**TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT TP. HỒ CHÍ MINH**

**KHOA ĐÀO TẠO CHẤT LƯỢNG CAO**



**BÁO CÁO MÔN HỌC**

**ĐIỆN TOÁN ĐÁM MÂY**

**AMAZON POLLY**

**Nhóm 9 :**

**Nguyễn Lê Bảo Thanh 19110019**

**Le Sok 19110C01**

**Trần Hoàng Sơn 19110509**

**GVHD : Ths. Trương Thị Ngọc Phượng**

Tp. Hồ Chí Minh, tháng 12 năm 2021

# **LỜI CẢM ƠN**

Để hoàn thành tốt đề tài và bài báo cáo này, nhóm chúng em xin gửi lời cảm ơn chân thành đến cô, người đã trực tiếp hỗ trợ chúng em trong suốt quá trình làm đề tài. Chúng em cảm ơn cô đã đưa ra những lời khuyên từ kinh nghiệm thực tiễn của mình để định hướng cho chúng em đi đúng với yêu cầu của đề tài đã chọn, sẵn sáng giải đáp thắc mắc và đưa ra những góp ý, chỉnh sửa kịp thời giúp chúng em khắc phục vấn đề và hoàn thành bài đúng thời hạn đã đề ra.

Đề tài và bài báo cáo được chúng em thực hiện trong khoảng thời gian ngắn, với những kiến thức còn hạn chế cùng nhiều hạn chế khác về mặt kỹ thuật và kinh nghiệm trong việc thực hiện. Do đó, trong quá trình làm nên đề tài có những thiếu sót là điều không thể tránh khỏi nên chúng em rất mong nhận được những ý kiến đóng góp quý báu của các quý thầy cô để kiến thức của chúng em được hoàn thiện hơn và chúng em có thể làm tốt hơn nữa trong những lần sau. Chúng em xin chân thành cảm ơn.

Lời cuối cùng, chúng em xin kính chúc quý thầy, quý cô thật nhiều sức khỏe, thành công và hạnh phúc.

**MỤC LỤC**

[**LỜI CẢM ƠN** 2](#_Toc89537191)

[**PHẦN 1: MỞ ĐẦU** 4](#_Toc89537192)

[**1. Lý do chọn đề tài** 4](#_Toc89537193)

[**2. Tổng quan đề tài** 4](#_Toc89537194)

[**3. Phân công công việc** 4](#_Toc89537195)

[**PHẦN 2: NỘI DUNG** 5](#_Toc89537196)

[**CHƯƠNG 1 :TÌM HIỂU TỔNG QUAN** 5](#_Toc89537197)

[**1. Azamon Polly là gì ?** 5](#_Toc89537198)

[**2. Cách sử dụng** 5](#_Toc89537199)

[**3. Danh sách các voice trong Amazon Polly** 6](#_Toc89537200)

[**3.1. Available Voices** 7](#_Toc89537201)

[**3.2. Bilingual Voices** 10](#_Toc89537202)

[**4. Danh sách ngôn ngữ Amazon Polly hỗ trợ** 12](#_Toc89537203)

[**5. Lợi ích khi sử dụng Amazon Polly** 14](#_Toc89537204)

[**CHƯƠNG 2 :TÌM HIỂU CÁCH SỬ DỤNG** 15](#_Toc89537205)

[**1. Getting start với Amazon Polly** 15](#_Toc89537206)

[**1.1. Tạo tài khoản AWS và tạo user** 15](#_Toc89537207)

[**1.2. Sử dụng bằng console** 15](#_Toc89537208)

[**1.3. Sử dụng bằng AWS CLI** 15](#_Toc89537209)

[**CHƯƠNG 3 : CÀI ĐẶT VÀ SỬ DỤNG** 16](#_Toc89537210)

[**PHẦN 3: KẾT LUẬN** 17](#_Toc89537211)

[**TÀI LIỆU THAM KHẢO** 19](#_Toc89537214)

# **PHẦN 1: MỞ ĐẦU**

## **1. Lý do chọn đề tài**

Với tình hình đại dịch COVID – 19 đang diễn ra vô cùng phức tạp nên hầu hết mọi hoạt động đều thực hiện qua internet như học tập, làm việc, mua sắm,.. Do đó ngày càng có nhiểu người truy cập vào các website để làm việc chẳn hạn như các website tin tức, học tập, mua sắm,.. Nhìn chung đa số các webiste này đều ở dạng văn bản, mà đa phần mọi người thường có xu hướng thích nghe hơn là đọc văn bản. Lẽ đó việc tích hợp Amazon Polly – dịch vụ giúp chuyển văn bản thành lời nói – sẽ giúp việc truy cập các website này sẽ không còn nhàm chán nữa. Thay vì phải đọc các trang văn bản đầy chữ nhàm chán, giờ đây bạn có thể trải nghiệm nó bằng nghe. Điều sẽ là một trải nghiệm vô cùng độc đáo.

## **2. Tổng quan đề tài**

Amazon Polly là dịch vụ đám mây cho phép chuyển văn bản thành giọng nói, đề tài sẽ tìm hiểu cách tích hợp Amazon Polly vào website cho phép mọi người nghe tin tức thay vì đọc.

## **3. Phân công công việc**

|  |  |  |  |
| --- | --- | --- | --- |
| STT | Họ và tên | Mô tả công việc | Đóng góp |
| 1 | Nguyễn Lê Bảo Thanh | - Trưởng nhóm phân công công việc  - Tìm hiểu về Amazon Polly  - Viết báo cáo: tìm hiểu tổng quan về Amazon Polly  - Tổng hợp và viết hoàn chỉnh báo cáo | 40% |
| 2 | Le Sok | - Tìm hiểu cách sử dụng Amazon Polly  - Viết slide báo cáo | 30% |
| 3 | Trần Hoàng Sơn | - Viết website tích hợp Amazon Polly  - Viết báo cáo: cách tích hợp Amazon Polly vào website | 30% |

# **PHẦN 2: NỘI DUNG**

# **CHƯƠNG 1 :TÌM HIỂU TỔNG QUAN**

## **1. Azamon Polly là gì ?**

Amazon Polly là một dịch vụ đám mây giúp chuyển đổi văn bản thành lời nói. Có thể sử dụng Amazon Polly để phát triển các ứng dụng giúp tăng mức độ tương tác và khả năng tiếp cận. Amazon Polly hỗ trợ nhiều ngôn ngữ và bao gồm nhiều loại giọng, vì vậy có thể xây dựng các ứng dụng hỗ trợ giọng nói cho nhiều vị trí địa lí và sử dụng giọng nói thích hợp cho khách hàng. Ngoài ra cũng có thể lưu vào bộ nhớ cache và phát lại giọng nói do Amazon Polly tạo ra.

Ngoài ra, Amazon Polly còn bao gồm một số giọng nói chuyển văn bản thành giọng nói Neural (Neural Text-to-Speech - NTTS), mang lại những cải tiến đột phá về chất lượng về giọng nói thông qua một cách tiếp cận máy học mới, do đó cung cấp cho khách hàng một giọng nói được chuyển từ văn bản, thành giọng nói tự nhiên và giống như con người nhất có thể. Công nghệ NTTS cũng hỗ trợ tính cách của giọng nói được điều chỉnh cho phù hợp với các trường hợp sử dụng tường thuật tin tức.

Các trường hợp phổ biến sử dụng Amazon Polly bao gồm các ứng dụng di động như trình đọc tin tức, trò chơi, nền tảng eLearning, ứng dụng trợ năng cho người khiếm thị và Internet vạn vật (IoT) đang phát triển nhanh chóng.

Amazon Polly được chứng nhận để sử dụng, được quy định HIPAA (Health Insurance Portability And Accountability Act), và Tiêu chuẩn Bảo mật Dữ liệu Ngành Thẻ Thanh toán (PCI DSS).

## **2. Cách sử dụng**

Amazon Polly chuyển văn bản đầu vào thành lời nói. Người dùng có thể chọn một trong các Neural Text-to-Speech (NTTS) hoặc Standard Text-to-Speech (TTS) và chỉ định định dạng âm thanh đầu ra. Amazon Polly sau đó tổng hợp văn bản được cung cấp thành luồng âm thanh giọng nói chất lượng cao.

* Input text (văn bản đầu vào) : Là nơi bạn cung cấp văn bản mà bạn muốn chuyển thành giọn nói và Amazon Polly trả về một luồng âm thanh chính là giọng đọc văn bản. Bạn có thể cung cấp đầu vào dưới dạng văn bản thuần túy hoặc ở định dạng Speech Synthesis Markup Language (SSML). Với SSML bạn có thể kiểm soát cách phát âm, âm lượng, cao độ và tốc độ nói.
* Available voices: Amazon Polly hỗ trợ nhiều ngôn ngữ và nhiều loại giọng nói, bao gồm giọng nói song ngữ (ví dụ giọng có thể nói cả tiếng Anh và tiếng Hindi). Đối với hầu hết các ngôn ngữ, bạn có thể chọn từ một số giọng cả nam và nữ. Khi sử dụng Amazon Polly, bạn chỉ định ID giọng nói, sau đó Amazon Polly sử dụng giọng nói này để chuyển văn bản thành giọng nói. Cần lưu ý rằng Amazon Polly không phải là một dịch vụ dịch thuật (Lời nói tổng hợp có cùng ngôn ngữ với văn bản) . Tuy nhiên, nếu văn bản không phải là chữ, các số được biểu thị dưới dạng chữ số (ví dụ: 53, không phải “năm mươi ba”) thì nó sẽ được đọc theo giọng của giọng nói chứ không phải theo văn bản bạn nhập.
* Output format: Amazon Polly có thể cung cấp các giọng ở nhiều định dạng. Bạn có thể chọn định dạng âm thanh phù hợp với nhu cầu của bạn. Ví dụ: bạn có thể chọn MP3 hoặc Ogg. Định dạng Vorbis sử dụng trong các ứng dụng web và điện thoại di động. Hoặc PCM dùng trong các thiết bị AWS IoT và telephony solutions.

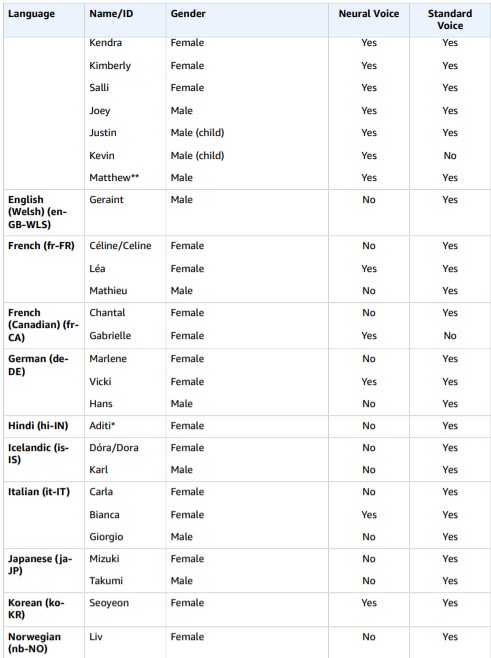
## **3. Danh sách các voice trong Amazon Polly**

Amazon Polly cung cấp một số giọng nói khác nhau để bạn sử dụng. Để nghe thử, bạn có thể vào <https://aws.amazon.com/vi/polly/> . Để nghe một giọng nói cụ thể, bạn có thể sử dụng Amazon Polly console.

### **3.1. Available Voices**

Amazon Polly cung cấp nhiều loại giọng nói khác nhau bằng nhiều ngôn ngữ để tổng hợp giọng nói từ văn bản







\* Voice này là song ngữ và có thể nói cả tiếng Anh và tiếng Hindi

\*\* Những voice này có thể được sử dụng với kiểu nói Newscaster khi được sử dụng với định dạng Neural.

Ngoài các voice trên, Amazon Polly có thể tạo cho bạn một Brand Voice tùy chỉnh phản ánh tính cách thương hiệu, mang đến cho bạn cơ hội đưa ra những giọng nói NTTS độc đáo và riêng biệt cho khách hàng. Để tìm hiểu thêm về Brand Voicee của Amazon Polly, bạn có thể truy cập <https://aws.amazon.com/vi/polly/features/#Brand_Voice>

### **3.2. Bilingual Voices**

Amazon Polly có hai loại giọng song ngữ (bilingual voice) :

• Giọng nói song ngữ có trọng âm (Accented bilingual voices)

• Giọng nói song ngữ hoàn toàn ( Fully bilingual voices)

#### **3.2.1. Accented bilingual voices**

Có thể tạo giọng nói song ngữ có trọng âm bằng bất kỳ giọng nói nào của Amazon Polly, nhưng chỉ khi sử dụng thẻ SSML.

Thông thường, tất cả các từ trong văn bản đầu vào được nói bằng ngôn ngữ mặc định của giọng nói mà bạn chỉ định đang sử dụng.

Ví dụ: nếu bạn đang sử dụng giọng nói của Joanna (người nói **tiếng Anh Mỹ**), Amazon Polly nói theo giọng Joanna không có **giọng Pháp**:

<speak>

Why didn't she just say, 'Je ne parle pas français?'

</speak>

Trong trường hợp này, các từ **Je ne parle pas français** được nói giống như là tiếng Anh.

Tuy nhiên, nếu bạn sử dụng giọng Joanna với thẻ <lang>, Amazon Polly sẽ nói theo giọng Joanna bằng tiếng Pháp có giọng Mỹ:

<speak>

Why didn't she just say, <lang xml:lang="fr-FR">'Je ne parle pas français?'</lang>.

</speak>

Vì Joanna không có giọng mẹ đẻ là tiếng Pháp nên cách phát âm dựa trên ngôn ngữ mẹ đẻ của cô, tiếng Anh Mỹ.

Ví dụ: mặc dù cách phát âm chuẩn chữ “r” trong từ français trong tiếng Pháp là “uvual trill”/ R / nhưng giọng tiếng Anh Mỹ của Joanna phát âm âm vị này là âm / r / thay vì / R /.

Nếu bạn sử dụng giọng nói của Giorgio, người nói tiếng Ý, với văn bản sau, Amazon Polly nói câu bằng giọng của Giorgio với cách phát âm tiếng Ý:

<speak>

Mi piace Bruce Springsteen.

</speak>

#### **3.3.2. Fully bilingual voices**

Một voice song ngữ hoàn toàn như Aditi (tiếng Anh Ấn Độ và tiếng Hindi) có thể nói thành thạo hai thứ tiếng. Điều này cho bạn khả năng sử dụng các từ và cụm từ từ cả hai ngôn ngữ trong một văn bản duy nhất bằng cách sử dụng cùng một giọng nói.

Hiện tại, Aditi là giọng nói song ngữ duy nhất có sẵn.

Aditi nói thông thạo cả tiếng Anh Ấn Độ (en-IN) và tiếng Hindi (hi-IN). Bạn có thể sử dụng voice này trong cả Tiếng Anh và tiếng Hindi, đồng thời giọng nói có thể chuyển đổi giữa hai ngôn ngữ ngay cả trong cùng một câu

Tiếng Hindi có thể được sử dụng ở hai dạng khác nhau:

* Devanagari: "उसेन कहँा, खेल तोह अब शुूर होगा"
* Romanagari (sử dụng bảng chữ cái Latin): "Usne kahan, khel toh ab shuru hoga"

Ngoài ra, có thể kết hợp tiếng Anh và tiếng Hindi của một hoặc cả hai dạng trong một câu duy nhất:

* Devanagari + tiếng Anh: "This is the song कभी कभी अदिति"
* Romanagari + tiếng Anh: "This is the song from the movie Jaane Tu Ya Jaane Na."
* Devanagari + Romanagari + tiếng Anh: "This is the song कभी कभी अदिति from the movie Jaane Tu Ya Jaane Na."

Vì Aditi là giọng nói song ngữ nên văn bản trong tất cả các trường hợp này sẽ được đọc chính xác, như Amazon Polly có thể phân biệt giữa các ngôn ngữ và script.

Amazon Polly cũng hỗ trợ số, ngày, giờ và tiền tệ bằng cả tiếng Anh (chữ số Ả Rập) và tiếng Hindi (chữ số Devanagari). Theo mặc định, các chữ số Ả Rập được đọc bằng tiếng Anh Ấn . Để Amazon Polly đọc chúng bằng tiếng Hindi, bạn phải sử dụng tham số mã ngôn ngữ hi-IN.

## **4. Danh sách ngôn ngữ Amazon Polly hỗ trợ**

|  |  |
| --- | --- |
| Language | Language Code |
| Arabic | arb |
| Chinese, Mandarin | cmn-CN |
| Danish | da-DK |
| Dutch | nl-NL |
| English, Australian | en-AU |
| English, British | en-GB |
| English, Indian | en-IN |
| English, New Zealand | en-NZ |

|  |  |
| --- | --- |
| Language | Language Code |
| Japanese | ja-JP |
| Korean | ko-KR |
| Norwegian | nb-NO |
| [Hindi](#_bookmark42) | hi-IN |
| English, South African | en-ZA |
| English, US | en-US |
| English, Welsh | en-GB-WLS |
| French | fr-FR |
| French, Canadian | fr-CA |
| German | de-DE |
| Icelandic | is-IS |
| Italian | it-IT |
| Polish | pl-PL |
| Portuguese, Brazilian | pt-BR |
| Portuguese, European | pt-PT |
| Romanian | ro-RO |
| Russian | ru-RU |
| Spanish, European | es-ES |
| Spanish, Mexican | es-MX |
| Spanish, US | es-US |
| Swedish | sv-SE |
| Turkish | tr-TR |
| Welsh | cy-GB |

## **5. Lợi ích khi sử dụng Amazon Polly**

* High quality: Amazon Polly cung cấp các neural TTS và công nghệ TTS tiêu chuẩn tốt nhất trong phân khúc để tổng hợp giọng nói tự nhiên vượt trội với phát âm có độ chính xác cao (bao gồm cả các từ viết tắt, tên viết tắt các tổ chức, ngày / giờ và từ giống nhau nhưng khác cách phát âm (homograph) ).
* Low latency: Amazon Polly đảm bảo phản hồi nhanh chóng, điều này làm cho nó trở thành một lựa chọn khả thi để sử dụng với độ trễ thấp các trường hợp chẳng hạn như hệ thống hộp thoại
* Support for a large portfolio of languages and voices: Amazon Polly hỗ trợ giọng nói bằng nhiều ngôn ngữ, cung cấp các tùy chọn giọng nói nam và nữ cho hầu hết các ngôn ngữ.
* Cost-effective: Amazon sử dụng mô hình trả tiền pay-per-use hay dùng bao nhiêu trả bấy nhiêu.
* Cloud-based solution: Các TTS trên thiết bị yêu cầu tài nguyên máy tính đáng kể, đáng chú ý là CPU, RAM và dung lượng ổ đĩa. Điều này có thể dẫn đến chi phí phát triển và công suất cao hơn cho các thiết bị như máy tính bảng, điện thoại thông minh, v.v. Ngược lại, TTS được thực hiện bằng AWS làm giảm đáng kể các yêu cầu về tài nguyên. Điều này cho phép hỗ trợ tất cả ngôn ngữ và giọng nói có sẵn ở chất lượng tốt nhất có thể. Hơn nữa, cải tiến giọng nói là có sẵn ngay lập tức cho tất cả người dùng cuối và không yêu cầu cập nhật bổ sung cho thiết bị.

# **CHƯƠNG 2 :TÌM HIỂU CÁCH SỬ DỤNG**

## **1. Getting start với Amazon Polly**

### **1.1. Tạo tài khoản AWS và tạo user**

#### **1.1.1. Tạo tài khoản AWS**

#### **1.1.2. Tạo IAM user**

### **1.2. Sử dụng bằng console**

#### **1.2.1. Synthesizing Speech**

#### **1.2.2. Synthesizing Speech với Plain Text**

### **1.3. Sử dụng bằng AWS CLI**

#### **1.3.1. Tạo AWS CLI**

#### **1.3.2. Sử dụng**

# **CHƯƠNG 3 : CÀI ĐẶT VÀ SỬ DỤNG**

# **PHẦN 3: KẾT LUẬN**



**TÀI LIỆU THAM KHẢO**