

**Đề tài**

**Nghiên cứu xây dựng hệ thống chuyển đổi tiếng nói tiếng Việt sang văn bản và công cụ hỗ trợ tốc ký trong xử lý nghiệp vụ chuyên môn**

***Tài liệu đặc tả yêu cầu phần mềm***

**Phân hệ 2: Nghiên cứu xây dựng hệ thống nhận dạng người nói và nhận dạng tiếng nói phục vụ chuyển đổi tiếng nói tiếng Việt sang văn bản**

**Phiên bản 1.0**

**08/05/2020**

# Lịch sử thay đổi tài liệu

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Ngày thay đổi | Mục, hình ảnh bảng thay đổi | Mô tả thay đổi | Lý do thay đôi | Thêm/Sửa/Xóa | Người thay đổi | Phiên bản mới |
|  |  |  |  |  |  |  |

# Mục lục

[Lịch sử thay đổi tài liệu 3](#_Toc532400977)

[Mục lục 4](#_Toc532400978)

[Danh mục hình vẽ 6](#_Toc532400979)

[Danh mục bảng 7](#_Toc532400980)

[Danh mục các từ viết tắt 8](#_Toc532400981)

[Danh mục thuật ngữ 9](#_Toc532400982)

[Phần 1 Tổng quan 10](#_Toc532400983)

[1.1 Mô tả chung 10](#_Toc532400984)

[1.2 Mục tiêu và phạm vi tài liệu 10](#_Toc532400985)

[1.2.1 Mục tiêu 10](#_Toc532400986)

[1.2.2 Phạm vi 10](#_Toc532400987)

[1.3 Đối tượng sử dụng tài liệu 10](#_Toc532400988)

[Phần 2 Tổng quan hệ thống 11](#_Toc532400989)

[2.1 Khảo sát hiện trạng 11](#_Toc532400990)

[2.2 Tổng quan chức năng hệ thống 12](#_Toc532400991)

[2.3 Đặc tả chức năng 12](#_Toc532400992)

[2.3.1 Biểu đồ use case tổng quan 12](#_Toc532400993)

[2.3.2 Biểu đồ use case phân rã XYZ 12](#_Toc532400994)

[2.3.3 Quy trình nghiệp vụ 12](#_Toc532400995)

[2.4 Đặc tả chức năng 12](#_Toc532400996)

[2.4.1 Đặc tả use case A 12](#_Toc532400997)

[2.4.2 Đặc tả chức năng B 14](#_Toc532400998)

[Phần 3 Các yêu cầu về giao tiếp bên ngoài 15](#_Toc532400999)

[3.1 Giao diện người dùng (user interface) 15](#_Toc532401000)

[3.2 Giao diện phần cứng (Hardware interface) 15](#_Toc532401001)

[3.3 Giao diện phần mềm (Software interface) 15](#_Toc532401002)

[3.4 Giao tiếp với các phân hệ / hệ thống khác (Communitcation interface) 16](#_Toc532401003)

[Phần 4 Các yêu cầu hệ thống 17](#_Toc532401004)

[4.1 Yêu cầu chức năng 17](#_Toc532401005)

[4.2 Yêu cầu tích hợp 18](#_Toc532401006)

[4.3 Yêu cầu phi chức năng 18](#_Toc532401007)

[4.3.1 Yêu cầu về hiệu năng 18](#_Toc532401008)

[4.3.2 Yêu cầu an toàn 18](#_Toc532401009)

[4.3.3 Yêu cầu về bảo mật 18](#_Toc532401010)

[4.4 Các yêu cầu khác 19](#_Toc532401011)

[Tài liệu tham khảo 20](#_Toc532401012)

[Phụ lục 21](#_Toc532401013)

# Danh mục hình vẽ

Hình 1: Biểu đồ use case …………………………………………………………….. 11

# Danh mục bảng

[Bảng 2.1 Bảng đặc tả chức năng thu 5 bản ghi 1](#_Toc532376363)1

[Bảng 2.2 Dữ liệu đầu vào gồm các trường dữ liệu của chức năng thu 5 bản ghi. 1](#_Toc532376364)2

[Bảng 2.3 Bảng đặc tả chức năng điểm danh 13](#_Toc532376365)

# Danh mục các từ viết tắt

|  |  |
| --- | --- |
| API | Application Programming Interface  Giao diện lập trình ứng dụng |
| **HTML** | HyperText Markup Language  Ngôn ngữ đánh dấu siêu văn bản |
| **CNTT** | Công nghệ thông tin |

# Danh mục thuật ngữ

|  |  |
| --- | --- |
| Browser | Trình duyệt |
| **Cache memory** | Bộ nhớ đệm |
| **E-commerce** | Thương mại điện tử |
| **Bloatware** | Ứng dụng nhà sản xuất tích hợp vào thiết bị |
| **Interpreter** | Trình thông dịch |
| **Compiler** | Trình biên dịch |

# Tổng quan

## Mô tả chung

Nhận diện giọng nói là một vấn đề đang rất được quan tâm, nó là vấn đề có rất nhiều ứng dụng trong thực tế, trên thế giới cũng đã có rất nhiều công trình nghiên cứu về xử lý tiếng nói. Đề tài là một sản phẩm nghiên cứu về nhận dạng người nói nhằm phục vụ quá trình điểm danh bằng giọng nói của một số cơ quan tổ chức. Và tài liệu này có nội dung là phân tích yêu cầu phầm mềm của sản phẩm.

## Mục tiêu và phạm vi tài liệu

### Mục tiêu

Nhận dạng người nói với khả năng nhận diện cao phục vụ cho chức năng điểm danh bằng tiếng nói.

### Phạm vi

Ở trong phạm vi môn học project 2, tài liệu tập trung nghiên cứu nhận dạng giọng nói của một vài người (10 – 20 người), tối ưu khả năng nhận diện một cách tối đa.

Tài liệu này mang tính chất làm quen với machine learning, các yêu cầu về số lượng người được nhận dạng, lọc nhiễu âm thanh sẽ không được chú trọng.

## Đối tượng sử dụng tài liệu

Tài liệu này chỉ được phép lưu hành nội bộ giữa các thành viên trong cùng một phân hệ hoặc một nhóm phân hệ với nhau. Tài liệu cũng là cơ sở để các thành viên hệ thống phối hợp làm việc với nhau, lập trình và tích hợp các module vào một hệ thống thống nhất.

Tài liệu còn là cơ sở để người dùng có thể sử dụng vào việc mở rộng hệ thống. Các quy định về sử dụng tài liệu đối với bên thứ ba là hoàn toàn không có điều kiện về tài chính, bản quyền và tuân theo các quy tắc về tài liệu mở.

# Tổng quan hệ thống

## Khảo sát hiện trạng

## Tổng quan chức năng hệ thống

* Chức năng ghi âm, điều khiển ghi âm
* Chức năng nhận dạng người nói trong file ghi âm

### Biểu đồ use case tổng quan

Ảnh có chứa văn bản, ảnh, bàn, người đàn ông

Mô tả được tạo tự động

Hình 1: Biểu đồ use case

Tác nhân tham gia hệ thống là người dùng tham gia nhận diện giọng nói (điểm danh)

* Mô tả use case:
  + Nếu đây là lần đầu bạn dùng sản phẩm, thì bạn cần nhập tên của bạn và bắt đầu ghi âm 5 bản thu để tạo ra 1 model riêng của bạn
  + Nếu bạn đã có model của mih trong hệ thống, thì bạn chỉ cần ghi âm 1 bản ghi để tiến hành nhận diện (điểm danh)

### Biểu đồ use case phân rã XYZ

### Quy trình nghiệp vụ

## Đặc tả chức năng

### Đặc tả use case thu 5 bản train

Bảng 2‑1 Bảng đặc tả chức năng thu 5 bản train

| Mã Usecase | UC0001 | Tên Usecase | Thu 5 bản train |
| --- | --- | --- | --- |
| Tác nhân hệ thống | Tất cả các tác nhân trong hệ thống đều có thể sử dụng chức năng này. | | |
| Tiền điều kiện | Chưa tồn tại model trong hệ thống | | |
| Luồng sự kiện chính | |  |  |  | | --- | --- | --- | | STT | Thực hiện | Hành động | | 1. | Người dùng mới | Chọn chức năng thu 5 bản ghi | | 2. | Người dùng mới | Nhập tên của bạn | | 2.1 | Người dùng mới | Bắt đầu thu âm | | 2.1.1 | Người dùng mới | Gủi bản ghi | | | |
| Luồng sự kiện thay thế | |  |  |  | | --- | --- | --- | | STT | Thực hiện | Hành động | | 2.2 | Hệ thống | Thông báo tên của bạn bị trùng, yêu cầu nhập tên khác | | | |
| Hậu điều kiện | Không | | |

Bảng 2‑2 Dữ liệu đầu vào gồm các trường dữ liệu của chức năng thu 5 bản train

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| STT | Trường dữ liệu | Mô tả | Bắt buộc | Ví dụ hợp lệ |
| 1 | Name | Tên của bạn | X | Hungdeptrai |
| 2 | Bản thu | Các bản thu âm của bạn | X |  |

### Đặc tả chức năng điểm danh

Bảng 2‑3 Đặc tả chức năng điểm danh

| Mã Usecase | UC0002 | Tên Usecase | Điểm danh |
| --- | --- | --- | --- |
| Tác nhân hệ thống | Tất cả các tác nhân trong hệ thống đều có thể sử dụng chức năng này. | | |
| Tiền điều kiện | Đã tồn tại model trong hệ thống | | |
| Luồng sự kiện chính | |  |  |  | | --- | --- | --- | | STT | Thực hiện | Hành động | | 1. | Người dùng | Chọn chức năng điểm danh | | 2.1 | Người dùng | Bắt đầu thu âm | | 2.1.1 | Người dùng | Gủi bản ghi | | | |
| Luồng sự kiện thay thế | Không | | |
| Hậu điều kiện | Không | | |

# Các yêu cầu về giao tiếp bên ngoài

## Giao diện người dùng (user interface)

## Giao diện phần cứng (Hardware interface)

Yêu cầu người dùng phải có một micro tốt để thực hiện chức năng ghi âm, phục vụ cho quá trình nhận dạng

## Giao diện phần mềm (Software interface)

## Giao tiếp với các phân hệ / hệ thống khác (Communitcation interface)

# Các yêu cầu hệ thống

## Yêu cầu chức năng

Bảng 4‑1 Hiển thị dữ liệu / Tích hợp giao diện người sử dụng

| Hiển thị dữ liệu / Tích hợp giao diện người sử dụng | |
| --- | --- |
| **Mô tả** | Hệ thống sẽ cho phép người dùng tương tác với bản đồ và báo cáo bằng cách sử dụng một giao diện người dùng đã được tích hợp. |
| **Ưu tiên** | Thiết yếu |
| **Cơ sở** | Thuận tiện cho người dùng tiếp cận tất cả các thông tin tại một nơi. |
| **Hàm ý** | * Giao diện người dùng cho phép người dùng tiếp cận với cả bản đồ và báo cáo. * Hệ thống sẽ cho phép quản lý (bật/tắt) tất cả các lớp bản đồ trên một bảng quản lý lớp duy nhất.   Trình bày:   * Người sử dụng phải có thể ẩn/hiện các lớp bản đồ * Các định dạng tỷ lệ cần được áp dụng cho các lớp bản đồ. Một định dạng mặc định sẽ được xác định cho từng loại đối tượng. Hệ thống cần phải có sẵn một loạt các định dạng để sử dụng. Ngoài ra, cũng có thể dùng một số định dạng từ các nguồn bên ngoài để sử dụng trong trường hợp cần thiết (VD: fonts TrueType và các file hình ảnh). |

## Yêu cầu tích hợp

## Yêu cầu phi chức năng

### Yêu cầu về hiệu năng

Nhận dạng giọng nói là một vấn đề hiện đang hot, có nhiều ứng dụng trong thực tiễn. Nhưng để hiệu năng nhận diện cao thì là một vấn đề khó

Yêu cầu của phần mềm là nhận diện người nói với độ chính xác cao nhất có thể, nghĩa là cần có data set đẹp (ít nhiễu,…), trích xuất tính âm thanh và model tốt

### Yêu cầu an toàn

Vì hiệu năng nhận dạng không thể 100%, nên khi nhận dạng xong, cần có sự xác nhận từ người dùng về độ chính xác của kết quả.

### Yêu cầu về bảo mật

Mọi thông tin về người dùng cần được bảo mật

## Các yêu cầu khác