ĐỒ ÁN CUỐI KÌ MÔN CCNLTHD

*Từ Điển Online*

# Thông tin nhóm thực hiện

## Thông tin cơ bản

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| STT | MSSV | Họ tên | Email | Điện thoại | Đóng góp |
| 1 | 0812543 | Trần Văn Tri | [trantri2006@gmail.com](mailto:trantri2006@gmail.com) | 0974009485 | 34% |
| 2 | 0812546 | Nguyễn Anh Trí | [nguyenanhtri@gmail.com](mailto:nguyenanhtri@gmail.com) | 01696873563 | 33% |
| 3 | 0812609 | Nguyễn Văn Việt | [vanviet.uos@gmail.com](mailto:vanviet.uos@gmail.com) | 01689938202 | 33% |

## Phân chia công việc

### 0812543 – Trần Văn Tri

-Trình bày ý tưởng chương trình.

-Thiết kế giao diện.

-Thực hiện phần từ điển dùng rút trích dữ liệu từ trang vdict.com

-Thực hiện tìm kiếm khái niệm từ Wikipedia.

### 0812546 – Nguyễn Anh Trí

-Nghiên cứu Google translate API.

-Thực hiện Chức năng Google dịch, dịch đoạn văn.

- Nhận dạng âm thanh thành văn bản.

### 0812609 – Nguyễn Văn Việt

-Phát âm từ, đọc đoạn văn.

-Nhận dạng ngôn ngữ đầu vào

# Thông tin chung về đề tài

* TuDienOnline là một ứng dụng desktop có một số chức năng như sau:
  + Từ điển đa ngôn ngữ
  + Phát âm
  + Dịch đoạn văn
  + Đọc đoạn văn
  + Chuyển âm thành text (nếu được)
  + Kết hợp từ các webservice API: google, Vdict, Wiki,…

# Thuyết minh nội dung

## Phát biểu vấn đề

Hiện nay vấn đề các học sinh, sinh viên nghiên cứu đọc dịch trên các tài liệu nước ngoài rất là phổ biến, nguồn tài liệu internet vô cùng phong phú, nhưng việc tìm hiểu rất khó khăn do rào cản ngôn ngữ, do đó phải có 1 công cụ hỗ trợ dịch thuật tốt hơn. Ứng dụng TuDienOline đáp ứng các nhu cầu sau:

* Cần có 1 từ điển đa ngôn ngữ để tra cứu và phát âm.
* Cần có 1 chương trình đọc và dịch 1 đoạn văn.
* Tìm hiểu nguồn gốc, khái niệm của 1 thông tin, vấn đề.

## Mục tiêu đề tài

* Xây dựng 1 ứng dụng đáp ứng càng nhiều chức năng kể trên càng tốt.
* Phải thực thi nhanh đáp ứng kết quả trong thời gian cho phép.

## Ý nghĩa thực hiện

* Giúp người dùng tra cứu từ điển, dịch tài liệu dễ dàng hơn.
* Giúp người dùng tra cứu nhanh hơn, chính xác hơn.

# Nội dung triển khai

## Kiến trúc đề tài

* Sử dụng kiến trúc WebService.

## Công nghệ sử dụng

* Web API
* Rút trích thông tin web
* Xứ lý âm thanh phát âm và Speech to text.

# Dự kiến kết quả

* Hoàn thành ứng dụng với các chức năng trên
* Google translate.
* Từ điển Vdict: với các chuyển đổi ngôn ngữ như:
  + English - Vietnamese
  + Vietnamese - English
  + Vietnamese - Vietnamese
  + French - Vietnamese
  + Vietnamese - French
  + Computing
  + English - English
  + Chinese - Vietnamese
  + Russian - Vietnamese
* Tìm kiếm trên Wikipedia

# Các khó khăn gặp phải và cách giải quyết

1. Khó khăn

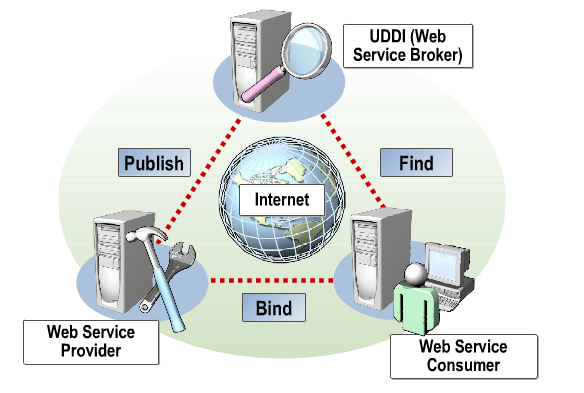
* Hạn chế dữ liệu đầu vào
* Nhận dạng chưa chuẩn, không như ý muốn
* Xử lý âm thanh phức tạp.

1. Cách giải quyết

* Cắt nối dữ liệu

# Kiến trúc hệ thống

-Sử dụng kiến trúc WebService gọi các hàm dịch vụ.



* Rút trích thông tin từ internet.

# Mô tả các chức năng

-Chức năng dịch từ điển, người dùng chọn loại ngôn ngữ rồi nhập từ khóa cần dịch.

-Phát âm từ vừa dịch.

-Chức năng dịch đoạn văn hỗ trợ nhiều ngôn ngữ.

-Chức năng đọc đoạn văn.

-Chức năng tra cứu từ bách khoa toàn thư wikipedia.

-Chức năng chuyển giọng nói thành văn bản.

# Hướng dẫn cài đặt

## Yêu cầu hệ thống

-Window 32 bit.

-Máy tính người dùng phải có kết nối internet.

-Hỗ trợ thư viện play file âm thanh wmp.dll.

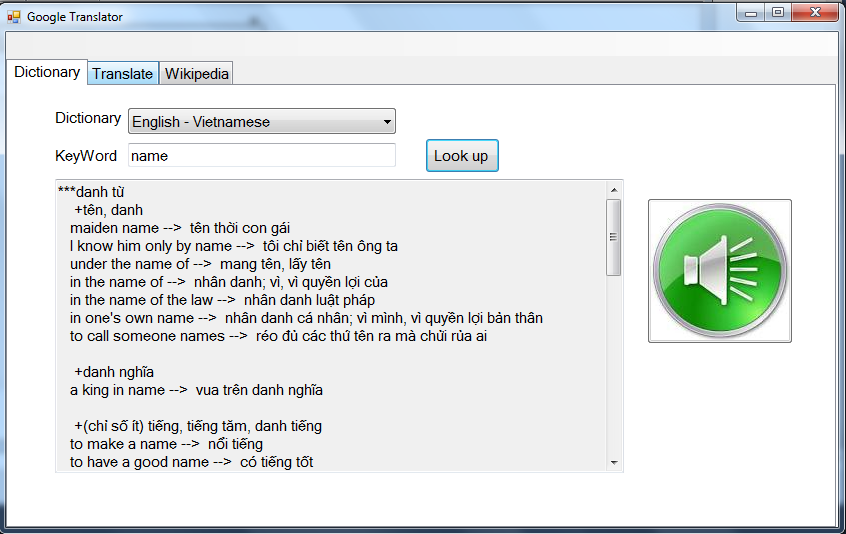
## Hướng dẫn cài đặt

-Chạy file cài đặt trong thư mục Release.

## Hướng dẫn sử dụng

# Trình bày nội dung đạt được

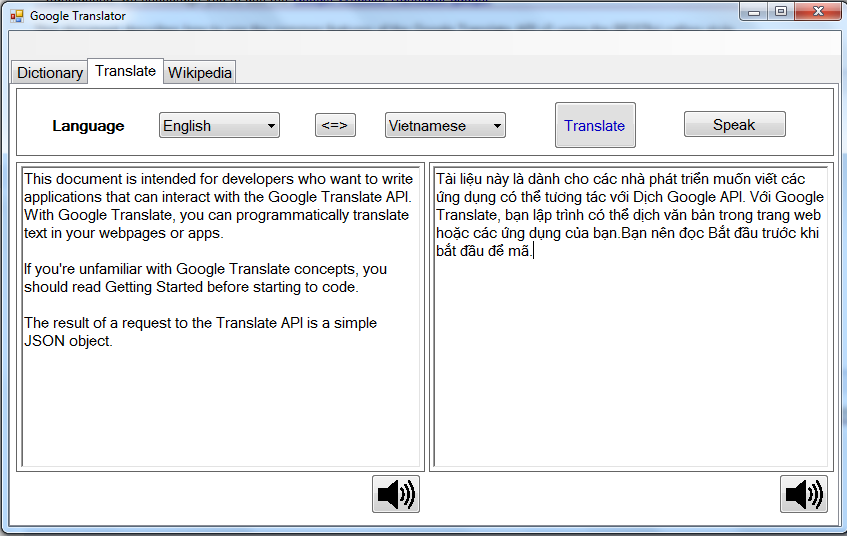
-Chức năng từ điển. Phát âm từ vừa dịch.



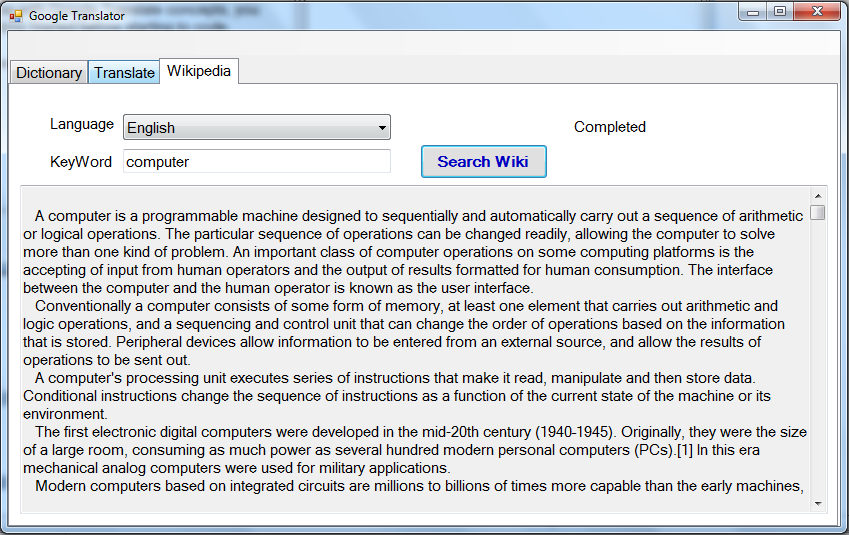
*-*Chức năng dịch đoạn văn (có thể dùng dịch nghĩa như từ điển).

-Chức năng đọc đoạn văn.

-Chức năng đổi giọng thành văn bản.



-Chức năng tra cứu từ Wikipedia.



# Kết luận và hướng phát triển

-Chương trình đã hoàn thiện các chức năng chính, và đúng ý nghĩa mục tiêu ban đầu.

-Chương trình sử dụng các công nghệ của Webservice và rút trích thông tin web.

-Hướng phát triển, chương trình này có thể dùng để viết các ứng dụng trên mobile để tăng sự tiện lơi khi cân tra từ, đặc biệt có ý nghĩa với chức năng Speech to Text.

-Có thể sử dụng Hook API để nhận biết sự kiện select text ở các ứng dụng khác cần dịch.

# Tài liệu tham khảo

* <http://code.google.com/apis/language/translate/overview.html>
* <http://www.wikipedia.org/>
* <http://vdict.com/>