

TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT TP. HỒ CHÍ MINH
KHOA CÔNG NGHỆ THÔNG TIN
BỘ MÔN ĐIỆN TOÁN ĐÁM MÂY



TÊN ĐỀ TÀI:

TÌM HIỂU AMAZON TRANSLATE VÀ VIẾT ỨNG DỤNG MINH HỌA

GVDH: TS. Huỳnh Xuân Phụng

SVTH:

Diệp Tấn Luân 19110055

Phạm Minh Mẫn 19110397

Nguyễn Đức Sáng 19110448

TP.HCM, tháng 12 năm 2021

PHIẾU NHẬN XÉT CỦA GIÁO VIÊN HƯỚNG DẪN

MSSV 1: 19110055

MSSV 2: 19110397

MSSV 3: 19110448

Tên đề tài: **Tìm hiểu Amazon Translate và viết ứng dụng minh họa.**

Họ và tên Giáo viên hướng dẫn: **TS. Huỳnh Xuân Phụng**

NHẬN XÉT

1. Về nội dung đề tài & khối lượng thực hiện:

2. Ưu điểm:

3. Khuyết điểm:

4. Đề nghị cho bảo vệ hay không?

5. Đánh giá loại:

6. Điểm:

Tp. Hồ Chí Minh, ngày 20 tháng 12 năm 2021

Giáo viên hướng dẫn
(Ký & ghi rõ họ tên)

PHIẾU NHẬN XÉT CỦA GIÁO VIÊN PHẢN BIÊN

Họ và tên Sinh viên 1: **Diệp Tấn Luân** MSSV 1: **19110055**

Họ và tên Sinh viên 2: **Phạm Minh Mẫn** MSSV 2: **19110397**

Họ và tên Sinh viên 3: Nguyễn Đức Sáng **MSSV 3: 19110448**

Ngành: Công nghệ Thông tin

Tên đề tài: Tìm hiểu Amazon Translate và viết ứng dụng minh họa.

Họ và tên Giáo viên phản biện: **TS. Huỳnh Xuân Phụng**

NHẬN XÉT

1. Về nội dung đề tài & khối lượng thực hiện:

.....

2. Ưu điểm:

.....

3. Khuyết điểm:

.....

4. Đề nghị cho bảo vệ hay không?

.....

5. Đánh giá loại:

.....

6. Điểm:

.....

Tp. Hồ Chí Minh, ngày tháng năm 2021

Giáo viên phản biện

(Ký & ghi rõ họ tên)

LỜI CẢM ƠN

Lời đầu tiên nhóm xin phép được gửi lời cảm ơn chân thành và sâu sắc nhất đến với Khoa Công Nghệ Thông Tin – Trường Đại Học Sư Phạm Kỹ Thuật Thành Phố Hồ Chí Minh đã tạo điều kiện cho nhóm chúng em được học tập, phát triển nền tảng kiến thức sâu sắc và thực hiện đề tài này.

Bên cạnh đó nhóm chúng em xin gửi đến thầy Huỳnh Xuân Phụng lời cảm ơn sâu sắc nhất. Trải qua một quá trình dài học tập và thực hiện đề tài trong thời gian qua. Thầy đã tận tâm chỉ bảo nhiệt tình nhóm chúng em trong suốt quá trình từ lúc bắt đầu cũng như kết thúc đề tài này.

Với sự hướng dẫn nhiệt tình, giảng dạy tận tình đầy đủ kiến thức của thầy Huỳnh Xuân Phụng, chúng em đã học tập và hiểu được những kiến thức cơ bản về Amazon Translate – một nền tảng của Amazon hỗ trợ dịch thuật trên nhiều ngôn ngữ. Qua đó chúng em biết cách cài đặt và sử dụng Amazon Translate.

Tuy nhiên lượng kiến thức là vô tận và với khả năng hạn hẹp chúng em đã rất cố gắng để hoàn thành một cách tốt nhất. Chính vì vậy việc xảy ra những thiếu sót là điều khó có thể tránh khỏi. Chúng em hy vọng nhận được sự góp ý tận tình của thầy để qua đó chúng em có thể rút ra được bài học kinh nghiệm và hoàn thiện và cải thiện nâng cấp lại sản phẩm của mình một cách tốt nhất có thể.

Chúng em xin chân thành cảm ơn!

Nhóm thực hiện

Diệp Tấn Luân – 19110055

Phạm Minh Mẫn – 19110397

Nguyễn Đức Sáng – 19110448

MỤC LỤC

LỜI CẢM ƠN	4
MỤC LỤC	5
1. Tính cấp thiết của đề tài	7
2. Đối tượng nghiên cứu.....	7
3. Phạm vi nghiên cứu.....	7
4. Bố cục đề tài	8
5. Kết quả dự kiến đạt được	8
PHẦN NỘI DUNG	9
CHƯƠNG 1: CƠ SỞ LÝ THUYẾT VỀ AMAZON TRANSLATE.....	9
1. Tổng quan về Amazon Translate	9
1.1. Khái niệm	9
1.2. Lợi ích mà Amazon Translate đem lại	9
1.3. Trường hợp sử dụng	10
1.4. Các đặc điểm, tính năng của Amazon Translate	11
2. Các bước sử dụng dịch vụ Amazon Translate.....	14
3. So sánh và phân tích dịch vụ Amazon Translate và Google Translate	15
4. Tìm hiểu một số dịch vụ kết hợp vào đề tài	17
CHƯƠNG 2: XÂY DỰNG ỨNG DỤNG MINH HỌA.....	19
1. Giới thiệu sản phẩm đề tài.....	19
1.1. Giới thiệu	19
1.2. Nền tảng công nghệ.....	19
1.3. Chức năng.....	20
2. Cài đặt môi trường.....	20
3. Cài đặt ứng dụng	21
4. Kết quả thực nghiệm.....	22
5. Deploy ứng dụng lên máy ảo EC2	24
PHẦN KẾT LUẬN.....	32
1. Kết quả đạt được	32
1.1. Nền tảng kiến thức	32
1.2. Vận dụng	32

2.	Ưu điểm	32
3.	Nhược điểm	33
4.	Hướng phát triển	33
TÀI LIỆU THAM KHẢO		34

PHẦN MỞ ĐẦU

1. Tính cấp thiết của đề tài

Ngày nay, vấn đề ngôn ngữ luôn là một vấn đề được quan tâm hàng đầu trong cuộc sống, bởi ngôn ngữ là “sợi chỉ đỏ” để kết nối con người người với con người, thậm chí là con người với công nghệ. Việc để ngôn ngữ trở thành một rào cản trong thời đại công nghệ 4.0 là một điều không đáng có, nhưng mà không phải ai cũng có thể học được một ngôn ngữ nào đó trước khi đi đến một quốc gia và giao tiếp với người bản địa.

Hiểu được vấn đề đó, con người luôn tìm tòi và sáng tạo ra các công cụ dịch thuật góp phần giúp con người kết nối với nhau vượt qua mọi rào cản và xóa đi mọi khoảng cách về ngôn ngữ một các hiệu quả như: Google dịch, Microsoft dịch,...

Khi được học môn Điện toán Đám mây, chúng em có cơ hội được tiếp xúc gần hơn với nền tảng và hệ sinh thái vô cùng mới mẻ của Amazon Web Services (AWS), sau khi tìm hiểu thì chúng em nhận thấy ở AWS có dịch vụ hỗ trợ dịch thuật là Amazon Translate. Vậy nên, với mong muốn làm chủ về ngôn ngữ trong thời đại chuyển đổi số nhóm chúng em đã quyết định thực hiện đề tài: ***“Tìm hiểu Amazon Translate và viết ứng dụng minh họa”***.

2. Đối tượng nghiên cứu

Đối với đề tài này, chúng em tập trung nghiên cứu các đối tượng sau:

- Dịch vụ Amazon Translate
- Dịch vụ Amazon Polly
- Dịch vụ Amazon Textract

3. Phạm vi nghiên cứu

Chúng em sẽ nghiên cứu các vấn đề cơ bản, đặc điểm của Amazon Translate, phương thức hoạt động, chi phí sử dụng và ứng dụng của Amazon Translate vào thực tiễn. So sánh Amazon Translate và Google dịch.

Bên cạnh đó, chúng em còn nghiên cứu thêm về dịch vụ của **Amazon Polly** và **Amazon Textract** để phần nào mở rộng thêm đề tài.

4. Bố cục đề tài

Chương 1: Cơ sở lý thuyết về Amazon Translate.

Chương 2: Xây dựng ứng dụng minh họa.

5. Kết quả dự kiến đạt được

Thông qua việc tìm hiểu đề tài này chúng em dự kiến sẽ hoàn thành những nhiệm vụ như sau:

- Tìm hiểu về tổng quan, cách thức hoạt động, đặc điểm, lợi ích, tính năng của dịch vụ Amazon Translate.

- Xây dựng được một trang web để dịch từ ngữ, văn bản trên nhiều ngôn ngữ sử dụng dịch vụ Amazon Translate.

- Có thể chuyển văn bản thành giọng nói để người dùng giao tiếp trực quan hơn với người khác nhờ tìm hiểu thêm dịch vụ Amazon Polly.

- Tìm hiểu thêm về dịch vụ Amazon Textract để tự động chuyển thành văn bản từ các tài liệu ảnh được đánh máy hoặc viết tay có đuôi png, jpg,...

PHẦN NỘI DUNG

CHƯƠNG 1: CƠ SỞ LÝ THUYẾT VỀ AMAZON TRANSLATE

1. Tổng quan về Amazon Translate

1.1. Khái niệm

Amazon Translate là dịch vụ máy dịch ứng dụng mạng nơ-ron để cung cấp bản dịch ngôn ngữ nhanh chóng, chất lượng cao, giá cả phải chăng và có thể tùy chỉnh. Dịch máy sử dụng mạng nơ-ron là hình thức tự động hóa công việc dịch thuật sử dụng mô hình deep learning để đưa ra bản dịch chính xác hơn và có văn phong tự nhiên hơn các thuật toán dịch thuật theo quy tắc và thống kê truyền thống.

Với Amazon Translate, chúng ta có thể bản địa hóa các nội dung như trang web và ứng dụng cho bộ phận người dùng đa dạng, dễ dàng dịch khối lượng lớn văn bản để phân tích, và cho phép giao tiếp đa ngôn ngữ một cách hiệu quả giữa nhiều người dùng.

Công nghệ sử dụng: **API deep learning dựa trên đám mây.**

Gần đây, Intento đã xếp hạng Amazon Translate là nhà cung cấp dịch máy hàng đầu trong năm 2020 với 14 cặp ngôn ngữ, 16 lĩnh vực công nghiệp và 8 loại nội dung.



1.2. Lợi ích mà Amazon Translate đem lại

- ***Độ chính xác cao và khả năng cải thiện liên tục***

Amazon Translate là dịch vụ dịch máy sử dụng mạng nơ-ron. Công cụ dịch thuật này liên tục cải thiện từ các bộ dữ liệu mới và mở rộng nhằm đưa ra bản dịch chính xác hơn cho nhiều trường hợp sử dụng đa dạng.

- ***Dễ dàng tích hợp vào ứng dụng***

Amazon Translate loại bỏ đi sự phức tạp của việc phải xây dựng các khả năng dịch thuật thời gian thực và dịch thuật hàng loạt trong ứng dụng của chúng ta chỉ bằng một lệnh gọi API đơn giản. Việc này khiến công tác bản địa hóa ứng dụng hoặc trang web, hoặc xử lý dữ liệu đa ngôn ngữ trong quy trình công việc hiện tại trở nên dễ dàng.

- ***Có thể tùy chỉnh***

Với tính năng Custom Terminology và Active Custom Translate, Amazon Translate cho phép chúng ta tùy chỉnh đầu ra máy dịch. Dùng Custom Terminology để xác định cách dịch tên thương hiệu, tên mẫu mã và các thuật ngữ đặc biệt khác. Dùng Active Custom Translation để tạo đầu ra máy dịch tùy chỉnh sao cho phù hợp với các nhu cầu cụ thể mà miền đó đặt ra. Chúng ta không cần xây dựng mô hình dịch tùy chỉnh, có thể cập nhật mô hình bất cứ khi nào chúng ta muốn và chỉ phải trả cho số lượng ký tự cần dịch.

- ***Tiết kiệm chi phí***

Với sức mạnh của dịch máy, Amazon Translate có giá rẻ hơn 1000 lần so với việc sử dụng thông dịch viên chuyên nghiệp để dịch một cách thủ công.

- ***Có khả năng mở rộng***

Dù chỉ là vài từ hay khối lượng văn bản lớn, Amazon Translate đều dễ dàng thay đổi quy mô theo nhu cầu dịch thuật với quá trình dịch nhanh chóng và đáng tin cậy.

- ***Linh hoạt***

Amazon Translate có thể dịch nhiều định dạng nội dung khác nhau, bao gồm văn bản Word, bài thuyết trình Powerpoint và bảng tính Excel.

1.3. Trường hợp sử dụng

- ***Bản địa hóa ngôn ngữ***

Các đội ngũ dịch giả rất khó có thể bắt kịp nội dung động hoặc thời gian thực. Tuy nhiên với Amazon Translate, chúng ta có thể dễ dàng dịch thuật khối lượng nội dung khổng lồ do người dùng tạo theo thời gian thực. Các trang web và ứng dụng có thể tự động sáng tạo nội dung như câu chuyện trên nguồn cấp tin, mô tả

hồ sơ và bình luận bằng ngôn ngữ mà người dùng muốn chỉ bằng một lần nhấp chuột vào nút “dịch”.

- ***Phân tích văn bản***

Với Amazon Translate, chúng ta không còn vướng phải rào cản ngôn ngữ. Có thể vừa nắm bắt được “tình cảm xã hội” của thương hiệu, sản phẩm hoặc dịch vụ của chúng ta, vừa theo dõi các cuộc hội thoại trực tuyến bằng nhiều ngôn ngữ khác nhau. Chúng ta chỉ cần dịch văn bản sang tiếng Anh rồi sử dụng ứng dụng xử lý nội dung tự nhiên (NLP) như Amazon Comprehend để phân tích nội dung theo ngữ cảnh bằng nhiều ngôn ngữ khác nhau.

- ***Giao tiếp***

Amazon Translate có thể cung cấp bản dịch tự động để tạo điều kiện giao tiếp đa ngôn ngữ giữa các người dùng. Bằng cách bổ sung khả năng dịch thuật thời gian thực vào tính năng trò chuyện, email, tổng đài hỗ trợ và ứng dụng nhận yêu cầu, đơn vị hoặc nhân viên nói tiếng Anh có thể giao tiếp với khách hàng bằng nhiều ngôn ngữ khác nhau.

1.4. Các đặc điểm, tính năng của Amazon Translate

- ***Broad Language Coverage***

Amazon Translate hỗ trợ dịch trên 75 ngôn ngữ: Afrikaans, Albanian, Amharic, Arabic, ... Giữa các cặp ngôn ngữ này, dịch vụ hỗ trợ khoảng 5550 tổ hợp dịch khác nhau.

- ***Neural Network-Based***

Amazon Translate sử dụng Deep Learning để tạo ra bản dịch chính xác và trôi chảy hơn so với các mô hình dịch dựa trên quy tắc và thống kê truyền thống. Hệ thống máy dịch nơ-ron được xây dựng dựa trên mạng nơ-ron chứa toàn bộ ngữ cảnh của ngôn ngữ nguồn và các bản dịch mà nó đã tạo ra cho đến nay. Trong khi đó các máy dịch truyền thống chỉ dịch ngữ cảnh của một vài từ trước và sau từ được dịch.

- ***Customized Machine Translation***

Sử dụng Active Custom Translation (ACT), Amazon Translate cho phép chúng ta kiểm soát đầu ra tốt hơn. Chúng ta có thể đưa dữ liệu của mình lên Amazon Translate để tùy chỉnh đầu ra đã dịch cho phù hợp với nhu cầu. Có thể cập nhật dữ liệu của mình thường xuyên nếu cần để bản dịch tùy chỉnh được cập nhật mà không phải trả phí.

- ***Named Entity Translation Customization***

Sử dụng Custom Terminology, Amazon Translate cho phép chúng ta định nghĩa các thuật ngữ hoặc tên tổ chức, vùng miền ... được dịch. Điều này giúp giảm số lượng bản dịch cần được chỉnh sửa bởi các dịch giả chuyên nghiệp, tiết kiệm chi phí và dịch nhanh hơn.

- ***Language Identification***

Amazon Translate có cơ chế xác định ngôn ngữ nguồn khi không được người dùng chỉ định với độ chính xác cao.

- ***Batch and Real-Time Translations***

Amazon Translate là lựa chọn tuyệt vời để thực hiện bản dịch hàng loạt khi có số lượng lớn văn bản sẵn có để dịch và bản dịch thời gian thực khi chúng ta muốn cung cấp bản dịch với nội dung theo yêu cầu.

- ***Secure Machine Translation***

Giao tiếp giữa trang web hoặc ứng dụng của chúng ta tới Amazon Translate được bảo vệ bằng mã hóa SSL. Bất kỳ nội dung nào do Amazon Translate xử lý đều được mã hóa và lưu trữ tại Amazon Region nơi mà chúng ta đang sử dụng dịch vụ. Quản trị viên cũng có thể kiểm soát quyền truy cập vào Amazon Translate thông qua chính sách cấp phép AWS Identity and Access Management (IAM) - đảm bảo rằng thông tin nhạy cảm được giữ an toàn và bảo mật.

- ***Pay-Per-Use***

Với Amazon Translate, chỉ cần trả tiền cho những gì chúng ta sử dụng, giúp dễ dàng tiết kiệm chi phí để mở rộng nhu cầu dịch của mình. Chi phí được tính dựa trên tổng số ký tự được gửi đến API để dịch.

1.5. Một số công cụ sử dụng để liên kết tới Amazon Translate

- ***AWS CLI (AWS Command Line Interface)***: giao diện dòng lệnh (CLI) của AWS là một công cụ mã nguồn mở, thống nhất để quản lý, tương tác với các dịch vụ AWS. Chỉ với một công cụ để tải xuống và cấu hình, chúng ta có thể kiểm soát nhiều dịch vụ AWS bằng một dòng lệnh và tự động hóa chúng thông qua các tập lệnh. AWS CLI cung

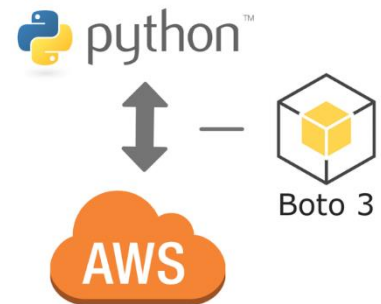


cấp quyền truy cập trực tiếp vào các API công khai của các dịch vụ AWS. Chúng

ta có thể khám phá các khả năng của dịch vụ với AWS CLI và phát triển các tập lệnh shell để quản lý tài nguyên. Ngoài các lệnh cấp thấp tương đương API, một số dịch vụ AWS cung cấp các tùy chỉnh cho AWS CLI. Các tùy chỉnh có thể bao gồm các lệnh cấp cao hơn giúp đơn giản hóa việc sử dụng dịch vụ có API phức tạp.

AWS CLI v2 là phiên bản chính mới nhất của AWS CLI cung cấp một số tính năng mới bao gồm trình cài đặt được cải tiến, các tùy chọn cấu hình mới như AWS Single Sign-On (SSO) và các tính năng tương tác khác nhau. AWS CLI v2 chỉ có sẵn để cài đặt dưới dạng trình cài đặt đi kèm. Mặc dù có thể tìm thấy nó trong trình quản lý gói, nhưng đây là những gói không được hỗ trợ và không chính thức, không được AWS sản xuất hoặc quản lý, chỉ nên cài đặt AWS CLI từ các điểm phân phối AWS chính thức.

• ***Boto 3 - AWS SDK (AWS Software Development Kit) dành cho Python:*** Boto 3 giúp chúng ta dễ dàng tích hợp ứng dụng, thư viện hoặc tập chỉ lệnh với các dịch vụ AWS, bao gồm Amazon S3, Amazon EC2, Amazon DynamoDB và nhiều dịch vụ khác.



Tính năng chính của Boto 3:

- **Resource APIs:** Boto 3 có 2 cấp API riêng biệt. Client APIs (hay còn gọi là “cấp thấp”) cung cấp ánh xạ một-một đến các hoạt động API HTTP ngàm. Resource APIs ẩn các lệnh gọi mạng rõ ràng nhưng thay vào đó sẽ cung cấp các đối tượng và bộ sưu tập tài nguyên để truy cập thuộc tính và thực hiện thao tác.

- **Giao diện đồng nhất và được cập nhật mới nhất:** các giao diện “client” và “resource” của Boto 3 có các lớp được tạo linh hoạt lấy các mô hình JSON mô tả các API AWS làm cơ sở. Việc này cho phép AWS cung cấp các bản cập nhật rất nhanh với độ đồng nhất cao trên tất cả các dịch vụ được hỗ trợ.

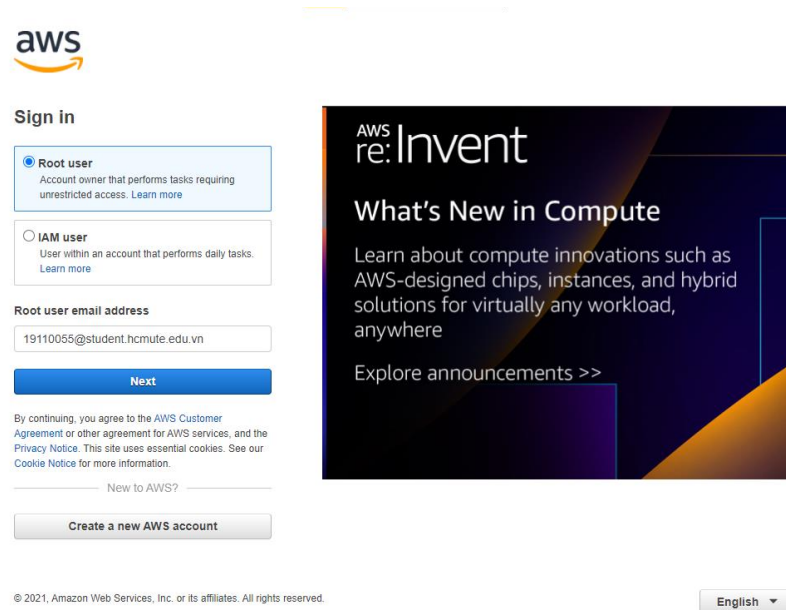
- **Hỗ trợ Python 2 và 3:** Boto 3 được viết từ đầu để có được khả năng hỗ trợ thuần cho các phiên bản Python 2.7+ và 3.4+.

- **Trình chờ (waiters):** Boto 3 được cung cấp kèm theo các “trình chờ”, có chức năng tự động thăm dò cho đến khi đạt được thay đổi trạng thái quy định trước trong tài nguyên AWS. Boto 3 có trình chờ cho cả Client và Resource APIs.

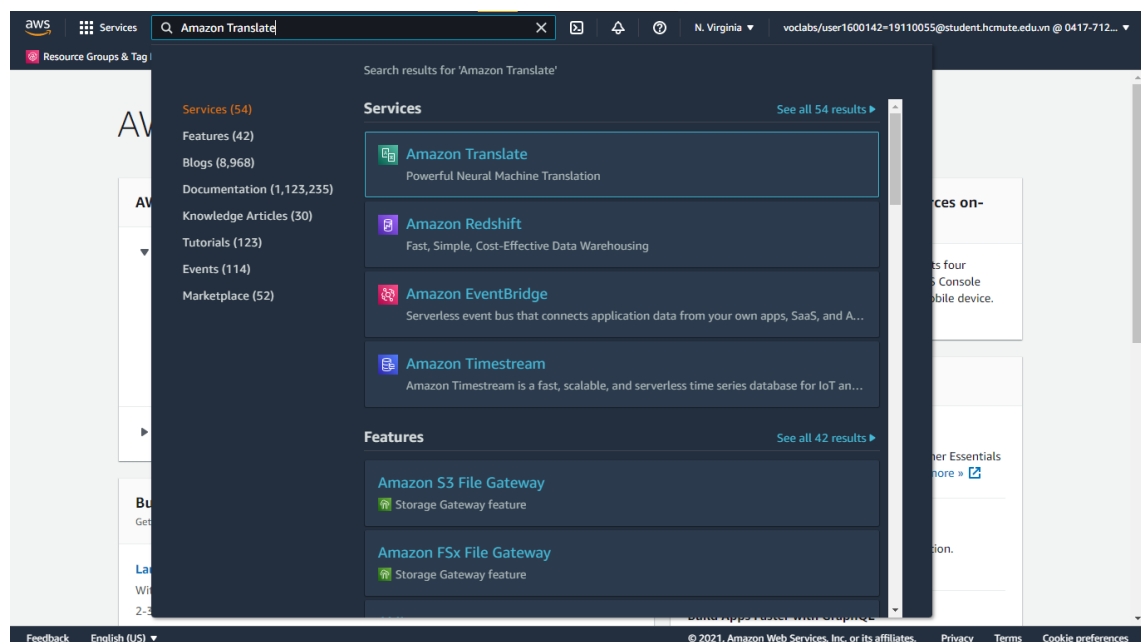
- **Nhiều tính năng cấp cao dành riêng cho dịch vụ:** Boto 3 đi kèm nhiều tính năng dành riêng cho dịch vụ như Automatic multi-part transfers dành cho Amazon S3 và Simplified query conditions cho Amazon DynamoDB.

2. Các bước sử dụng dịch vụ Amazon Translate

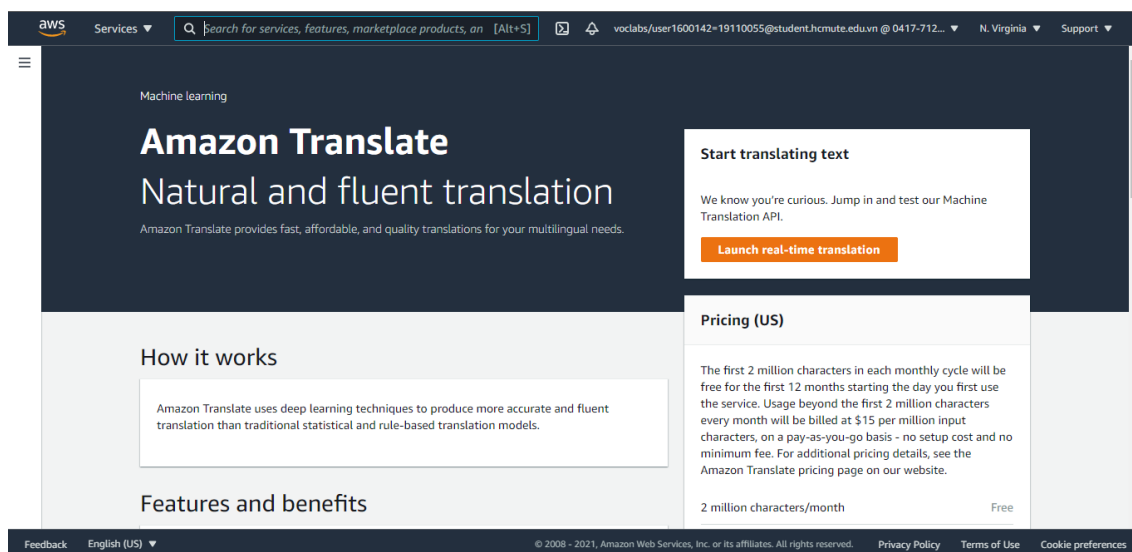
- **Bước 1:** Đăng nhập vào tài khoản AWS.



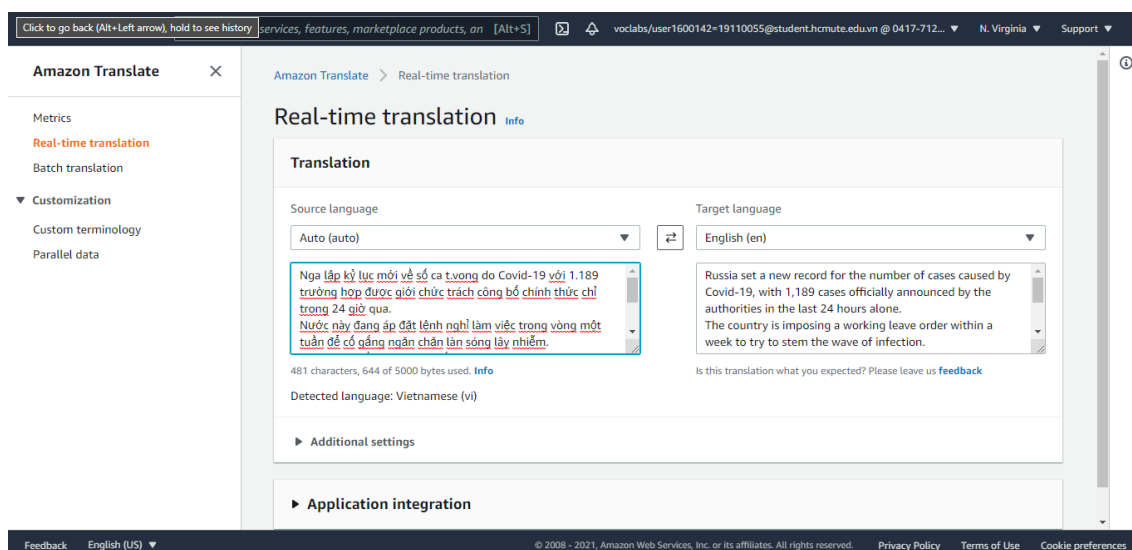
- **Bước 2:** Tại ô tìm kiếm nhập từ khóa “*Amazon Translate*” để truy cập vào dịch vụ dịch thuật của Amazon.



- Bước 3: Để bắt đầu sử dụng dịch vụ chọn **“Launch real-time translation”**.



- Bước 4: Gõ hoặc dán văn bản cần dịch vào, Amazon Translate sẽ tự động phát hiện ngôn ngữ đầu vào và bạn có thể tùy chỉnh kết quả đầu ra.



3. So sánh và phân tích dịch vụ Amazon Translate và Google Translate

Có thể nói được ngôn ngữ là chìa khóa của sự kết nối thành công, nhưng ... thật khó để chọn công cụ dịch đáng tin cậy nhất khi tất cả đều tự nhận là tốt nhất. Đi sâu vào so sánh **Google Translate** và **Amazon Translate** sẽ giúp bạn tìm ra công cụ dịch hoàn hảo cho nhu cầu của mình.

Phân tích	Google Translate	Amazon Translate
Logo nhận diện thương hiệu		

Khái niệm	Google Translate là một dịch vụ dịch máy miễn phí do Google phát triển để giúp người dùng internet hàng ngày dịch văn bản, tài liệu nhỏ hơn và trang web từ ngôn ngữ này sang ngôn ngữ khác. Nó cung cấp một giao diện trang web, một ứng dụng di động cho Android và IOS, và một giao diện lập trình ứng dụng cho các nhà phát triển. Nó cho phép người dùng internet dịch sang 109 ngôn ngữ khác nhau.	Amazon Translate là một dịch vụ dịch thuật do Amazon.com, Inc. cung cấp , sử dụng công nghệ máy học tiên tiến để cung cấp bản dịch chất lượng cao theo yêu cầu. Dịch vụ Amazon Translate có thể được sử dụng để dịch các tài liệu văn bản trong DOCX, PPTX, XSLX và các tệp tương tự. Nó cũng có thể được sử dụng để dịch các ứng dụng di động và máy tính sang nhiều ngôn ngữ.
Dùng thử miễn phí	✓	✓
Thanh toán	Miễn phí	Sử dụng bao nhiêu trả bấy nhiêu dựa trên số lượng ký tự văn bản mà bạn đã xử lý. (Dịch tối đa 2 triệu ký tự hàng tháng - miễn phí trong 12 tháng đầu tiên)
Số loại tệp được hỗ trợ	11 (.doc, .docx, .odf, .pdf, .ppt, .pptx, .ps, .rtf, .txt, .xls, .xlsx)	2 (.docx và .pdf nhưng phải kết hợp thêm với các dịch vụ khác như: Amazon S3, Amazon Textract. AWS Lambda)
Giới hạn tải lên tài liệu	150MB	20MB
Tích hợp dịch máy	✓	✓
Số lượng ngôn ngữ được hỗ trợ	108 ngôn ngữ	71 ngôn ngữ

Công cụ trực tuyến	✓	✓
Bộ nhớ dữ liệu đã dịch	✗	✓
Lưu trữ trên đám mây	✗	✓
Chỉnh sửa theo ngữ cảnh	✗	✓
Tự động phát hiện ngôn ngữ	✓	✓

4. Tìm hiểu một số dịch vụ kết hợp vào đề tài

4.1. Dịch vụ Amazon Polly

Amazon Polly là dịch vụ chuyên đổi văn bản thành giọng nói chân thực, cho phép bạn tạo các ứng dụng có thể nói chuyện và phát triển những thể loại sản phẩm được trang bị khả năng nói hoàn toàn mới.

Polly là dịch vụ Text-to-Speech (TTS) (chuyển văn bản thành lời nói) sử dụng công nghệ deep learning tiên tiến để tổng hợp thành lời nói tự nhiên của con người.



Với hàng chục giọng nói chân thực theo **31 ngôn ngữ**, bạn có thể xây dựng ứng dụng có giọng nói hoạt động ở nhiều nước khác nhau. Tuy nhiên, Amazon Polly vẫn chưa hỗ trợ ngôn ngữ Việt Nam.

Ứng dụng vào đề tài nhằm giúp chuyển đổi văn bản thành giọng nói và để đọc những ngôn ngữ đặc trưng theo từng quốc gia để trang web dịch thuật trở nên tiện lợi hơn.

4.2. Dịch vụ Amazon Textract

Amazon Textract là một dịch vụ máy học (ML) sử dụng tính năng nhận diện ký tự quang học (OCR) để tự động trích xuất văn bản, chữ viết tay và dữ liệu từ các tài liệu được quét như PNG.

Dịch vụ có thể chuyển đổi hình ảnh chữ viết tay và chữ đánh máy thành văn bản với số lượng



ngôn ngữ hạn chế như: Tiếng Anh, Pháp, Đức, Ý, Bồ Đào Nha và Tây Ban Nha.

Ứng dụng vào đề tài nhằm giúp khâu nhập văn bản đầu vào được nhanh hơn, tiện lợi hơn, đa dạng hóa các nguồn đầu vào bằng hình ảnh và văn bản.

CHƯƠNG 2: XÂY DỰNG ỨNG DỤNG MINH HỌA

1. Giới thiệu sản phẩm đề tài

1.1. Giới thiệu

Trong cuộc sống hiện đại ngày nay, việc dịch thuật các ngôn ngữ để con người có thể dễ dàng thông hiểu và giao tiếp với nhau là một điều hết sức quan trọng và cần thiết. Đặc biệt là đối với các bạn sinh viên thường xuyên tiếp cận với các tài liệu học tập, nghiên cứu, tham khảo với nhiều ngôn ngữ thì công cụ dịch càng trở nên quan trọng hơn.

Hiểu được điều đó, ứng dụng “DTL Translatr App” ra đời với mục đích xóa tan mọi rào cản về ngôn ngữ, giúp người dùng có thể dịch, nghe, chuyển hình ảnh bằng văn bản một cách dựa trên các dịch vụ của AWS như: Amazon Translate, Amazon Polly, Amazon Textract,...

1.2. Nền tảng công nghệ

Ứng dụng được viết dựa trên các công nghệ như:

Công nghệ	Mô tả
Python 3.8	Ngôn ngữ chính được sử dụng là Python 3.8 bằng công cụ PyCharm 2021.
Flask Framework	Cung cấp cho ta tất cả các công cụ, thư viện và công nghệ cho phép bạn xây dựng một ứng dụng web.
HTML/CSS	HTML, HyperText Markup Language, cung cấp cấu trúc nội dung và ý nghĩa bằng cách xác định nội dung đó, ví dụ như tiêu đề, đoạn văn hoặc hình ảnh. CSS, hay Cascading Style Sheets, là ngôn ngữ trình bày được dùng để tạo kiểu cho sự xuất hiện của nội dung sử dụng, ví dụ như phông chữ hoặc màu sắc.
Bootstrap 4	Bootstrap 4 (viết tắt là BS4) là phiên bản mới của Bootstrap, là framework HTML, CSS và JavaScript phổ biến nhất để thiết kế web đáp ứng
Thư viện boto3	Giúp ta dễ dàng tích hợp ứng dụng, thư viện hoặc tập chỉ lệnh với các dịch vụ AWS, bao gồm Amazon S3, Amazon EC2, Amazon DynamoDB và nhiều dịch vụ khác.
AWS CLI (AWS Command Line Interface)	Giao diện dòng lệnh (CLI) của AWS là một công cụ mã nguồn mở, thống nhất để quản lý, tương tác với các dịch vụ AWS.

Amazon Translate	Giúp dịch các văn bản sang nhiều ngôn ngữ khác nhau.
Amazon Textract	Chuyển hình ảnh chữ viết tay hoặc đánh máy thành văn bản.
Amazon Polly	Chuyển văn bản thành giọng nói với nhiều giọng đọc và ngôn ngữ khác nhau.

1.3. Chức năng

- Dịch văn bản từ ngôn ngữ này sang ngôn ngữ khác (**với 71 ngôn ngữ được hỗ trợ**).

- Chuyển văn bản thành giọng nói với nhiều ngôn ngữ (**với 11 ngôn ngữ được hỗ trợ**).

- Chuyển hình ảnh file .png, .jpeg chữ viết tay hoặc đánh máy sang dạng văn bản để tiến hành việc dịch thuật (với số lượng ngôn ngữ truy xuất hạn chế như: Tiếng Anh, Pháp, Đức, Ý, Bồ Đào Nha và Tây Ban Nha).

2. Cài đặt môi trường

- Cài đặt thư viện giao diện dòng lệnh AWS CLI MSI bản dành cho Windows (64-bit): <https://awscli.amazonaws.com/AWSCLIV2.msi>

```
C:\WINDOWS\system32\cmd.exe
Microsoft Windows [Version 10.0.19042.1415]
(c) Microsoft Corporation. All rights reserved.

C:\Users\USER>aws --version
aws-cli/2.4.7 Python/3.8.8 Windows/10 exe/AMD64 prompt/off
```

- Cài đặt cấu hình người dùng với các *AWS Access Key ID*, *AWS Secret Access Key*, *Default region name*, *Default output format* tương ứng.

```
C:\Users\USER>aws configure
AWS Access Key ID [*****A66J]: ASIAQT00B77BEUMJYDS7
AWS Secret Access Key [*****Bnep]: KEOyVPv582v+QcExM3y0g6fH7UKOY89FHH3wi09a
Default region name [us-east-1]: us-east-1
Default output format [json]: json
```

- Cài đặt thư viện boto3 bằng lệnh `pip install boto3`.

- Cài đặt thư viện flask framework bằng lệnh `pip install flask`.

- Cài đặt thư viện Bootstrap 4 framework HTML, CSS và JavaScript:

```
<link rel="stylesheet"
href="https://maxcdn.bootstrapcdn.com/bootstrap/4.0.0
/css/bootstrap.min.css" integrity="sha384-
```

```
Gn5384xqQ1aoWXA+058RXPxPg6fy4IWvTNh0E263XmFcJlSAwiGgF  
AW/dAiS6JXm" crossorigin="anonymous">
```

3. Cài đặt ứng dụng

- **Bước 1:** Clone dự án về máy

<https://github.com/dieptanluangithub/NHOM27-TRANSLATE-APP.git>

- **Bước 2:** Tạo tài khoản AWS

- **Bước 3:** Ghi nhớ các thông tin cấu hình như:

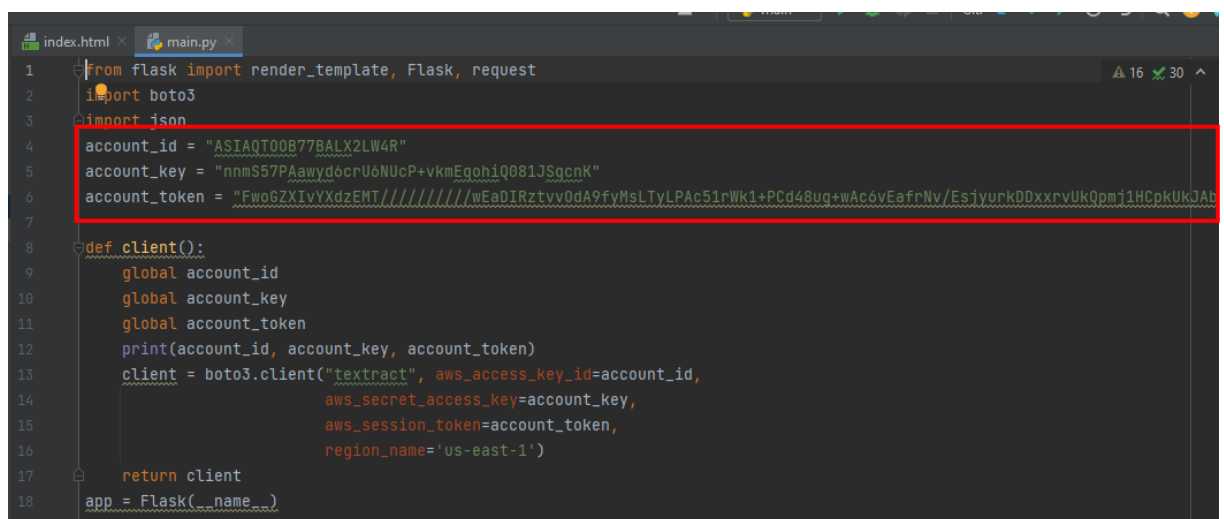
aws_access_key_id

aws_secret_access_key

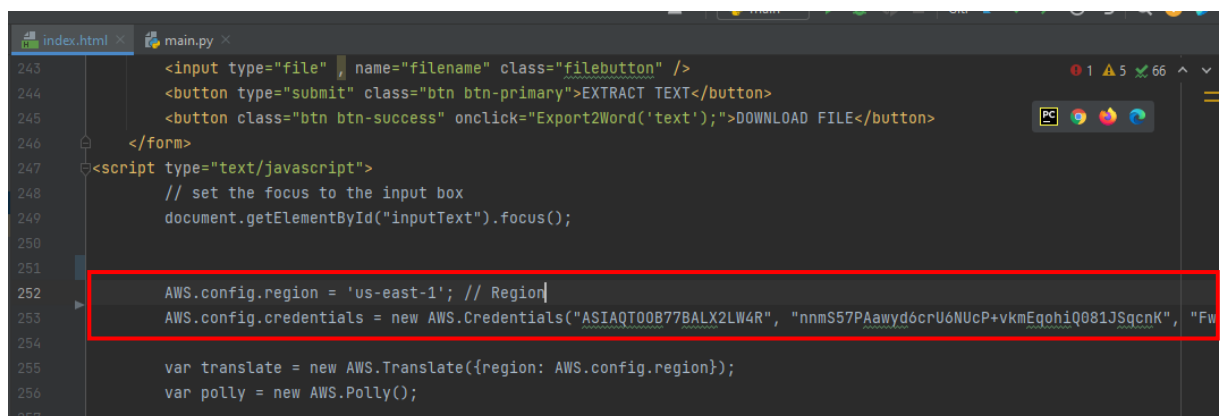
aws_session_token

aws_region

- **Bước 4:** Mở file main.py và file index.html để cấu hình thông tin tài khoản AWS của bạn

A screenshot of a code editor showing the main.py file. The code imports Flask, boto3, and json. It defines three global variables: account_id, account_key, and account_token, each assigned a long alphanumeric string. A def client() function is defined, which prints these values and creates a boto3 client for 'texttranscribe' in the 'us-east-1' region, passing the account credentials. The Flask app is initialized with __name__ as the template folder.

```
1 from flask import render_template, Flask, request
2 import boto3
3 import json
4 account_id = "ASIAQT00B77BALX2LW4R"
5 account_key = "nnmS57PAawydc6rU6NUcP+vkmEgohiQ081JSqcnK"
6 account_token = "Fwo6ZXivYXdzEMT////////wEa0IRztvv0dA9fyMsLTyLPac51rWk1+PCd48ug+wAc0vEafrNv/EsjyurkDDxxrvUkQpmj1HCpkUkJAAb"
7
8 def client():
9     global account_id
10    global account_key
11    global account_token
12    print(account_id, account_key, account_token)
13    client = boto3.client("texttranscribe", aws_access_key_id=account_id,
14                          aws_secret_access_key=account_key,
15                          aws_session_token=account_token,
16                          region_name='us-east-1')
17    return client
18 app = Flask(__name__)
```

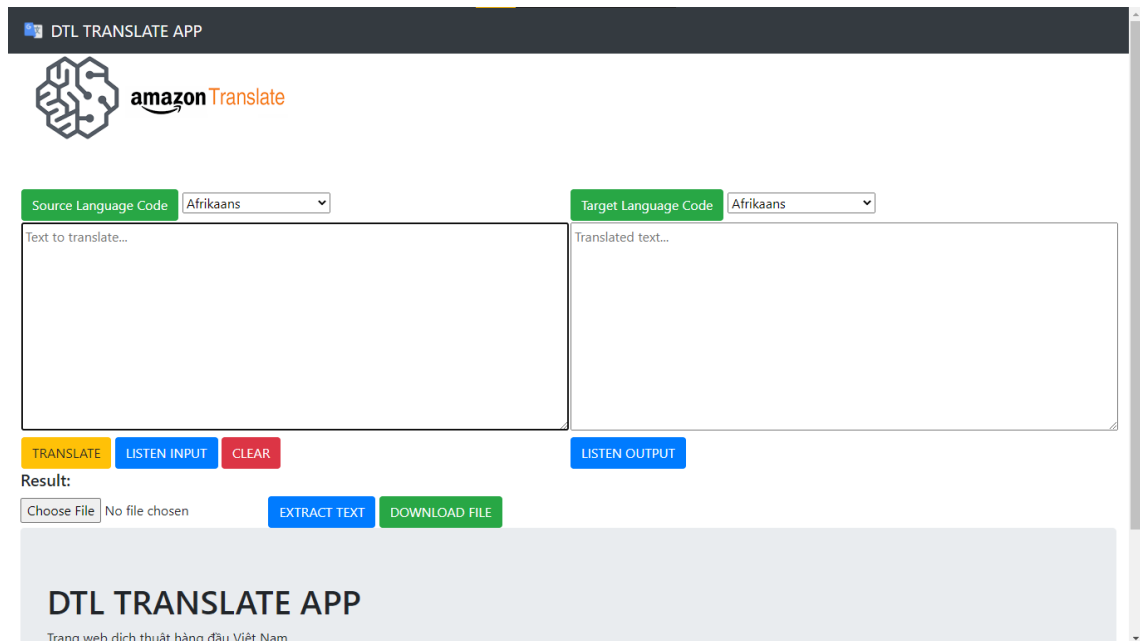
A screenshot of a code editor showing the index.html file. It contains HTML for a file upload form with a submit button and a success button. A JavaScript script sets focus to the input box. At the bottom, there is a red box highlighting the configuration of the AWS SDK, including the region and credentials.

```
243 <input type="file" name="filename" class="filebutton" />
244 <button type="submit" class="btn btn-primary">EXTRACT TEXT</button>
245 <button class="btn btn-success" onclick="Export2Word('text');">DOWNLOAD FILE</button>
246 </form>
247 <script type="text/javascript">
248 // set the focus to the input box
249 document.getElementById("inputText").focus();
250
251
252 AWS.config.region = 'us-east-1'; // Region
253 AWS.config.credentials = new AWS.Credentials("ASIAQT00B77BALX2LW4R", "nnmS57PAawydc6rU6NUcP+vkmEgohiQ081JSqcnK", "Fw
254
255 var translate = new AWS.Translate({region: AWS.config.region});
256 var polly = new AWS.Polly();
257
```


- **Bước 5:** Chạy file main.py để xem kết quả.

4. Kết quả thực nghiệm

- Giao diện trang web



DTL TRANSLATE APP

 amazonTranslate

Source Language Code: Afrikaans

Target Language Code: Afrikaans

Text to translate...

Translated text...

TRANSLATE LISTEN INPUT CLEAR LISTEN OUTPUT

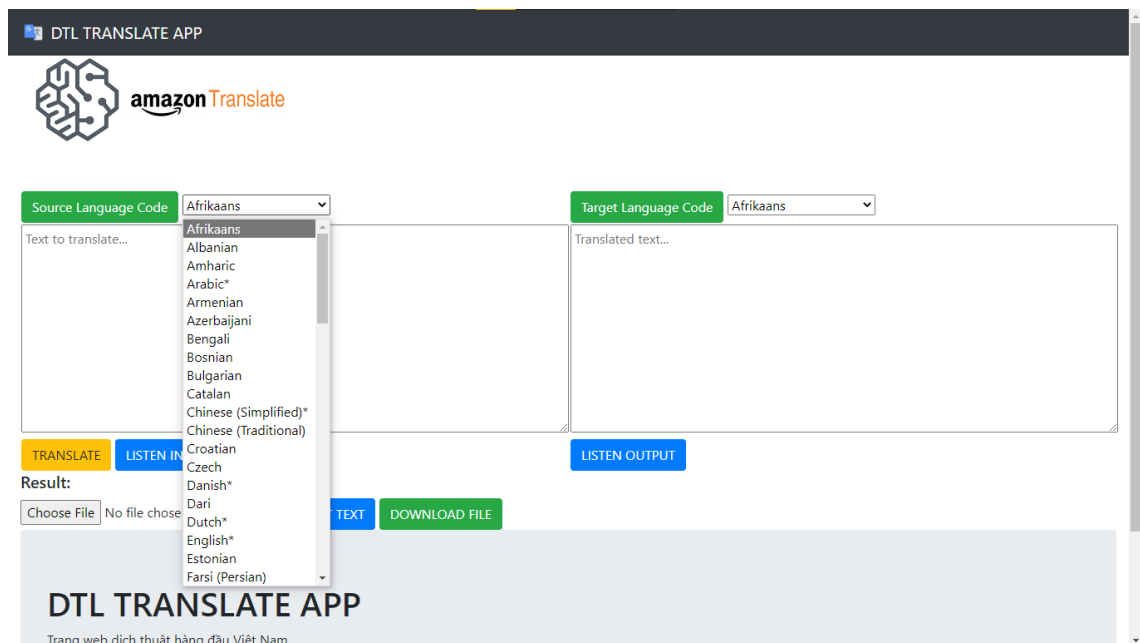
Result:

Choose File No file chosen EXTRACT TEXT DOWNLOAD FILE


DTL TRANSLATE APP

Trang web dịch thuật hàng đầu Việt Nam

- Có thể chọn nhiều ngôn ngữ khác nhau



DTL TRANSLATE APP

 amazonTranslate

Source Language Code: Afrikaans

Target Language Code: Afrikaans

Text to translate...

Translated text...

TRANSLATE LISTEN INPUT CLEAR LISTEN OUTPUT

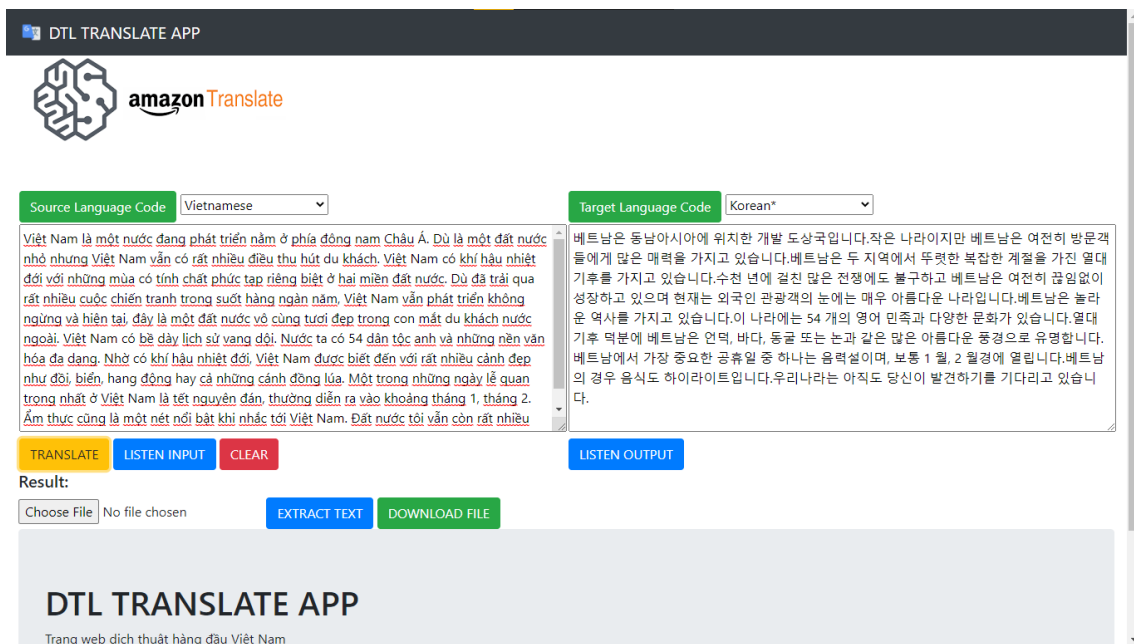
Result:

Choose File No file chosen EXTRACT TEXT DOWNLOAD FILE

DTL TRANSLATE APP

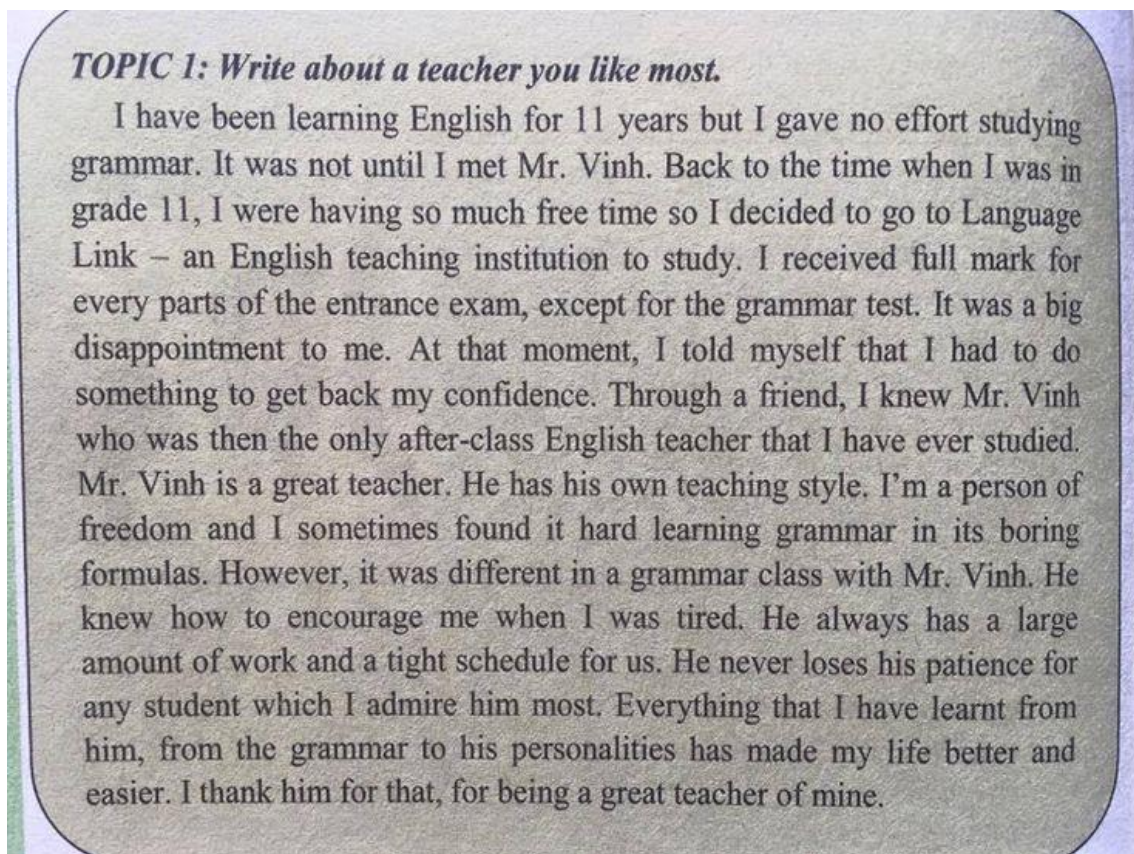
Trang web dịch thuật hàng đầu Việt Nam

- Dịch ngôn ngữ và chuyển văn bản thành giọng nói.




- Chuyển file hình ảnh thành văn bản.

+ Đầu vào:



+ Đầu ra:

DTL TRANSLATE APP



Source Language Code: Afrikaans Target Language Code: Afrikaans

Text to translate... Translated text...

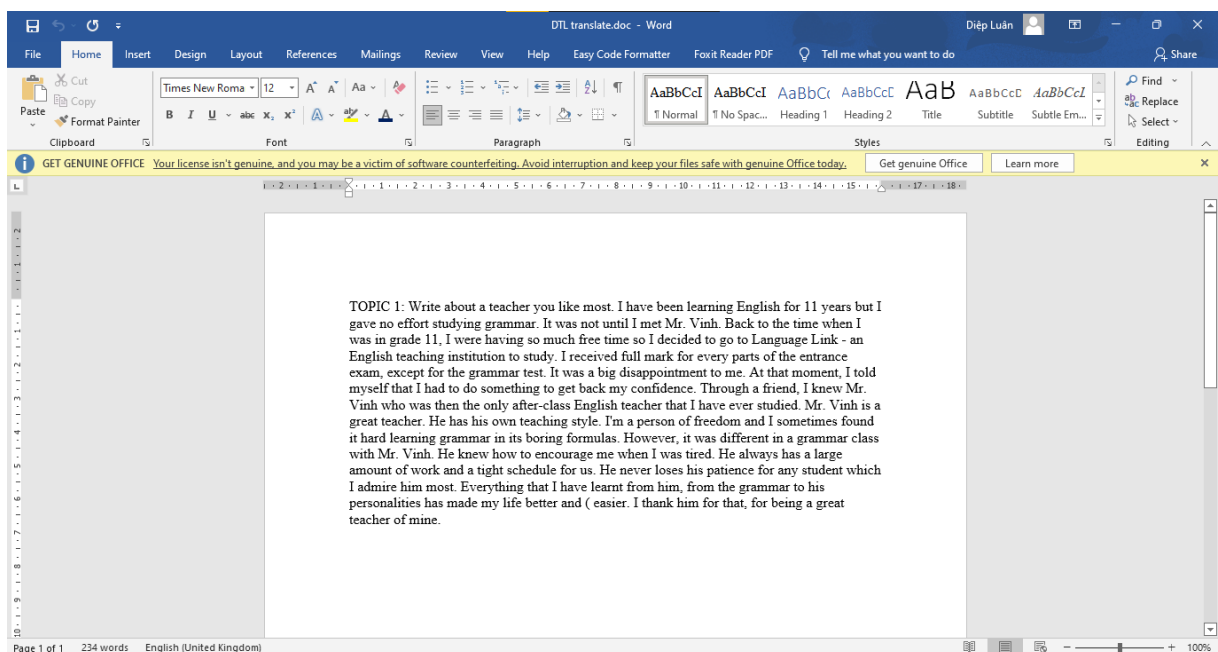
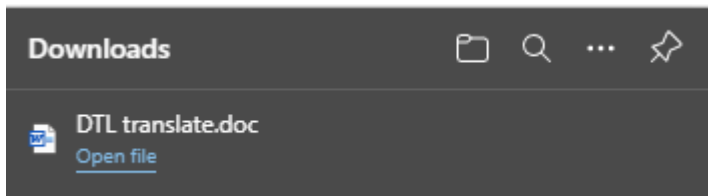
TRANSLATE LISTEN INPUT CLEAR LISTEN OUTPUT

Result:

TOPIC 1: Write about a teacher you like most. I have been learning English for 11 years but I gave no effort studying grammar. It was not until I met Mr. Vinh. Back to the time when I was in grade 11, I were having so much free time so I decided to go to Language Link - an English teaching institution to study. I received full mark for every parts of the entrance exam, except for the grammar test. It was a big disappointment to me. At that moment, I told myself that I had to do something to get back my confidence. Through a friend, I knew Mr. Vinh who was then the only after-class English teacher that I have ever studied. Mr. Vinh is a great teacher. He has his own teaching style. I'm a person of freedom and I sometimes found it hard learning grammar in its boring formulas. However, it was different in a grammar class with Mr. Vinh. He knew how to encourage me when I was tired. He always has a large amount of work and a tight schedule for us. He never loses his patience for any student which I admire him most. Everything that I have learnt from him, from the grammar to his personalities has made my life better and (easier. I thank him for that, for being a great teacher of mine.

Choose File No file chosen EXTRACT TEXT DOWNLOAD FILE

+ Download file vừa chuyển từ ảnh sang văn bản



5. Deploy ứng dụng lên máy ảo EC2

- **Bước 1:** Đăng nhập vào tài khoản AWS.

aws

Sign in

☒ **Root user**
Account owner that performs tasks requiring unrestricted access. [Learn more](#)

☐ **IAM user**
User within an account that performs daily tasks. [Learn more](#)

Root user email address

19110055@student.hcmute.edu.vn

Next

By continuing, you agree to the [AWS Customer Agreement](#) or other agreement for AWS services, and the [Privacy Notice](#). This site uses essential cookies. See our [Cookie Notice](#) for more information.

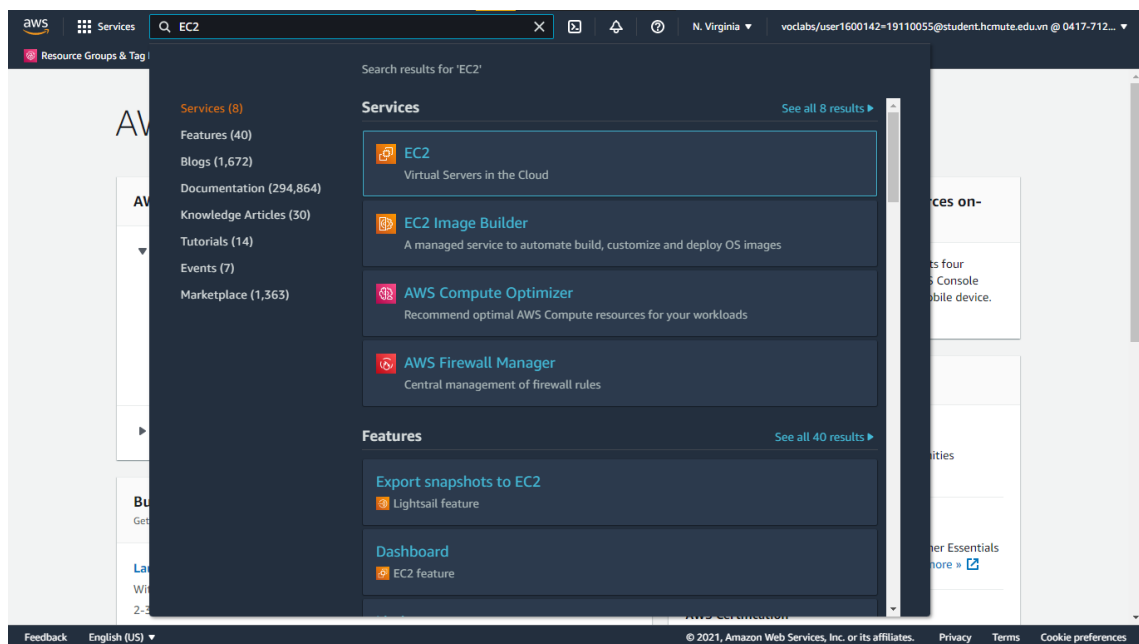
[New to AWS?](#)

Create a new AWS account

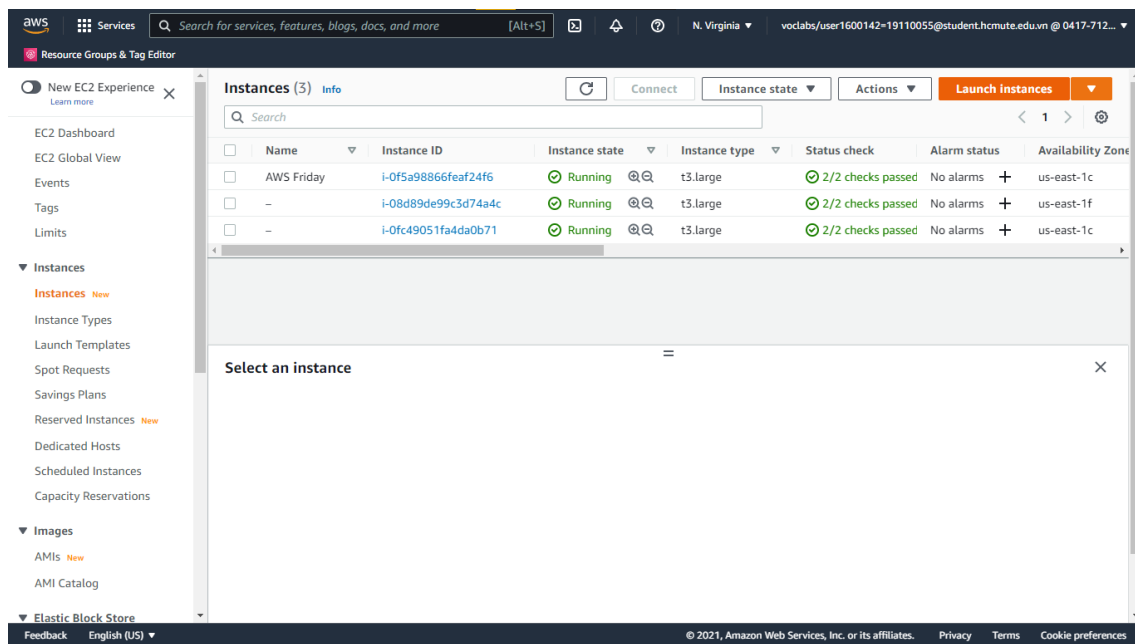
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

English

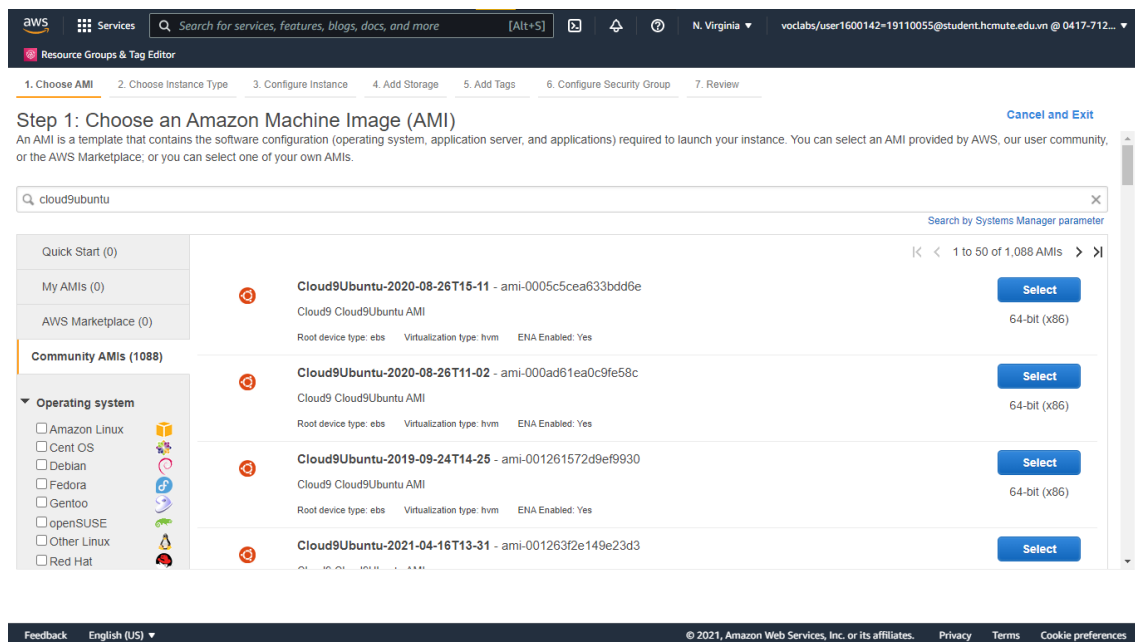
- **Bước 2:** Tại ô tìm kiếm nhập từ khóa “**EC2**” để truy cập vào dịch vụ máy ảo của Amazon.



- **Bước 3:** Tạo một máy ảo mới bằng cách chọn “Launch instances”



- Bước 4: Tại ô tìm kiếm nhập từ khóa “Cloud9ubuntu” và chọn “Select”



- Bước 4: Chọn loại phiên bản “t3.large” với bộ nhớ là 8GB

aws Services Search for services, features, blogs, docs, and more [Alt+S] N. Virginia voclabs/user1600142~19110055@student.hcmute.edu.vn @ 0417-712...

Resource Groups & Tag Editor

1. Choose AMI 2. Choose Instance Type 3. Configure Instance 4. Add Storage 5. Add Tags 6. Configure Security Group 7. Review

Step 2: Choose an Instance Type

Instance Type	Instance Size	Instance Type	VCpus	Memory (GiB)	Storage	Network	Price per Hour (USD)	On-Demand Price per Hour (USD)
<input type="checkbox"/>	t2	t2.small	1	2	EBS only	-	Low to Moderate	Yes
<input type="checkbox"/>	t2	t2.medium	2	4	EBS only	-	Low to Moderate	Yes
<input type="checkbox"/>	t2	t2.large	2	8	EBS only	-	Low to Moderate	Yes
<input type="checkbox"/>	t2	t2.xlarge	4	16	EBS only	-	Moderate	Yes
<input type="checkbox"/>	t2	t2.2xlarge	8	32	EBS only	-	Moderate	Yes
<input type="checkbox"/>	t3	t3.nano	2	0.5	EBS only	Yes	Up to 5 Gigabit	Yes
<input type="checkbox"/>	t3	t3.micro	2	1	EBS only	Yes	Up to 5 Gigabit	Yes
<input type="checkbox"/>	t3	t3.small	2	2	EBS only	Yes	Up to 5 Gigabit	Yes
<input type="checkbox"/>	t3	t3.medium	2	4	EBS only	Yes	Up to 5 Gigabit	Yes
<input checked="" type="checkbox"/>	t3	t3.large	2	8	EBS only	Yes	Up to 5 Gigabit	Yes
<input type="checkbox"/>	t3	t3.xlarge	4	16	EBS only	Yes	Up to 5 Gigabit	Yes
<input type="checkbox"/>	t3	t3.2xlarge	8	32	EBS only	Yes	Up to 5 Gigabit	Yes
<input type="checkbox"/>	t3a	t3a.nano	2	0.5	EBS only	Yes	Up to 5 Gigabit	Yes

Cancel Previous **Review and Launch** Next: Configure Instance Details

Feedback English (US) © 2021, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

- Bước 5: Cấu hình một số chi tiết phù hợp với nhu cầu

aws Services Search for services, features, blogs, docs, and more [Alt+S] N. Virginia voclabs/user1600142~19110055@student.hcmute.edu.vn @ 0417-712...

Resource Groups & Tag Editor

1. Choose AMI 2. Choose Instance Type 3. Configure Instance 4. Add Storage 5. Add Tags 6. Configure Security Group 7. Review

Step 3: Configure Instance Details

Additional charges will apply for selected tenancy.

Elastic Inference ☐ Add an Elastic Inference accelerator
Additional charges apply.

Credit specification ☒ Unlimited
Additional charges may apply.

File systems

▼ Advanced Details

Enclave ☐ Enable

Metadata accessible ☒ Enabled

Metadata version V1 and V2 (token optional)

Metadata token response hop limit 1

User data ☒ As text ☐ As file ☐ Input is already base64 encoded

```
#!/bin/bash
# Install Apache Web Server and PHP
yum install -y httpd mysql php
# Download Lab files
wget https://aws-tc-largeobjects.s3.us-west-2.amazonaws.com/CUR-TF-100-ACCLFO-2/2-lab2-vpc/s3/lab-app.zip
```

Cancel Previous **Review and Launch** Next: Add Storage

Feedback English (US) © 2021, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

```
#!/bin/bash
```

```
# Install Apache Web Server and PHP
```

```
yum install -y httpd mysql php
```

```
# Download Lab files
```

```
wget https://aws-tc-largeobjects.s3.us-west-2.amazonaws.com/CUR-TF-100-ACCLFO-2/2-lab2-vpc/s3/lab-app.zip
```

```
unzip lab-app.zip -d /var/www/html/
```

```
# Turn on web server
```

```
chkconfig httpd on
```

```
service httpd start
```

- **Bước 6: Thêm bộ nhớ**

aws

Services

Search for services, features, blogs, docs, and more

[Alt+S]

N. Virginia

voclabs/user1600142~19110055@student.hcmute.edu.vn @ 0417-712...

Resource Groups & Tag Editor

1. Choose AMI

2. Choose Instance Type

3. Configure Instance

4. Add Storage

5. Add Tags

6. Configure Security Group

7. Review

Step 4: Add Storage

Your instance will be launched with the following storage device settings. You can attach additional EBS volumes and instance store volumes to your instance, or edit the settings of the root volume. You can also attach additional EBS volumes after launching an instance, but not instance store volumes. [Learn more](#) about storage options in Amazon EC2.

Volume Type	Device	Snapshot	Size (GiB)	Volume Type	IOPS	Throughput (MB/s)	Delete on Termination	Encryption
Root	/dev/sda1	snap-0668359b73426edcb	45	General Purpose SSD (gp2)	135 / 3000	N/A	<input checked="" type="checkbox"/>	Not Encrypt

Add New Volume

Free tier eligible customers can get up to 30 GB of EBS General Purpose (SSD) or Magnetic storage. [Learn more](#) about free usage tier eligibility and usage restrictions.

Shared file systems

You currently don't have any file systems on this instance. Select "Add file system" button below to add a file system.

Add file system

Cancel

Previous

Review and Launch

Next: Add Tags

https://aws.amazon.com/free/

© 2021, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

- **Bước 7: Thêm thẻ**

aws

Services

Search for services, features, blogs, docs, and more

[Alt+S]

N. Virginia

voclabs/user1600142~19110055@student.hcmute.edu.vn @ 0417-712...

Resource Groups & Tag Editor

1. Choose AMI

2. Choose Instance Type

3. Configure Instance

4. Add Storage

5. Add Tags

6. Configure Security Group

7. Review

Step 5: Add Tags

A tag consists of a case-sensitive key-value pair. For example, you could define a tag with key = Name and value = Webserver. A copy of a tag can be applied to volumes, instances or both. Tags will be applied to all instances and volumes. [Learn more](#) about tagging your Amazon EC2 resources.

Key	Value	Instances	Volumes	Network Interfaces
name	translateapp	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

Add another tag

(Up to 50 tags maximum)

Cancel

Previous

Review and Launch

Next: Configure Security Group

Feedback English (US)

© 2021, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

- **Bước 8: Cấu hình nhóm bảo mật như Port, địa chỉ truy cập đến trang web**

aws

Services

Search for services, features, blogs, docs, and more

[Alt+S]

N. Virginia

voclabs/user1600142~19110055@student.hcmute.edu.vn @ 0417-712...

Resource Groups & Tag Editor

1. Choose AMI

2. Choose Instance Type

3. Configure Instance

4. Add Storage

5. Add Tags

6. Configure Security Group

7. Review

Step 6: Configure Security Group

A security group is a set of firewall rules that control the traffic for your instance. On this page, you can add rules to allow specific traffic to reach your instance. For example, if you want to set up a web server and allow Internet traffic to reach your instance, add rules that allow unrestricted access to the HTTP and HTTPS ports. You can create a new security group or select from an existing one below. [Learn more](#) about Amazon EC2 security groups.

Assign a security group: ☒ Create a **new** security group
☐ Select an **existing** security group

Security group name:

Description:

Type	Protocol	Port Range	Source	Description
SSH	TCP	22	Custom 0.0.0.0/0	e.g. SSH for Admin Desktop
Custom TCP	TCP	5000	Anywhere 0.0.0.0, ::/0	e.g. SSH for Admin Desktop

[Add Rule](#)

Warning

Rules with source of 0.0.0.0/0 allow all IP addresses to access your instance. We recommend setting security group rules to allow access from known IP addresses only.

[Cancel](#) [Previous](#) [Review and Launch](#)

Feedback

English (US)

© 2021, Amazon Web Services, Inc. or its affiliates.

Privacy

Terms

Cookie preferences

- Bước 9: Xem lại cấu hình máy ảo vừa tạo

aws

Services

Search for services, features, blogs, docs, and more

[Alt+S]

N. Virginia

voclabs/user1600142~19110055@student.hcmute.edu.vn @ 0417-712...

Resource Groups & Tag Editor

1. Choose AMI

2. Choose Instance Type

3. Configure Instance

4. Add Storage

5. Add Tags

6. Configure Security Group

7. Review

Step 7: Review Instance Launch

Please review your instance launch details. You can go back to edit changes for each section. Click **Launch** to assign a key pair to your instance and complete the launch process.

Improve your instances' security. Your security group, launch-wizard-3, is open to the world.

Your instances may be accessible from any IP address. We recommend that you update your security group rules to allow access from known IP addresses only. You can also open additional ports in your security group to facilitate access to the application or service you're running, e.g., HTTP (80) for web servers. [Edit security groups](#)

Your instance configuration is not eligible for the free usage tier

To launch an instance that's eligible for the free usage tier, check your AMI selection, instance type, configuration options, or storage devices. Learn more about [free usage tier](#) eligibility and usage restrictions.

[Don't show me this again](#)

AMI Details

Cloud9Ubuntu-2020-08-26T15-11 - ami-0005c5cea633bdd6e

Cloud9 Cloud9Ubuntu AMI

Root Device Type: ebs Virtualization type: hvm

[Edit AMI](#)

Instance Type

Instance Type	ECUs	vCPUs	Memory (GiB)	Instance Storage (GB)	EBS-Optimized Available	Network Performance
t3.large	-	2	8	EBS only	Yes	Up to 5 Gigabit

[Edit instance type](#)

[Cancel](#) [Previous](#) [Launch](#)

Feedback

English (US)

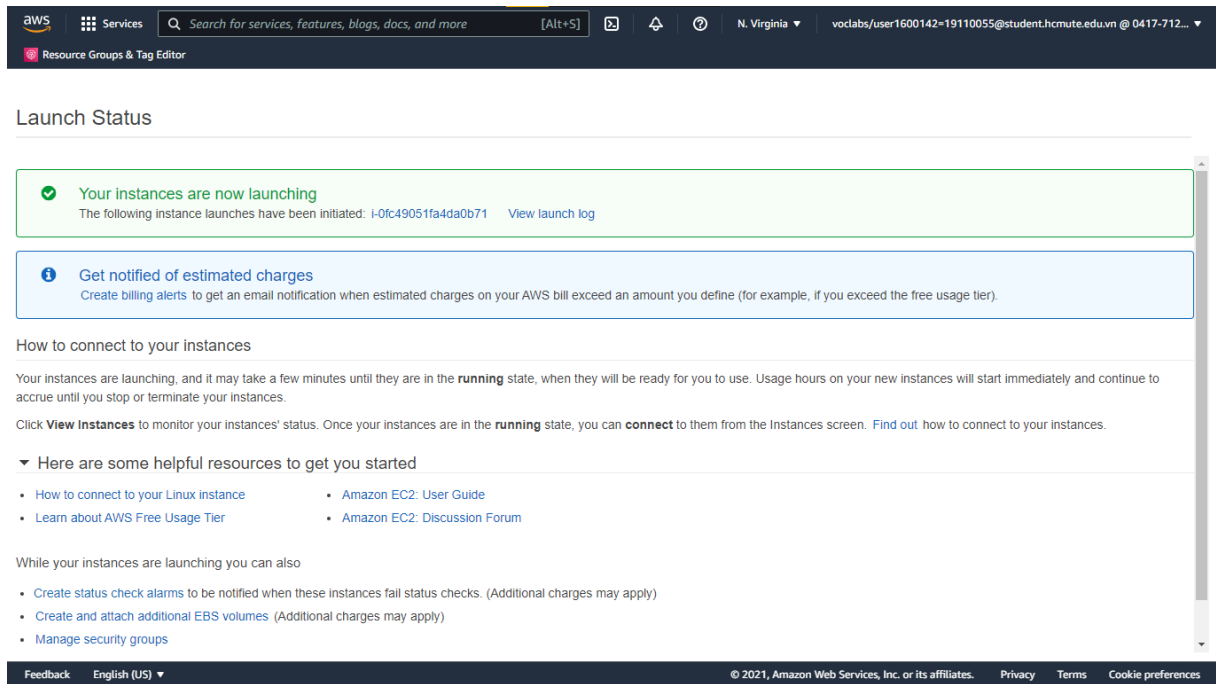
© 2021, Amazon Web Services, Inc. or its affiliates.

Privacy

Terms

Cookie preferences

- Bước 10: Tạo máy ảo EC2 thành công



- Bước 11: Kết nối đến máy ảo thông qua ssh -i "translateapp.pem" ubuntu@ec2-34-204-40-116.compute-1.amazonaws.com

```
OpenSSH SSH client
Microsoft Windows [Version 10.0.19042.1415]
(c) Microsoft Corporation. All rights reserved.

C:\Users\USER>cd Downloads

C:\Users\USER\Downloads>ssh -i "translateapp.pem" ubuntu@ec2-54-234-47-62.compute-1.amazonaws.com
The authenticity of host 'ec2-54-234-47-62.compute-1.amazonaws.com (54.234.47.62)' can't be established.
ECDSA key fingerprint is SHA256:ck8uyW0Fp57d1p2NkuBdsNFsQxQFhaFQ5aLcpZfwP4.
Are you sure you want to continue connecting (yes/no/[fingerprint])? yes
Warning: Permanently added 'ec2-54-234-47-62.compute-1.amazonaws.com,54.234.47.62' (ECDSA) to the list of known hosts.
Welcome to Ubuntu 18.04.5 LTS (GNU/Linux 5.3.0-1033-aws x86_64)

 * Documentation:  https://help.ubuntu.com
 * Management:    https://landscape.canonical.com
 * Support:       https://ubuntu.com/advantage

System information disabled due to load higher than 2.0

 * Ubuntu Pro delivers the most comprehensive open source security and
   compliance features.

   https://ubuntu.com/aws/pro

Get cloud support with Ubuntu Advantage Cloud Guest:
http://www.ubuntu.com/business/services/cloud

236 packages can be updated.
179 updates are security updates.

The programs included with the Ubuntu system are free software;
the exact distribution terms for each program are described in the
individual files in /usr/share/doc/*/copyright.

Ubuntu comes with ABSOLUTELY NO WARRANTY, to the extent permitted by
applicable law.

ubuntu@ubuntu:~$
```

- Bước 12: Cài đặt môi trường cho máy ảo

- + sudo apt-get update
- + sudo apt-get install python3
- + sudo apt-get install python3-pip
- + sudo pip3 install flask
- + sudo pip3 install boto3
- + sudo apt-get install nginx

+ sudo apt-get install gunicorn3

(Có thể thử các lệnh sau nếu quá trình cài đặt môi trường bị lỗi)

+sudo rm /var/lib/apt/lists/lock

+sudo rm /var/cache/apt/archives/lock

+sudo rm /var/lib/dpkg/lock*

- Bước 13: Chạy ứng dụng

+Tải project từ github git clone

+ Di chuyển vào thư mục chứa project dùng ls để kiểm tra có file main.py chưa

+ python3 main.py

PHẦN KẾT LUẬN

1. Kết quả đạt được

Sau một thời gian nghiên cứu, tìm hiểu và thực hiện đề tài “**Tìm hiểu Amazon Translate và viết ứng dụng minh họa**”, nhóm chúng em đã đạt được cơ bản những kết quả như sau:

1.1. Nền tảng kiến thức

Nắm bắt được các kiến thức cũng như những vấn đề liên quan về khái niệm, lợi ích, các sử dụng, đặc điểm, điểm mạnh và tính năng của dịch vụ Amazon Translate.

Hiểu cơ bản về một số dịch vụ kết hợp vào để kết hợp với Amazon Translate như: thư viện giao diện dòng lệnh AWS CLI, thư viện Boto3, Amazon Polly để chuyển văn bản thành giọng nói, Amazon Textract để chuyển hình ảnh thành văn bản, AWS EC2 để deploy ứng dụng lên server của AWS.

1.2. Vận dụng

Xây dựng hoàn chỉnh cơ bản một trang web dịch thuật ứng dụng dịch vụ Amazon Translate, Amazon Polly, Amazon Textract với các chức năng cơ bản như:

- Dịch văn bản từ ngôn ngữ này sang ngôn ngữ khác (**với 71 ngôn ngữ được hỗ trợ**).
- Chuyển văn bản thành giọng nói với nhiều ngôn ngữ (**với 11 ngôn ngữ được hỗ trợ**).
- Chuyển hình ảnh file .png, .jpeg chữ viết tay hoặc đánh máy sang dạng văn bản để tiến hành việc dịch thuật (với số lượng ngôn ngữ truy xuất hạn chế như: Tiếng Anh, Pháp, Đức, Ý, Bồ Đào Nha và Tây Ban Nha).

2. Ưu điểm

- Dịch được các từ vựng và đoạn văn bản trên nhiều ngôn ngữ, độ chính xác cao và văn phong trôi chảy nhờ dịch vụ Amazon Translate.
- Có tính năng chuyển văn bản thành giọng nói giúp thuận tiện trong giao tiếp với người bản địa.
- Chuyển hình ảnh thành văn bản thuận tiện cho việc sao lưu tài liệu.

3. Nhược điểm

- Chưa có tính năng dịch cho tệp tin (.txt, .docx, ...).
- Chưa tìm hiểu được công nghệ tự động phát hiện ngôn ngữ.
- Amazon Polly chỉ hỗ trợ cho 11 ngôn ngữ nên việc đọc các văn bản còn hạn chế.
- Amazon Textract còn hạn chế hỗ trợ truy xuất các hình ảnh có ngôn ngữ khác ngoài các ngôn ngữ được hỗ trợ.
- Giao diện chỉ ở mức cơ bản, còn hạn chế.

4. Hướng phát triển

- Tích hợp trang web này vào các trang báo online để phát huy hơn nữa tính năng của nó.
- Mở rộng dữ liệu đầu vào có thể thể là dạng file .txt, .pdf và giọng nói.
- Mở rộng cho nhiều ngôn ngữ.
- Cải thiện hơn về giao diện.

TÀI LIỆU THAM KHẢO

[1]. Amazon Translate - Dịch máy chính xác và thông thạo

Link: <https://aws.amazon.com/vi/translate/>

[2]. AWS Command Line Interface

Link: <https://aws.amazon.com/vi/cli/>

[3]. What is the AWS Command Line Interface

Link: <https://docs.aws.amazon.com/cli/latest/userguide/cli-chap-welcome.html>

[4]. AWS SDK dành cho Python (Boto 3)

Link: <https://aws.amazon.com/vi/sdk-for-python/>