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

**KHOA ĐÀO TẠO CHẤT LƯỢNG CAO**



**BÁO CÁO MÔN HỌC**

**ĐIỆN TOÁN ĐÁM MAY**

**AMAZON POLLY**

**Nhóm 9 :**

**Nguyễn Lê Bảo Thanh 19110019**

**Le Sok**

**Trần Hoàng Sơn**

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

Tp. Hồ Chí Minh, tháng 12 năm 2021

# **LỜI CẢM ƠN**

**MỤC LỤC**

[**LỜI CẢM ƠN** 2](#_Toc89024269)

[**CHƯƠNG 1:** **TÌM HIỂU TỔNG QUAN** 6](#_Toc89024270)

[**CHƯƠNG 2:** **TÌM HIỂU CÁCH SỬ DỤNG** 7](#_Toc89024271)

[**CHƯƠNG 3:** **CÀI ĐẶT VÀ SỬ DỤNG** 8](#_Toc89024272)

[**CHƯƠNG 4:** **KẾT LUẬN** 9](#_Toc89024273)

[**TÀI LIỆU THAM KHẢO** 10](#_Toc89024276)

**DANH MỤC HÌNH**

**DANH MỤC BẢNG**

# **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 sống động như thật. 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. Với Amazon Polly, chỉ trả tiền cho văn bản bạn tổng hợp. Ngoài ra cũng có thể lưu vào bộ nhớ cache và phát lại giọng nói do Amazon Polly tạo ra mà không cần phải trả chi phí bổ sung.

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 với khối lượng công việc 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 sống động như thật. 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 nói và Amazon Polly trả về một luồng âm thanh chính 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 (cho 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.

# **TÌM HIỂU CÁCH SỬ DỤNG**

# **CÀI ĐẶT VÀ SỬ DỤNG**

# **KẾT LUẬN**



**TÀI LIỆU THAM KHẢO**