mỹ trị giá 7,5 tỷ USD. Dự kiến sẽ tăng trưởng ba lần trong năm năm tới.
* Voice control ngày càng phổ biến và được ứng dụng tích hợp vào phần mềm quản lý smarthome. Chủ nhà có thể điều khiển các thiết bị trong nhà qua giọng nói, thông qua các thiết bị nhà thông minh mà không cần dùng tới smartphone. Khi bạn đi vắng có thể kiểm soát mọi thông tin, điều khiển các thiết bị từ xa bằng điện thoại smartphone, tablet ở bất kỳ đâu.

## **2.5. Nhược điểm của công nghệ điều khiển bằng tiếng nói**

* Độ chính xác còn thấp do một số tác động bên ngoài

+ Môi trường ồn ào: Trong các môi trường có nhiều tiếng ồn, hệ thống có thể khó khăn trong việc nhận diện chính xác các lệnh thoại.

+ Âm giọng đa dạng: Các giọng nói khác nhau, đặc biệt là giọng trẻ em, người già hoặc người có giọng nói đặc biệt, có thể gây khó khăn cho hệ thống.

+ Phát âm không chuẩn: Nếu người dùng phát âm không rõ ràng hoặc sai ngữ pháp, hệ thống có thể hiểu nhầm ý.

+ Ngôn ngữ và giọng địa phương: Hệ thống chưa hỗ trợ tốt các ngôn ngữ và giọng địa phương đa dạng, gây khó khăn cho người dùng.

+ Từ vựng chuyên ngành: Hệ thống có thể gặp khó khăn trong việc hiểu các thuật ngữ chuyên ngành hoặc các từ ngữ ít phổ biến.

+ Hệ thống có thể hiểu sai ý nghĩa của câu lệnh nếu không nắm bắt được ngữ cảnh cụ thể.

+ Thiếu linh hoạt chưa thể xử lý các câu hỏi mở hoặc các yêu cầu phức tạp đòi hỏi suy luận logic.

* Lo ngại về bảo mật và quyền riêng tư

+ Rò rỉ thông tin: Việc thu thập và lưu trữ dữ liệu giọng nói có thể tiềm ẩn nguy cơ rò rỉ thông tin cá nhân.

+ An ninh mạng chưa vững: Hệ thống có thể bị tấn công và dữ liệu giọng nói có thể bị kẻ xấu lợi dụng.

+ Một số thiết bị có thể ghi âm liên tục để cải thiện khả năng nhận diện, điều này có thể gây ra lo ngại về quyền riêng tư.

* Chi phí cao

+ Các thiết bị hỗ trợ công nghệ điều khiển bằng giọng nói thường có giá thành cao hơn so với các thiết bị thông thường.

+ Việc phát triển và nâng cấp các hệ thống nhận dạng giọng nói đòi hỏi chi phí lớn.

# **CHƯƠNG 3:** **TRIỂN KHAI / ỨNG DỤNG VỀ CÔNG NGHỆ ĐIỀU KHIỂN BẰNG TIẾNG NÓI VỚI LOA THÔNG MINH.**

## **3.1. Triển khai công nghệ điều khiển bằng tiếng nói với loa thông minh**

Công nghệ điều khiển bằng giọng nói đã trở thành một xu hướng phổ biến, đặc biệt khi kết hợp với các thiết bị loa thông minh. Để triển khai một ứng dụng như vậy, bạn cần hiểu rõ các thành phần chính và quy trình thực hiện.

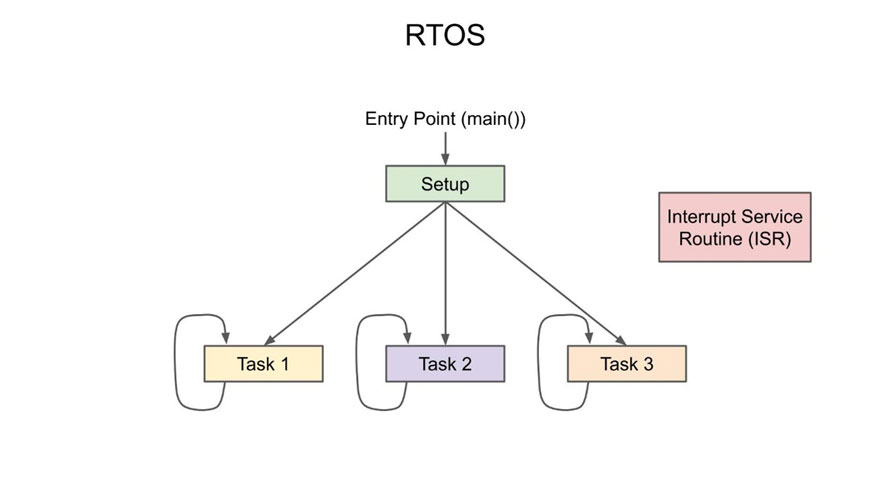
### 3.1.1. Các thành phần chính

**-** Thiết bị: Loa thông minh với micro chất lượng cao để thu âm giọng nói.

**-** Phần cứng: Bộ vi xử lý mạnh mẽ để xử lý các thuật toán nhận dạng giọng nói.

**-** Phần mềm:

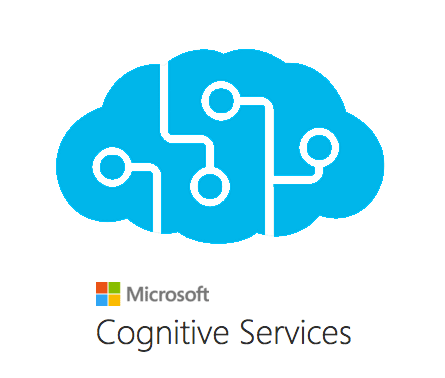
+ Hệ điều hành: Một hệ điều hành nhúng hoặc một hệ điều hành thời gian thực (RTOS) được tối ưu hóa cho các thiết bị IoT.



Sơ đồ 3:Hệ điều hành thời gian thực (RTOS).

+ SDK (Software Development Kit): Bộ công cụ phát triển phần mềm cung cấp các API và thư viện để tương tác với phần cứng và các dịch vụ đám mây.

+ Thư viện nhận dạng giọng nói: Các thư viện như Google Cloud Speech-to-Text, Amazon Transcribe, Microsoft Azure Speech Services, hoặc các thư viện mã nguồn mở khác.

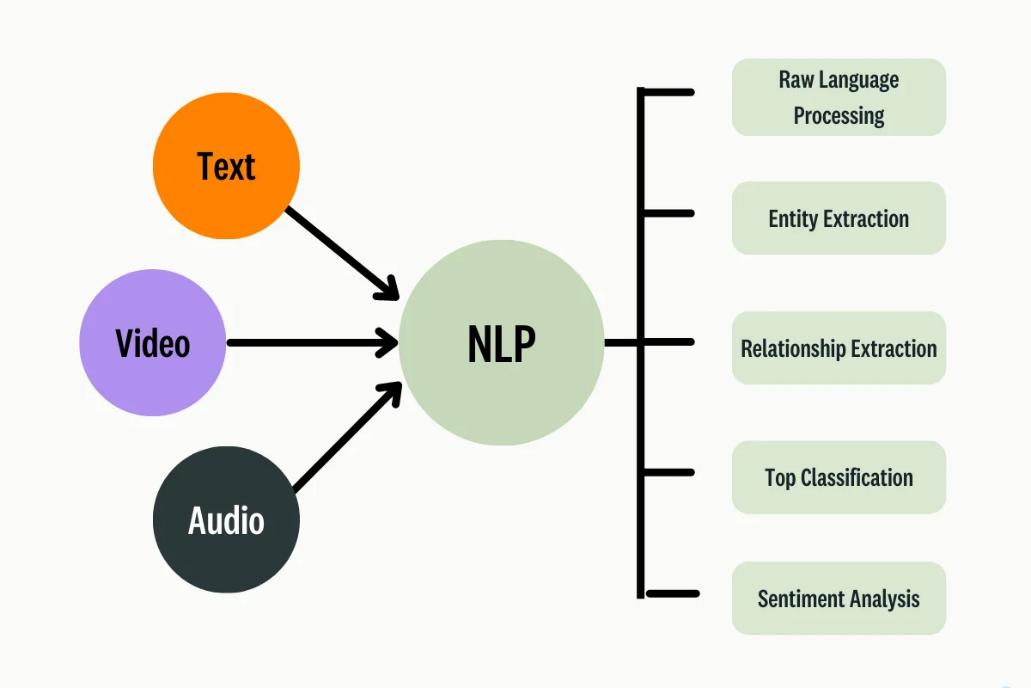
Hình 4:Thư viện nhận dạng giọng nói.

+ Khung ứng dụng: Một khung ứng dụng để xây dựng và quản lý các ứng dụng giọng nói.

* Dịch vụ đám mây:

+ Nhận dạng giọng nói: Chuyển đổi tín hiệu âm thanh thành văn bản.

+ Xử lý ngôn ngữ tự nhiên (NLP): Hiểu ý nghĩa của văn bản và trích xuất thông tin.



Sơ đồ 4:Xử lý ngôn ngữ tự nhiên (NLP).

+ Tương tác với các dịch vụ khác: Kết nối với các dịch vụ bên thứ ba như dịch vụ âm nhạc, dịch vụ nhà thông minh,…

### 3.1.2. Quy trình triển khai

- Thiết kế giao diện giọng nói:

+ Xác định các lệnh: Liệt kê các lệnh mà người dùng có thể sử dụng để điều khiển thiết bị.

+ Xây dựng ngữ cảnh: Xác định các ngữ cảnh khác nhau mà lệnh có thể được sử dụng.

+ Xây dựng các câu trả lời: Chuẩn bị các câu trả lời phù hợp cho từng lệnh.

* Phát triển ứng dụng:

+ Thu thập dữ liệu: Thu thập một lượng lớn dữ liệu âm thanh để huấn luyện mô hình nhận dạng giọng nói.

+ Huấn luyện mô hình: Sử dụng các thư viện nhận dạng giọng nói để huấn luyện mô hình nhận dạng giọng nói tùy chỉnh.

+ Tích hợp với phần cứng: Kết nối phần mềm với phần cứng của loa thông minh.

+ Xây dựng logic ứng dụng: Viết mã để xử lý các lệnh, tương tác với các dịch vụ khác và điều khiển thiết bị.

* Kiểm thử và tối ưu hóa:

+ Kiểm thử độ chính xác: Đánh giá độ chính xác của hệ thống nhận dạng giọng nói trong các điều kiện khác nhau.

+ Kiểm thử chức năng: Kiểm tra xem các lệnh có được thực hiện đúng không.

+ Tối ưu hóa hiệu năng: Tối ưu hóa hệ thống để giảm thiểu độ trễ và tăng cường tính ổn định.

* Triển khai:

+ Đóng gói ứng dụng: Đóng gói ứng dụng thành một gói cài đặt hoặc một hình ảnh hệ thống.

+ Cài đặt lên thiết bị: Cài đặt ứng dụng lên loa thông minh.

## **3.2. Ứng dụng công nghệ điều khiển bằng tiếng nói với loa thông minh**

Điều khiển nhà thông minh: Các thiết bị thông minh mà loa có thể điều khiển theo ngữ cảnh khác nhau và điều khiển được là điều hòa, đèn , ti vi, rèm cửa,… Không chỉ thế, loa thông minh còn có khả năng điều khiển các ngữ cảnh khác nhau, tạo kịch bản hay điều khiển hàng loạt các thiết bị chỉ trong một câu lệnh. Ví dụ khi khách đến chơi, bạn muốn đèn phòng khách sáng, điều hòa bật 26 độ và rèm cửa mở, bạn chỉ cần nói “Maika ơi, Tiếp khách” ; “Ok Google, Turn on the Table Light”; “Hey Google, What is the weather like today”,… thì mọi thứ sẽ diễn ra theo ý của bạn

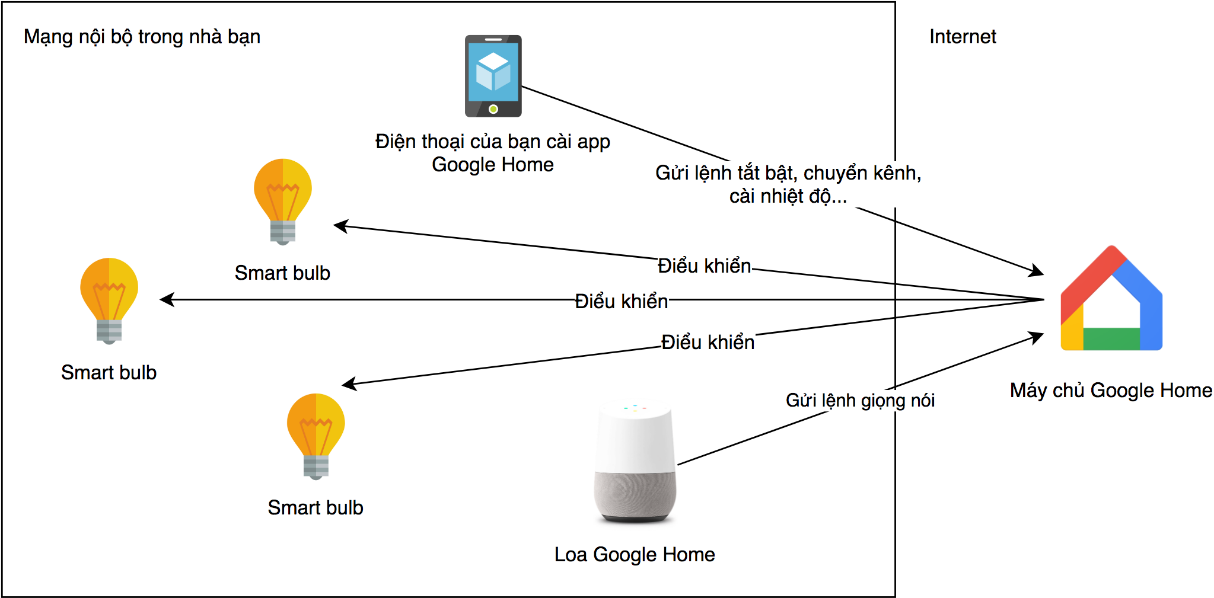
Phát nhạc và giải trí: Loa thông minh có các tính năng giúp bạn và gia đình có thể thư giãn sau những giờ làm việc căng thằng hay một ngày hoạt động mệt mỏi nhứ nghe nhạc, nghe radio, podcast, kể chuyện, cung cấp trò chơi,…

Yêu cầu thông tin và tra cứu: Loa có khả năng truy cập vào các dịch vụ trực tuyến để cung cấp thông tin và tra cứu thông tin cho bạn. Bạn có thể yêu cầu Maika cung cấp thời thiết, tin tức, kết quả thể thao, thông tin về sản phẩm, dịch vụ hay địa điểm chỉ cần qua giọng nói.

Tính năng trợ lý: Bạn co quá nhiều thứ phải làm trong ngày và cần có trợ lý ghi nhớ, nhắc nhở, sắp xếp danh sách công việc.

Các tiện ích khác: Bạn có thể yêu cầu gửi tin nhắn văn bản hoặc tin nhắn thoại cho các liên hệ trong danh bạ của bạn. Ngoài ra loa thông minh còn có thể gọi điện giữa các loa nhờ kết nối wifi hỗ trợ Bluetooth 5.0, dịch thuật, quy đổi tiền tệ,….

Tương tác với các ứng dụng và dịch vụ thông minh: Loa thông minh có khả năng tích hợp và tương tác với các ứng dụng và dịch vụ thông minh khác. Bạn có thể yêu cầu Google Assistant gửi tin nhắn, gọi điện, tìm kiếm thông tin trực tuyến, đặt hàng mua sắm trực tuyến, kiểm tra lịch trình và thực hiện nhiều tác vụ khác thông qua các ứng dụng và dịch vụ được hỗ trợ.



Sơ đồ 5:Tương tác Google Assistant.

Gửi lệnh từ xa: Nếu bạn kết nối loa thông minh với ứng dụng di động Google Home, bạn có thể điều khiển và tương tác với nhà thông minh ngay cả khi bạn không ở gần loa. Bằng cách sử dụng ứng dụng di động, bạn có thể bật/tắt thiết bị, thay đổi cài đặt, tạo lịch trình và thực hiện các tác vụ khác từ xa.

## **3.3. Các nền tảng phát triển ứng dụng công nghệ điều khiển bằng tiếng nói**

### 3.3.1. Google Assistant

Để tương tác với trợ lý ảo của mình, chúng ta có thể nói chuyện thông qua câu“OK Google”hoặc nhập qua bàn phím và tương tác trên màn hình thiết bị.

Với thuật toán xử lý ngôn ngữ tự nhiên của Google, Assistant có thể tham gia trò chuyện hai chiều giữa phần mềm và người dùng một cách tự nhiên và chân thực, đây cũng là cải tiến lớn nhất so với đàn anh Google Now - cũng do Google phát triển trước đó.

Những tiện ích của Google Assistant mang lại:

- Tìm kiếm thông tin bằng giọng nói giúp tiết kiệm thời gian.

- Nhận thông tin chỉ đường ngay lập tức bằng giọng nói mà không cần cầm vào điện thoại.

- Tìm câu trả lời: Công thức, kết quả phép tính, đổi đơn vị, xem lịch xe chạy,...

- Chạy trên nhiều nền tảng, thiết bị: Điện thoại, máy tính bảng, loa thông minh, đồng hồ thông minh, tivi,...

- Luôn an toàn và bảo mật.

### 3.3.2. Amazon Alexa

Tính năng cốt lõi của Alexa chính là khả năng điều khiển bằng giọng nói. Người dùng chỉ cần nói câu lệnh “Alexa” hoặc từ khóa đánh thức khác, sau đó đưa ra yêu cầu. Alexa sẽ xử lý và thực hiện hành động tương ứng.

Điều khiển bằng giọng nói mang lại sự tiện lợi đáng kể, đặc biệt khi tay đang bận hoặc không thể thao tác trực tiếp trên thiết bị. Ví dụ, khi đang nấu ăn, có thể yêu cầu Alexa đặt hẹn giờ hoặc chuyển bài hát mà không cần dừng công việc đang làm.

Alexa có khả năng nhận diện giọng nói khá tốt, ngay cả trong môi trường có tiếng ồn. Tuy nhiên, để đạt hiệu quả cao nhất, nên nói rõ ràng và sử dụng các câu lệnh chuẩn. Amazon cũng liên tục cải thiện khả năng nhận diện giọng nói của Alexa, giúp trợ lý ảo này ngày càng hiểu người dùng tốt hơn.

### 3.3.3. Microsoft Azure

* Ưu điểm:

+ Azure Cognitive Services: Cung cấp các dịch vụ nhận dạng giọng nói, tổng hợp giọng nói, và dịch thuật chất lượng cao.

+ Tích hợp với các sản phẩm Microsoft: Dễ dàng tích hợp với các sản phẩm khác của Microsoft như Office 365.

+ Đa nền tảng: Có thể sử dụng trên nhiều nền tảng khác nhau như Windows, Linux, và macOS.

* Ứng dụng: Phát triển các ứng dụng doanh nghiệp, chatbot, và các ứng dụng cần đến khả năng xử lý ngôn ngữ tự nhiên phức tạp.

### 3.3.4. Nền Tảng Mở Nguồn

* Ưu điểm:

+ Linh hoạt: Có thể tùy chỉnh và mở rộng theo nhu cầu của bạn.

+ Miễn phí: Không phải trả phí bản quyền.

+ Cộng đồng lớn: Có một cộng đồng người dùng lớn, giúp hỗ trợ và phát triển.

* Ví dụ: Kaldi công cụ nhận dạng giọng nói mạnh mẽ ; SpeechRecognition thư viện Python để nhận dạng giọng nói.
* Ứng dụng: Phát triển các ứng dụng nghiên cứu, các ứng dụng tùy chỉnh cao, và các ứng dụng cần đến tính bảo mật cao.

## **3.4. Thách thức và giải pháp của công nghệ điều khiển bằng tiếng nói với loa thông minh**

### 3.4.1. Thách thức

* Độ chính xác của nhận dạng giọng nói

Tiếng ồn xung quanh có thể làm giảm độ chính xác của việc nhận dạng giọng nói. Các giọng nói khác nhau, đặc biệt là giọng trẻ em, người già hoặc người có giọng nói đặc biệt, có thể gây khó khăn cho hệ thống. Nếu người dùng phát âm không rõ ràng hoặc sai ngữ pháp, hệ thống có thể hiểu nhầm ý. Hệ thống chưa hỗ trợ tốt các ngôn ngữ và giọng địa phương đa dạng.

* Bảo mật và quyền riêng tư:

Việc thu thập và lưu trữ dữ liệu giọng nói có thể tiềm ẩn nguy cơ rò rỉ thông tin cá nhân. Hệ thống có thể bị tấn công và dữ liệu giọng nói có thể bị kẻ xấu lợi dụng.

* Hiểu ngữ cảnh

Hệ thống khó hiểu được ý nghĩa của câu nói nếu không nắm bắt được ngữ cảnh cụ thể. Nhiều từ và câu có thể có nhiều nghĩa khác nhau, gây khó khăn cho việc xác định ý nghĩa chính xác.

* Tích hợp với các thiết bị và dịch vụ

Không phải tất cả các thiết bị và dịch vụ đều hỗ trợ công nghệ điều khiển bằng giọng nói. Chưa có một tiêu chuẩn chung cho các giao thức và giao diện điều khiển bằng giọng nói.

### 3.4.2. Giải pháp

* Cải thiện độ chính xác

Thu thập một lượng lớn dữ liệu âm thanh đa dạng để huấn luyện mô hình. Áp dụng các thuật toán học sâu tiên tiến như mạng thần kinh tuần hoàn (RNN), mạng thần kinh hội tụ (CNN) để cải thiện khả năng nhận dạng. Áp dụng các kỹ thuật xử lý tín hiệu để giảm nhiễu và tăng cường tín hiệu giọng nói.

* Bảo vệ bảo mật

Mã hóa dữ liệu giọng nói để bảo vệ thông tin cá nhân. Yêu cầu người dùng xác thực bằng nhiều cách khác nhau để tăng cường bảo mật. Sửa chữa các lỗ hổng bảo mật và cập nhật các tính năng bảo mật mới.

* Nâng cao khả năng hiểu ngữ cảnh

Tích hợp kiến thức về thế giới thực vào mô hình để hiểu rõ hơn về ngữ cảnh. Tạo ra các đồ thị kiến thức để liên kết các khái niệm và thông tin.

* Tăng cường khả năng tương tác

Kết hợp với các phương thức tương tác khác như cử chỉ, biểu cảm khuôn mặt. Tùy chỉnh trải nghiệm người dùng dựa trên lịch sử tương tác và sở thích.

## **3.5. Tương lai của công nghệ điều khiển bằng tiếng nói với loa thông minh**

Công nghệ điều khiển bằng giọng nói đang ngày càng trở nên phổ biến và phát triển mạnh mẽ. Với sự ra đời của các loa thông minh như Amazon Echo, Google Nest, và Apple HomePod, chứng kiến một cuộc cách mạng trong cách tương tác với các thiết bị điện tử.

* Khả năng hỗ trợ đa ngôn ngữ.

Các hệ thống nhận dạng giọng nói hàng đầu hiện nay cung cấp khả năng đa ngôn ngữ mạnh mẽ, cho phép người dùng tương tác bằng ngôn ngữ chính của họ. Điều này mở rộng phạm vi và tính khả dụng của các thiết bị và ứng dụng được điều khiển bằng giọng nói. Tuy nhiên, đây vẫn là một vấn đề lớn và liên tục cần giải quyết. Ở Hoa Kỳ một mình, có đến 30 ngôn ngữ địa phương chính. Ngay cả trong những phương ngôn đó, còn sự biến đổi từ người nói này sang người nói khác dựa trên giới tính, trình độ giáo dục, địa vị kinh tế và nhiều yếu tố dân số khác, chưa kể đến người nói tiếng thứ hai với chất giọng đặc biệt hoặc các cập nhật mới trong ngôn ngữ theo thời gian (ví dụ: từ mới hoặc ngôn ngữ lóng).

Vì vậy, để tạo ra công nghệ giọng nói hiểu được mọi người, các thuật toán nhận

dạng giọng nói cần được đào tạo trên dữ liệu giọng nói từ những người có đủ đa

dạng dân số.

* Cải thiện xử lý tiếng ồn

Các hệ thống nhận dạng giọng nói đã tiến xa trong việc loại bỏ tiếng ồn và xử lý tiếng ồn. Chúng có thể lọc ra âm thanh xung quanh, giúp dễ dàng bắt và hiểu chính xác giọng nói ngay cả trong môi trường ồn ào.

# KẾT LUẬN

CHƯƠNG 1:GIỚI THIỆU TỔNG QUÁT CÔNG NGHỆ ĐIỀU KHIỂN BẰNG TIẾNG NÓI VỚI LOA THÔNG MINH.

Nhu cầu về kiểm soát hệ thống thiết bị điện và điều khiển bằng tiếng nói ngày càng phổ biến như kiểm tra trạng thái của đèn, quạt, máy lạnh, các thiết bị khác, có thể mở hay tắt và các thiết bị điện trong nhà từ xa bằng thiết bị điện thoại di động, thiết bị máy tính thông qua mạng internet.

Mục tiêu nghiên cứu : Tăng cường trải nghiệm người dung, mở rộng khả năng ứng dụng, cải thiện hiệu suất, đảm bảo tính bảo mật. So sánh hiệu suất của các hệ thống khác nhau, đánh giá trải nghiệm người dùng.

Đối tượng nghiên cứu là công nghệ điều khiển bằng tiếng nói. Nghiên cứu sẽ tập trung vào các công nghệ và ứng dụng hiện tại, trong khoảng thời gian từ 2020 đến 2024.

Phương pháp nghiên cứu: Xây dựng hệ thống mẫu, thu thập và xử lý dữ liệu, phân tích và tối ưu hóa, nghiên cứu định lượng, nghiên cứu định tính, nghiên cứu so sánh.

CHƯƠNG 2: PHÂN TÍCH CƠ SỞ LÝ THUYẾT VỀ CÔNG NGHỆ ĐIỀU KHIỂN BẰNG TIẾNG NÓI VỚI LOA THÔNG MINH.

Công nghệ điều khiển bằng giọng nói đã trở thành một phần không thể thiếu trong cuộc sống hiện đại, đặc biệt với sự phát triển của các loa thông minh như Amazon Echo, Google Nest và Apple HomePod.

Vai trò của Công nghệ Điều Khiển bằng Tiếng Nói trong Đời sống tăng tính tiện lợi, giải phóng đôi tay, tăng tốc độ, mở rộng khả năng tiếp cận, tăng cường trải nghiệm người dung, mở ra nhiều ứng dụng mới

Sự tiện lợi và tính tự nhiên của phương thức tương tác này đã nhanh chóng chiếm được cảm tình của người dùng trên toàn thế giới. Sự tiện lợi, tính tự nhiên và đa dạng ứng dụng đã khiến công nghệ này trở nên phổ biến và không thể thiếu trong cuộc sống hiện đại.

Lợi ích của công nghệ điều khiển bằng tiếng nói trong cuộc sống tăng cường trải nghiệm khách hàng, tính bảo mật, năng suất. Cải thiện khả năng truy cập, tính khả dụng.Công nghệ điều khiển bằng tiếng nói có nhiều ưu điểm như tối ưu hiệu suất làm việc, tiết kiệm thời gian, dễ dàng sử dụng, đa nền tảng và đa ngôn ngữ. Voice control ngày càng phổ biến và được ứng dụng tích hợp vào phần mềm quản lý smarthome. Bên cạnh đó vẫn còn một số nhược điểm như độ chính xác còn thấp do một số tác động bên ngoài, lo ngại về bảo mật và quyền riêng tư, chi phí thiết bị, nâng cấp lớn.

CHƯƠNG 3: TRIỂN KHAI / ỨNG DỤNG VỀ CÔNG NGHỆ ĐIỀU KHIỂN BẰNG TIẾNG NÓI VỚI LOA THÔNG MINH.

Công nghệ điều khiển bằng giọng nói đã và đang trở thành một phần không thể thiếu trong cuộc sống hiện đại. Với sự phát triển của các trợ lý ảo như Google Assistant, Amazon Alexa và Siri, việc tương tác với thiết bị điện tử chỉ bằng giọng nói đã trở nên đơn giản và tiện lợi hơn bao giờ hết.

Các ứng dụng phổ biến của công nghệ điều khiển bằng giọng nói: Điều khiển nhà thông minh, truyền thông, lập lịch và nhắc nhở, giải trí, an ninh, y tế, giáo dục.

Thách thức độ chính xác của nhận dạng giọng nói trong môi trường ồn ào, các giọng nói khác nhau và các ngôn ngữ khác nhau, bảo vệ dữ liệu người dùng và ngăn chặn các cuộc tấn công, nâng cao khả năng hiểu ngôn ngữ tự nhiên phức tạp, bao gồm cả các câu hỏi mở và các câu hỏi đòi hỏi sự suy luận logic. Giải pháp sử dụng các kỹ thuật tăng cường dữ liệu, học chuyển tiếp và mô hình ngôn ngữ lớn, mã hóa dữ liệu, xác thực người dùng bằng nhiều yếu tố, cập nhật phần mềm thường xuyên, phát triển các mô hình ngôn ngữ lớn hơn và phức tạp hơn, sử dụng kiến thức thế giới để hiểu ngữ cảnh.

Tương lai của công nghệ điều khiển bằng giọng nói tích hợp sâu hơn vào cuộc sống hàng ngày, trợ lý ảo thông minh hơn, cá nhân hóa cao, tương tác đa phương thức.

# TÀI LIỆU THAM KHẢO

1. DOFA SMART, “Điều Khiển Bằng Giọng Nói: Giải Pháp Tiện Lợi Cho Công Tắc Thông Minh”, 03/ 09/ 2024, <https://dofa.vn/tin-tuc/dieu-khien-bang-giong-noi-giai-phap-tien-loi-cho-cong-tac-thong-minh>
2. WE SMART, “Voice Control - Công nghệ làm thay đổi cuộc sống con người”,

24/ 08/ 2020, <https://wesmart.vn/tin-tuc/voice-control-cong-nghe-lam-thay-doi-cuoc-song-con-nguoi.html>

1. HINACO Solutions, “Điều khiển nhà thông minh bằng giọng nói”, 07/08/2020, <https://hinaco.vn/dieu-khien-nha-thong-minh-bang-giong-noi>
2. Ms. Huyen, “Tìm hiểu tính năng điều khiển bằng giọng nói siêu đơn giản”, FIBARO a nice brand, 02/12/2023, <https://fibarovn.com/dieu-khien-bang-giong-noi/>
3. LUCI, “Các ứng dụng của công nghệ nhận dạng giọng nói”, 18/02/2024, <https://luci.vn/cac-ung-dung-cua-cong-nghe-nhan-dang-giong-noi>
4. ThS. Hồ Lê Anh Hoàng, “Hệ thống điều khiển thiết bị bằng giọng nói”, 18/05/2021, <https://ktcn.vhu.edu.vn/vi/tin-moi-1/he-thong-dieu-khien-thiet-bi-bang-giong-noi>