PROJECT VISION

Xây dựng mô hình dịch máy từ tiếng Anh sang tiếng Việt

Version 1.1

Sinh viên thực hiện:

1612689 – Trương Phạm Nhật Tiến

1612726 – Nguyễn Minh Trí

**Bảng ghi nhận thay đổi tài liệu**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Ngày** | **Phiên bản** | **Mô tả** | **Người thay đổi** | |
| 21/12/2019 | 1.0 | Project vision phần 3, 4, 5 | Nguyễn Minh Trí |
| 22/12/2019 | 1.1 | Project vision phần 1, 2 | Trương Phạm Nhật Tiến |
|  |  |  |  |
|  |  |  |  |

Danh mục

[1. Giới thiệu 3](#_Toc23032854)

[1.1 Mục đích của tài liệu 3](#_Toc23032855)

[1.2 Đối tượng độc giả 3](#_Toc23032856)

[2. Vấn đề 3](#_Toc23032857)

[2.1 Đặt vấn đề 3](#_Toc23032858)

[2.2 Phạm vi của đề tài 3](#_Toc23032859)

[2.3 Lý do của đề tài 3](#_Toc23032860)

[2.4 Những lợi ích đạt được khi giải quyết vấn đề 3](#_Toc23032861)

[2.5 Giả định trường hợp vấn đề 3](#_Toc23032862)

[3. Sản phẩm 3](#_Toc23032863)

[3.1 Những bên liên quan 3](#_Toc23032864)

[3.2 Các sản phẩm liên quan 3](#_Toc23032865)

[3.3 Thời gian giao dự kiến 4](#_Toc23032866)

[3.4 Sản phẩm bàn giao 4](#_Toc23032867)

[3.5 Chi phí và giá cả 4](#_Toc23032868)

[3.6 Môi trường phát triển 4](#_Toc23032869)

[4. Tính năng của sản phẩm 5](#_Toc23032870)

[5. Kết luận 5](#_Toc23032871)

# Giới thiệu

## Mục đích của tài liệu

* Tài liệu thu thập, phân tích và xác định yêu cầu ở mức cao “Xây dựng mô hình dịch máy từ tiếng Anh sang tiếng Việt”.
* Tài liệu tập trung vào nhu cầu của các bên liên quan và người dùng mục tiêu cũng như nguyên nhân của nhu cầu đó.
* Tài liệu cung cấp cái nhìn tổng quan về hệ thống (vấn đề, phạm vi, lý do,…) cũng như hướng phát triển trong tương lai. Nó bao gồm mục đích, phạm vi và các vấn đề liên quan.
* Sau khi đọc tài liệu này thì khách hàng, đội ngũ phát triển và các bên liên quan sẽ có được một cái nhìn tổng quan về các tính năng cơ bản của phần .

## Đối tượng độc giả

* **Khách hàng.**
  + Thầy Ngô Huy Biên – giảng viên trường Đại học Khoa Học Tự Nhiên.
* **Đội ngũ xây dựng và phát triển hệ thống.**
  + Nguyễn Minh Trí – 1612726.
  + Trương Phạm Nhật Tiến – 1612689**.**
* **Người dùng phần mềm:** người dùng phần mềm do khách hàng quyết định.

# Vấn đề

## Đặt vấn đề

* Tiếng Anh của tôi không đủ tốt nên tôi phải thường xuyên tra từ điển Anh-Việt khi gặp nội dung tiếng Anh. Có cách nào để tôi có thể chỉ cần sao chép một từ, một câu, 1 đoạn văn bản hay nguyên cả văn bản rồi có thể dịch nó sang tiếng Việt không?
* Tôi đang học tiếng Anh và khi gặp 1 từ không hiểu tôi phải tra từ điển Anh-Việt nhưng tìm kiếm từ đó khá tốn thời gian. Liệu có cách nào dịch nhanh cho tôi được không?

## Phạm vi của đề tài

* Mô hình chỉ dịch từ tiếng Anh sang tiếng Việt.
* Mô hình chỉ sử dụng được khi có kết nối internet.

## Lý do của đề tài

* Mô hình được xây dựng nhằm giúp những người muốn dịch từ Anh sang Việt nhanh chóng và thuận lợi đạt được nhu cầu.

## Những lợi ích đạt được khi giải quyết vấn đề

* Người dùng có thể hiểu được nội dung tiếng Anh nhanh hơn.
* Hỗ trợ nhu cầu học tiếng Anh của người dùng.
* Có thể kết hợp vào những lĩnh vực liên quan để hỗ trợ hoặc cùng phát triển.

## Giả định trường hợp vấn đề

* **Trường hợp 1:** Trí là một học sinh có nhu cầu nâng cao vốn từ vựng tiếng Anh của mình. Khi đang học tiếng Anh thì Trí gặp phải những từ, câu mà Trí không biết. Việc tra từ điển giấy gặp nhiều khó khăn nên Trí tìm đến công cụ của chúng tôi, Trí chỉ cần sao chép những từ, câu mà trí chưa hiểu và đưa sang công cụ của chúng tôi. Thế là Trí sẽ có nghĩa của những từ, câu mà Trí mong muốn. Điều này giúp cho Trí tiết kiệm được nhiều thời gian tra từ hơn từ đó nâng cao hiệu suất học tập.
* **Trường hợp 2:** Tiến là một nhân viên văn phòng, thời gian tới Tiến được sếp giao cho tìm hiểu một lĩnh vực mới chưa có nhiều tài liệu tiếng Việt liên quan. Nếu Tiến không tìm hiểu được thì sẽ ảnh hưởng tới sự nghiệp của Tiến nhưng vốn tiếng Anh của Tiến rất yếu không thể đọc được. Tiến tìm đến công cụ của chúng tôi, Tiến có thể dịch được từng đoạn văn trong tài liệu của Tiến. Nó giúp Tiến đọc hiểu được các tài liệu về lĩnh vực Tiến tìm hiểu. Nhờ đó giúp tiến có cơ hội thăng tiến trong công việc.

# Sản phẩm

## Những bên liên quan

* **Khách hàng:** thầy Ngô Huy Biên, giảng viên khoa Công nghệ thông tin, trường đại học Khoa Học Tự Nhiên. Bên khách hàng sẽ kiểm tra, đánh giá liên tục sản phẩm và đưa ra các nhận định đánh giá, giúp nhóm hoàn thiện sản phẩm qua các tuần, đồng thời là những người sẽ kiểm định sản phẩm cuối cùng, quyết định sản phẩm có được chấp nhận hay không.
* **Nhóm phát triển:** Gồm 2 thành viên là Nguyễn Minh Trí và Trương Phạm Nhật sẽ chịu trách nhiệm lên kế hoạch, phát triển và kiểm duyệt phần mềm đảm bảo đúng các yêu cầu từ bên khách hàng đưa ra.
* **Người dùng phần mềm:** khách hàng là người dùng có quyền sở hữu phần mềm, có quyền quyết định ai sẽ là người có thể sử dụng sản phẩm này.

## Các sản phẩm liên quan

* Hiện này trên thị trường đã có một số đơn vị triển khai như:
* **Google Translate**
* Điểm mạnh:
  + Dịch được qua lại hơn 90 ngôn ngữ.
  + Có phát âm đầy đủ, chính xác.
  + Hỗ trợ dịch từ ảnh chụp với 26 ngôn ngữ.
  + Có thể tự động phát hiện ngôn .
  + Có thể tự động hiện ra kết quả không cần ấn thêm.
* Điểm yếu:
  + Dịch Anh – Việt còn nhiều sai sót.
* **Cambridge Dictionary**
* Điểm mạnh:
  + Dịch được 23 ngôn ngữ.
  + Có thể tự động phát hiện ngôn ngữ.
  + Kết quả khá chính xác.
* Điểm yếu:
  + Miễn phí tối đa 160 kí tự một lần và 2000 kí tự một ngày.
  + Khi nhập xong cần phải ấn nút Dịch để thực hiện dịch.
* Sản phẩm dự kiến
* **Điểm mạnh**:
  + - Tập trung vào dịch từ tiếng Anh sang tiếng Việt.
    - Kết quả chính xác.
* **Điểm yếu**: chưa phục vụ được nếu không có internet.

## Thời gian giao dự kiến

* Thời gian dự kiến trong 4 tháng, bắt đầu từ ngày 01/01/2020 đến ngày 05/05/2020

## Sản phẩm bàn giao

* Dịch vụ Web API của mô hình dịch máy từ tiếng Anh sang tiếng Việt cho phép tải lên một từ, câu hoặc đoạn văn và trả về dạng chữ tiếng Việt của từ, câu, đoạn văn đó..
* Website mẫu việc sử dụng API của mô hình dịch máy từ tiếng Anh sang tiếng Việt đã xây dựng.

## Chi phí và giá cả

* Số người dự kiến: 2 người.
* Thời gian dự kiến: 4 giờ / ngày / người \* 4 tháng \* 2 người = 960 giờ.
* Chi phí dự kiến:
* Web mẫu sử dụng Amazon EC2 tốn 0đ.
* Web API tải lên máy chủ (server) của Amazon EC2 được miễn phí 1 năm.
* Card màn hình rời 8Gb dùng để huấn luyện API nhận dạng âm thanh có giá khoảng 2.000.000 vnđ.
* Tổng chi phí: 2.000.000 vnđ.

## Môi trường phát triển

* Web API:
* Ngôn ngữ: nodejs, python.
* Giao diện: không có.
* Nền tảng: linux, macos, windows.
* Web mẫu:
* Ngôn ngữ: nodejs.
* Giao diện: ReactJS.
* Nền tảng: Amazon EC2.

# Tính năng của sản phẩm

* Web API:
* Cho phép tải lên một đoạn văn bản và trả về văn bản tiếng Việt của đoạn văn bản đó.
* Cho phép nhiều truy cập cùng lúc.
* Web mẫu:
* Tải lên một đoạn văn vản và trả về văn bản tiếng Việt của đoạn văn bản đó.

# Kết luận

* Với tất cả các sản phẩm sẽ triển khai và xây dựng sắp tới, chúng tôi hy vọng dự án sẽ thành công với tất cả các sản phẩm đạt chất lượng cao như mong đợi. Ngoài ra, chúng tôi có hy vọng rằng những sản phẩm này sẽ thỏa mãn được nhu cầu của tất cả khách hàng, tạo được niềm tin nơi chúng tôi với khách hàng.