|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **BỘ CÔNG THƯƠNG**  **TRƯỜNG ĐẠI HỌC ĐIỆN LỰC**  **KHOA CÔNG NGHỆ THÔNG TIN**  **C:\Users\VCIS\Desktop\logo-epu-inkythuatso-14-15-47-22.jpg**  **ĐỒ ÁN TỐT NGHIỆP**  **XÂY DỰNG & PHÁT TRIỂN CÔNG CỤ TÁCH NỘI DUNG TỪ PPTX SANG H5P**   |  |  | | --- | --- | | **Giảng viên hướng dẫn** | **: TS. LÊ HOÀN** | | **Sinh viên thực hiện** | **: TRẦN TUẤN ANH** | | **Mã sinh viên** | **: 1981031026** | | **Ngành** | **: CÔNG NGHỆ THÔNG TIN** | | **Chuyên ngành** | **: CÔNG NGHỆ PHẦN MỀM** | | **Lớp** | **: D14CNPM1** | | **Khóa** | **: 2019 - 2024** |   ***Hà Nội, 14 tháng 01 năm 2024*** |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **BỘ CÔNG THƯƠNG**  **TRƯỜNG ĐẠI HỌC ĐIỆN LỰC**  **KHOA CÔNG NGHỆ THÔNG TIN**  **C:\Users\VCIS\Desktop\logo-epu-inkythuatso-14-15-47-22.jpg**  **ĐỒ ÁN TỐT NGHIỆP**  **XÂY DỰNG & PHÁT TRIỂN CÔNG CỤ TÁCH NỘI DUNG TỪ PPTX SANG H5P**   |  |  | | --- | --- | | **Giảng viên hướng dẫn** | **: TS. LÊ HOÀN** | | **Sinh viên thực hiện** | **: TRẦN TUẤN ANH** | | **Mã sinh viên** | **: 1981031026** | | **Ngành** | **: CÔNG NGHỆ THÔNG TIN** | | **Chuyên ngành** | **: CÔNG NGHỆ PHẦN MỀM** | | **Lớp** | **: D14CNPM1** | | **Khóa** | **: 2019 - 2024** |   ***Hà Nội, 14 tháng 01 năm 2024*** |

**DANH SÁCH NHỮNG NGƯỜI THAM GIA THỰC HIỆN ĐỀ TÀI**

|  |  |  |
| --- | --- | --- |
| **TT** | **Họ và tên** | **Cơ quan/tổ chức** |
| 1 | Trần Tuấn Anh | Sinh viên đại học Điện Lực |
| 2 |  |  |
| 3 |  |  |
| 4 |  |  |
| 5 |  |  |
| 6 |  |  |
| 7 |  |  |
| 8 |  |  |
| 9 |  |  |
| 10 |  |  |

***Lời cảm ơn***

Em xin dành một khoảnh khắc để bày tỏ lòng biết ơn sâu sắc đến Thầy giáo TS Lê Hoàn và nhà trường trong quá trình thực hiện nghiên cứu này. Thầy giáo không chỉ là người hướng dẫn nghiên cứu, mà còn là nguồn động viên và người đồng hành quan trọng, giúp em vượt qua những thách thức và phát triển khả năng nghiên cứu của mình.

Sự tận tâm và sự sáng tạo của Thầy giáo đã mang lại không chỉ kiến thức sâu rộng mà còn là sự khích lệ và chiến lược hữu ích cho hành trình nghiên cứu của em. Những góp ý chân thành đã giúp em hiểu rõ hơn về đề tài và phát triển những ý tưởng mới.

Em cũng muốn bày tỏ lòng biết ơn đặc biệt đến nhà trường, nơi mà em đã có cơ hội học tập và nghiên cứu đồ án tốt nghiệp. Sự hỗ trợ toàn diện và không ngừng từ nhà trường đã tạo ra một môi trường lý tưởng cho em khám phá, sáng tạo và áp dụng kiến thức vào thực tế.

Những cơ hội nghiên cứu và hoàn thành đồ án mà nhà trường đã tạo ra không chỉ giúp em phát triển kỹ năng mà còn mở ra những cánh cửa mới trong sự hiểu biết và ứng dụng. Em biết ơn vì sự hỗ trợ này đã giúp em không chỉ hoàn thành đồ án mà còn đào sâu và xây dựng nền tảng vững chắc cho sự nghiệp sau này.

Sinh viên

Trần Tuấn Anh

Mục Lục

[BẢNG CHÚ GIẢI CÁC CHỮ VIẾT TẮT, KÝ HIỆU 5](#_Toc155585469)

[DANH MỤC HÌNH ẢNH 6](#_Toc155585470)

[LỜI MỞ ĐẦU 7](#_Toc155585471)

[CHƯƠNG 1: TỔNG QUAN ĐỀ TÀI 8](#_Toc155585472)

[1. Đặt vấn đề 8](#_Toc155585473)

[2. Mục tiêu đề tài 8](#_Toc155585474)

[CHƯƠNG 2: TỔNG QUAN TÌNH HÌNH NGHIÊN CỨU TRONG, NGOÀI NƯỚC 9](#_Toc155585475)

[1. Tình hình nghiên cứu ngoài nước. 9](#_Toc155585476)

[2. Tình hình nghiên cứu trong nước. 10](#_Toc155585477)

[CHƯƠNG 3: NỘI DUNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU 11](#_Toc155585478)

[1. Nội dung nghiên cứu 11](#_Toc155585479)

[2. Phương pháp nghiên cứu 12](#_Toc155585480)

[2.1 Phương pháp nghiên cứu Power Point 12](#_Toc155585481)

[2.2 Phương pháp nghiên cứu H5P (html 5 package) 13](#_Toc155585482)

[2.3 Các giải pháp khả thi phù hợp cho xây dựng công cụ chuyển đổi. 13](#_Toc155585483)

[3. Địa điểm nghiên cứu 14](#_Toc155585484)

[4. Thời gian nghiên cứu công cụ chuyển đổi 14](#_Toc155585485)

[5. Kết quả và đánh giá 15](#_Toc155585486)

[CHƯƠNG 4: KẾT LUẬN VÀ KIẾN NGHỊ 18](#_Toc155585487)

[TÀI LIỆU THAM KHẢO 19](#_Toc155585488)

# BẢNG CHÚ GIẢI CÁC CHỮ VIẾT TẮT, KÝ HIỆU

|  |  |
| --- | --- |
| Ký hiệu | Ý nghĩa |
| PPTX | Tệp PowerPoint |
| H5P | Html 5 package |
| XML | Extensible Markup Language |
| Html | HyperText Markup Language |
| LMS | Learning Management System |
| JSON | JavaScript Object Notation |
| Shape | Các định dạng kiểu dáng |
| Slide | Trang chiếu trong PowerPoint |

# DANH MỤC HÌNH ẢNH

[Hình 1: Một vài loại nội dung H5P - I (https://H5P.org/) 9](#_Toc154192849)

[Hình 2: Các loại nội dung H5P – II (https://H5P.org/) 10](#_Toc154192850)

[Hình 3: cấu trúc một tệp PowerPoint 12](#_Toc154192851)

[Hình 4: : cấu trúc một tệp H5P 13](#_Toc154192852)

[Hình 5: Slide ban đầu của tệp PowerPoint 15](#_Toc154192853)

[Hình 6: Slide của tệp H5P đã được chuyển đổi 15](#_Toc154192854)

[Hình 7: Slide ban đầu của tệp PowerPoint 16](#_Toc154192855)

[Hình 8: Slide sau khi sử dụng công cụ chuyển đổi 16](#_Toc154192856)

## LỜI MỞ ĐẦU

Trong thời đại công nghệ ngày nay, việc tích hợp công nghệ vào quá trình giáo dục là một xu hướng không thể phủ nhận. Cùng với sự phổ biến của các tệp PowerPoint (PPTX) và nền tảng tương tác trực tuyến H5P, việc xây dựng và phát triển công cụ giúp tách nội dung từ PPTX và tích hợp vào H5P trở thành một yếu tố quan trọng, hứa hẹn mang lại hiệu quả tương tác cao trong lĩnh vực giáo dục.

Đề tài "Xây dựng & Phát triển Công cụ Tách Nội Dung từ PPTX Sang H5P" tập trung vào việc nghiên cứu chi tiết về kiến trúc của tệp PowerPoint (PPTX) để hiểu rõ cấu trúc và cách tổ chức thông tin trong chúng. Đồng thời, nghiên cứu về kiến trúc của H5P nhằm phát triển thuật toán cho việc thêm nội dung vào tệp H5P một cách hiệu quả. Mục tiêu của đề tài là tạo ra một công cụ hỗ trợ người dùng tạo tệp H5P một cách thuận tiện và nhanh chóng.

Với sự phổ biến của H5P trong môi trường giáo dục trực tuyến, công cụ mà chúng tôi phát triển hứa hẹn mang lại lợi ích đáng kể cho cộng đồng giáo viên và người học. Việc tự động hóa quá trình chuyển đổi từ PPTX sang H5P không chỉ giúp tiết kiệm thời gian mà còn tăng cường khả năng tương tác và hiệu quả trong quá trình giảng dạy.

Bằng việc kết hợp sự hiểu biết vững về cấu trúc tệp PowerPoint và nền tảng H5P, chúng tôi hy vọng rằng đề tài của mình không chỉ đáp ứng nhu cầu ngày càng tăng về sự linh hoạt trong việc tạo tài liệu giảng dạy mà còn góp phần vào sự phát triển của giáo dục trực tuyến hiện đại.

# TỔNG QUAN ĐỀ TÀI

## Đặt vấn đề

* Microsoft PowerPoint là một phần mềm trình chiếu do Microsoft phát triển. PowerPoint là một phần của gói ứng dụng văn phòng Microsoft Office. Nó có thể cài đặt và sử dụng được trên cả máy tính dùng hệ điều hành Windows lẫn Mac OS X. Nó thường được sử dụng trong các buổi thueytes trình cũng như trong các tiết học trên lớp. Do hình thành từ lâu nên được hỗ trợ rất nhiều ví dụ như các nền tảng hỗ trợ tạo template nhanh để người dùng có thể ngay lập tức có được một bài thuyết trình. Tuy nhiên thì nó có nhiều hạn chế khi thương mại hóa, vì nó rất nặng nên mỗi khi muốn xem một file như thế trên web sẽ gây tốn dữ liệu và thời gian.
* H5P là viết tắt của HTML5 Package. Đây là một cách đơn giản để tạo và chia sẻ những nội dung phong phú và để tương tác với nội dung trong quá trình giảng dạy. Chính xác hơn, H5P là bộ công cụ mã nguồn mở giúp người dùng biên soạn nhiều loại hình nội dung tương tác cho các khóa học trực tuyến, trang web hoặc blog cá nhân. Mỗi nội dung tương tác là một gói HTML5 bao gồm HTML, CSS và Javascript đóng gói cùng nhau. H5P là sản phẩm của một dự án do cộng đồng dẫn lái giúp phát triển những công cụ có nhiều tính năng linh hoạt để tạo các nội dung học tập phong phú, tùy biến, dễ chia sẻ và có thể tái sử dụng. Có thể nói rằng bộ công cụ này không chỉ là một ngôi sao đang lên mà còn là một ‘game changer’ thay thế cho phương thức tương tác truyền thống dựa trên nền tảng Adobe Flash. Điểm mạnh của bộ công cụ này là: dễ sử dụng, linh hoạt và mã nguồn mở.
* Vì H5P tiện lợi như vậy, còn PowerPoint lại có từ lâu, nhiều nền tảng hỗ trợ tạo nhanh template,… Kế thừa các template chúng em viết phần mềm hỗ trợ bóc tách nội dung từ PowerPoint chuyển sang H5P gồm các text, images, shape.

## Mục tiêu đề tài

* Đề tài tập trung vào việc nghiên cứu và phát triển kỹ thuật bóc tách nội dung từ PowerPoint và chuyển đổi nó sang định dạng H5P. Mục tiêu chính là hiểu rõ cách bóc tách những thành phần như text, hình ảnh, và shape từ PowerPoint, sau đó thực hiện quá trình chuyển đổi chúng vào định dạng H5P. Điều này nhằm mục đích hỗ trợ tạo ra các file H5P tương tác, giúp tăng cường trải nghiệm học tập và giảng dạy trong môi trường giáo dục.

# TỔNG QUAN TÌNH HÌNH NGHIÊN CỨU TRONG, NGOÀI NƯỚC

## 1. Tình hình nghiên cứu ngoài nước.

- H5P là một mô-đun dựa trên HTML5 mã nguồn mở được tạo bởi cộng đồng các nhà giáo dục để làm cho việc học trực tuyến trở nên tương tác. Nó hỗ trợ học tập tích cực, một phương pháp giảng dạy đã được chứng minh và rất phổ biến được sử dụng trong học tập trực tuyến. Việc tích hợp nội dung H5P trong bất kỳ khóa học trực tuyến nào sẽ khiến khóa học đó trở nên thú vị và mang tính cộng tác. Nó tạo cơ hội cho người học tương tác và suy nghĩ chín chắn về những gì họ đang học trên máy tính mà không có sự hiện diện trực tiếp và/hoặc trực tiếp của giáo viên. Một lợi ích tuyệt vời khác của H5P là nó cho phép bạn dễ dàng chia sẻ và sử dụng lại nội dung (Singleton & Charlton, 2020). H5P hỗ trợ các hoạt động hoặc loại nội dung thú vị sau: A screenshot of a cell phone

Description automatically generated

Hình 1: Một vài loại nội dung H5P - I ([https://H5P.org/](https://h5p.org/))

* Tất cả các hoạt động trên có thể dễ dàng được tích hợp vào bất kỳ khóa học MODLE, Elearning nào. Ví dụ: Điểm truy cập hình ảnh có thể tạo nhiều vùng thông tin trong một hình ảnh. Các loại nội dung không giới hạn ở đề cập ở trên mà còn bao gồm các nội dung khác như Ghép nối hình ảnh, Thanh trượt hình ảnh, Sách tương tác và Chuyến tham quan ảo (360). Những thay đổi nhanh chóng trong thế giới công nghệ ngày càng tăng và giáo dục từ xa trên thế giới, cũng như việc chấp nhận các phát minh và kỹ thuật mới, công nghệ và phương pháp tiếp cận mới đang dần thay thế các truyền thống kỹ thuật cũ (Iqbal, Mir, & Janjua, 2020).

A screenshot of a computer

Description automatically generated

Hình 2: Các loại nội dung H5P – II ([https://H5P.org/](https://h5p.org/))

* Hoạt động video tương tác dựa trên HTML5 này cho phép người dùng thêm hoặc nhúng video hiện có và thêm tương tác trong video đó, chẳng hạn như đặt câu hỏi trắc nghiệm hoặc điền vào ngân hàng trong video vào một thời điểm cụ thể. Nó làm cho video trở nên hấp dẫn và tương tác, đồng thời phục vụ tính năng của lớp đồng bộ trong video được ghi/không đồng bộ. Theo dõi bối cảnh học tập trực tuyến và video tương tác H5P.

- Vì nhiều lợi ích mang lại từ H5P, đã nhiều đơn vị cố gắng phát triển chuyển đổi nội dung từ các nền tảng như Canvas, PowerPoint nhưng chưa thành công, chưa có gì nổi trội. Không thể tạo các nội dung cũng như câu hỏi tự động. Kế thừa các images, background từ các slide.

## 2. Tình hình nghiên cứu trong nước.

- Hiện tại có nhiều doanh nghiệp, các cá nhân tổ chức thấy được tiềm năng của H5P (html5 package), họ đã ứng dụng nó vào các nền tảng giáo dục đào tạo trực tuyến như Moodle, các hệ thống giáo dục trực tuyến LMS, các hệ thống quản lý khóa học trực tuyến elearning.

- Các nội dung bài giảng được kết hợp sử dụng H5P để tăng tương tác giữa người học và người dạy, nhằm tăng hiệu quả của buổi học. Nhiều doanh nghiệp, các đơn vị tổ chức, cá nhân đã ứng dụng H5P vào giảng dạy, tối ưu tài nguyên hệ thống. Vốn dĩ nếu sử dụng các nền tảng cũ như PowerPoint sẽ gây tốn tài nguyên hơn khi triển khai trên các hệ thống webapp. Khi sử dụng H5P sẽ giúp tối ưu tài nguyên, cũng như tăng trải nghiệm hiệu quả học tập khi sử dụng cho mục đích giảng dạy.

- Hiện tại, ở Việt Nam cũng chưa có đơn vị nào công khai, hay các bài báo nào em tìm thấy có thể chuyển đổi, bóc tách nội dung từ PowerPoint, Canvas sang H5P. Căn cứ vào đấy em đã nghiên cứu và phát triển công cụ tách nội dung từ PowerPoint sang định dạng H5P hỗ trợ tính tương tác giữa người học và giảng viên. Nhằm hỗ trợ cho người dùng chuyển đổi file H5P bằng cách kế thừa nội dung từ tệp PowerPoint, Canvas đã thiết kế.

# NỘI DUNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU

## 1. Nội dung nghiên cứu

- Nghiên cứu này tập trung vào phân tích và triển khai các chức năng chuyển đổi nội dung từ định dạng PowerPoint sang định dạng H5P, tạo ra một công cụ hỗ trợ quá trình này. Các công việc chính trong nghiên cứu bao gồm:

* **Phân tích các đối tượng JSON:**
* Tiến hành phân tích các đối tượng JSON liên quan đến nội dung trong H5P như text, video, và hình ảnh.
* **Viết hàm chuyển đổi đối tượng:**
* Xây dựng các hàm chuyển đổi để chuyển đổi đối tượng từ định dạng PowerPoint sang định dạng H5P, bao gồm cả text, video, và hình ảnh.
* **Tạo template cho bóc tách nội dung:**
* Phát triển template để hỗ trợ quá trình bóc tách nội dung từ PowerPoint, giúp đơn giản hóa và tăng cường hiệu suất công việc.
* **Phân tích cú pháp XML trong Tệp pptx sử dụng OPENXML SDK:**
* Nghiên cứu và phân tích cú pháp XML và sử dụng OPENXML SDK để tạo và chỉnh sửa tệp PowerPoint. Trích xuất các đối tượng ảnh, video, văn bản.
* **Sử dụng Python-pptx xử lý Background Slide:**
* Áp dụng thư viện python-pptx để xử lý nền của các slide trong PowerPoint.
* **Tìm hiểu kỹ thuật làm ảnh trong suốt:**
* Nghiên cứu mã nguồn để hiểu cách tạo ảnh trong suốt và tích hợp chúng vào quá trình chuyển đổi.
* **Lập Báo Cáo:**
* Tổng hợp kết quả và trình bày chúng trong báo cáo, mô tả quy trình chuyển đổi và những cải tiến đạt được.
* Sử dụng công cụ cho việc thử nghiệm chuyển đổi bên cạnh đó sẽ đề xuất ra các giải pháp cải thiện công cụ.

- Nghiên cứu này đề xuất một giải pháp toàn diện để chuyển đổi nội dung PowerPoint sang định dạng H5P, kết hợp giữa phân tích cú pháp XML, sử dụng các thư viện như OPEN XML SDK và python-pptx, và tối ưu hóa quy trình thông qua việc xây dựng template và hàm chuyển đổi.

## 2. Phương pháp nghiên cứu

### 2.1 Phương pháp nghiên cứu Power Point

#### 2.1.1 XML & PowerPoint

- Đầu tiên chúng ta phải làm quen khái niệm XML. XML là gì? Được tạo ra bởi World Wide Web Consortium (W3C), XML hay Extensible Markup Language, dịch sang tiếng Việt chính là “ngôn ngữ đánh dấu mở rộng” có chức năng truyền đạt và mô tả nhiều loại dữ liệu khác nhau. Từ đó, việc chia sẻ dữ liệu sẽ giữa các hệ thống và nền tảng thông qua mạng Internet sẽ diễn ra vô cùng đơn giản.

- Ví dụ: Khi bạn xây dựng ứng dụng A bằng PHP và ứng dụng B bằng Java. Lúc này, hai ngôn ngữ của 2 ứng dụng sẽ không thể hiểu nhau, vì vậy ta cần đến file XML để trao đổi và mã hóa dữ liệu.

- Tệp Power Point bản chất là một tệp viết bằng XML, được đóng gói nội dung lại. Dưới đây là một cấu trúc đơn giản của tệp PowerPoint.

A close up of a number

Description automatically generated

Hình 3: cấu trúc một tệp PowerPoint

#### 2.1.2 Xử lý tệp PowerPoint

- Khi đã hiểu được rằng tệp PowerPoint được viết bằng XML và đóng gói lại. Chúng ta có thể khai thác khía cạnh đấy để nghiên cứu các đối tượng trong PowerPoint. Chúng ta sẽ tiến hành nghiên cứu các đối tượng văn bản, các khối hình, các ảnh, các video được lưu trữ như thế nào, ký hiệu và đánh dấu ra sao. Từ đó có thể xây dựng các kỹ thuật để trích xuất các đối tượng cần thiết từ tệp PowerPoint ra.

- Phương pháp xử lý tệp PowerPoint sẽ chia thành các bước sau:

1) Truy cập vào tệp Slide Layouts, trích xuất ra layout của các slide.

2) Truy cập vào tệp font, media, để trích xuất các đối tượng hình ảnh, video, các đối tượng kiểu font, màu chữ được sử dụng trong PowerPoint.

3) Truy cập vào tệp Slides để trích xuất vị trí các đối tượng video, văn bản, khối hình được lưu và đánh dấu ở trong.

4) Sau khi có được các đối tượng shape, images, text, video từ từng slide. Ta tiến hành nén và lưu lại. Chuẩn bị cho việc chuyển đổi sang tệp H5P (html5 package).

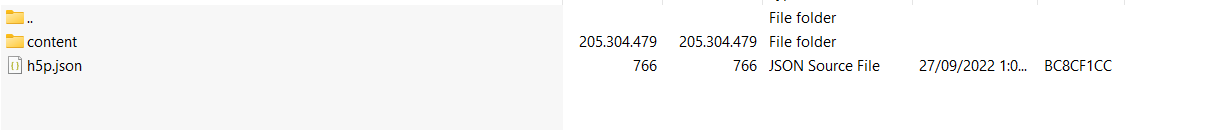
### 2.2 Phương pháp nghiên cứu H5P (html 5 package)

#### 2.2.1 JSON & H5P ( html 5 package)

- JavaScript Object Notation hay JSON là một kiểu dữ liệu mở trong JavaScript. Kiểu dữ liệu này bao gồm chủ yếu là văn bản, có thể đọc được theo dạng cặp "thuộc tính - giá trị". Về cấu trúc, nó mô tả một vật thể bằng cách bọc những vật thể con trong vật thể lớn hơn trong dấu ngoặc nhọn.

- Trong H5P hay html 5 packages, các tệp được lưu dữ liệu dưới dạng JSON. Dựa trên khái niệm này, ta có thể truy cập vào tệp JSON, thay đổi các nội dung trong tệp H5P.

#### 2.2.2 Xử lý tệp H5P (html 5 package)

Dưới đây là hình ảnh cấu trúc tệp cơ bản của một tệp H5P. 

Hình 4: : cấu trúc một tệp H5P

* Ta có thể truy cập vào từng tệp JSON trong H5P, thay đổi nội dung trong file sao cho phù hợp với cấu trúc có sẵn. Sau đây là các bước xử lý tệp H5P:
  + 1. Lấy các dữ liệu đã lưu lại được trích xuất từ các nội dung trong tệp PowerPoint gồm các ảnh, khối hình, video, văn bản trong từng slide.
    2. Thêm lần lượt chúng vào từng tệp JSON trong H5P. Làm sao cho phù hợp với cấu trúc được định nghĩa sẵn trong H5P.
    3. Sau đó tiến hành nén chúng lại, thành một tệp H5P duy nhất. Ta sẽ có được tệp H5P sau chuyển đổi.

### 2.3 Các giải pháp khả thi phù hợp cho xây dựng công cụ chuyển đổi.

- Chúng ta sẽ chia ra các bài toán nhỏ để giải quyết, từ đây lựa chọn ra giải pháp nào là phù hợp nhất:

1) Kế thừa toàn bộ nội dung có trong từng slide dưới dạng hình ảnh

2) Chuyển đổi toàn bộ nội dung gồm các khối, hình ảnh, nền trang chiếu, văn bản, video.

- Hai giải pháp trên là giữ nguyên toàn bộ nội dung vốn có của tệp PowerPoint đã tạo trước đó. Nếu tiến hành được hai bước này hoàn hảo, thì sẽ mang lại hiệu quả tuyệt vời. Chúng ta có thể kế thừa toàn bộ nội dung đã tạo sẵn từ các nền tảng như PowerPoint, Canvas. Các định dạng tệp đã được tạo sẵn.

## 3. Địa điểm nghiên cứu

- Vì đây là nghiên cứu khoa học xây dựng công cụ hỗ trợ chuyển đổi nội dung trên trên các phần mềm với nhau. Vì thế có chỉ cần một máy tính cá nhân và một đội nhóm hỗ trợ nhau giải quyết vấn đề thì hoàn toàn có thể đẩy nhanh tiến độ nghiên cứu khoa học.

## 4. Thời gian nghiên cứu công cụ chuyển đổi

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **TT** | **Nội dung** | **Người thực hiện** | **Thời gian**  **(từ ÷ đến)** | **Kết quả** |
| 1 | Tìm hiểu cấu trúc định dạng PowerPoint (cấu trúc tệp tin pptx) | Trần Tuấn Anh | Từ tháng 9 đến tháng 10 | Báo cáo phân tích định dạng cấu trúc tổng quát của PowerPoint và đưa ra kết quả về định dạng cần chuyển đổi |
| 2 | Tìm hiểu cấu trúc định dạng H5P (cấu trúc tệp tin H5P) | Trần Tuấn Anh | Từ tháng 10 đến tháng 11 | Báo cáo phân tích định dạng cấu trúc tổng quát của H5P và đưa ra kết quả về định dạng H5P cho việc chuyển đổi |
| 3 | Phân tích một số nền tàng chứa nội dụng tạo bài trình diễn có thể chuyển về định dạng pptx (ví dụ như Canvas)  Phân tích khả năng tương tác của H5P và Thực hiện cài đặt tách nội dung từ PowerPoint sang H5P | Trần Tuấn Anh | Từ tháng 10 đến tháng 11 | Các mô hình và các yêu cầu hệ thống, biểu đồ phân tích và kết quả chuyển đổi các đối tượng trong phần mềm trình chiếu từ PowerPoint sang H5P |
| 4 | Viết báo cáo kết quả nghiên cứu | Trần Tuấn Anh | Tháng 12 | Bản báo cáo và phần mềm thực nghiệm |
|  |  |  |  |  |

## 5. Kết quả và đánh giá

- Chúng ta sẽ lần lượt đánh giá các giải pháp dưới đây:

1) Kế thừa toàn bộ nội dung có trong từng slide dưới dạng hình ảnh

- Dưới đây là các hình ảnh của 01 slide được chuyển từ PowerPoint sang H5P:

A cartoon characters with a magnifying glass

Description automatically generated

Hình 5: Slide ban đầu của tệp PowerPoint



Hình 6: Slide của tệp H5P đã được chuyển đổi

- Có thể thấy kết quả này công cụ đã làm rất tốt, tuy nhiên thì sẽ bị hạn chế về hoạt ảnh. Vì H5P không hỗ trợ các hoạt ảnh chuyển động đặc sắc như PowerPoint nên nó sẽ hơi mất mỹ quan.

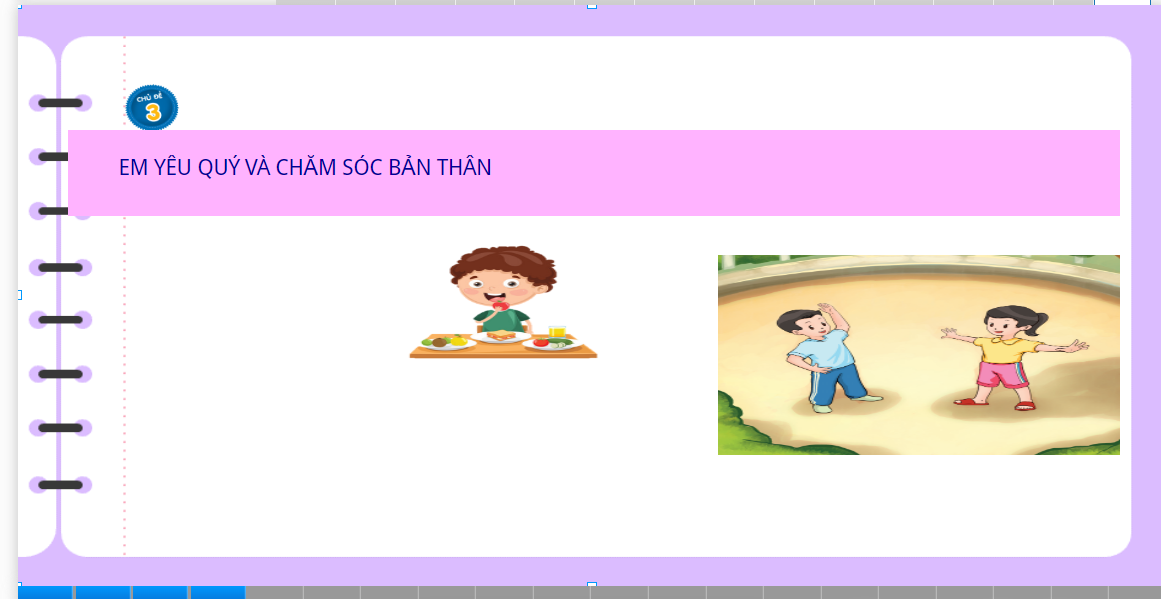
2) Chuyển đổi toàn bộ nội dung gồm các khối, hình ảnh, nền trang chiếu, văn bản, video.

- Có thể hiểu kỹ thuật này là bóc tách từng đối tượng trong tệp PowerPoint ra sau đó tiến hành chuyển đổi từng đối tượng sang H5P. Sau đây là kết quả của giải pháp này:

A cartoon of a child

Description automatically generated

Hình 7: Slide ban đầu của tệp PowerPoint



Hình 8: Slide sau khi sử dụng công cụ chuyển đổi

* Ở giải pháp này rất khó tiến hành đạt hiệu quả cao, vì có nhiều xung đột trong quá trình bóc tách nội dung giữa hai nền tảng H5P và PowerPoint.

Cụ thể cho các lý do gây gián đoạn quá trình chuyển đổi:

- Về khung hình thì 2 nền tảng này có độ dài rộng của khung hình trình chiếu khác nhau nên sẽ khó chuyển nội dung một cách tương đồng về tỷ lệ, gây mất thẩm mỹ. Đặc biệt là đối với các tỷ lệ hình ảnh bị vỡ, các dòng văn bản căn lề bị lệch.

- Trong quá trình bóc tách các ảnh, có các ảnh bị dính nền trắng, sẽ gây mất mỹ quan. Từ đó phải thực hiện quá trình chuyển ảnh sang ảnh trong suốt (ảnh không nền).

- Về font chữ, trong H5P chỉ hỗ trợ 1 font duy nhất, nên không thể chuyển đổi font từ PowerPoint sang H5P được.

- Trong tệp PowerPoint, được các người dung tạo các bố cục mẫu, sử dụng nhiều kiểu hình ảnh định dạng khác nhau như JPEG, PNG, JPG, … và một vài trong số chúng không được duyệt trong định dạng H5P vì H5P không hỗ trợ định dạng ảnh này.

- Trong PowerPoint có sử dụng rất nhiều các đối tượng kiểu shape, các đối tượng này có rất nhiều định dạng trong PowerPoint, tuy nhiên nó lại không có nhiều các loại định dạng trong H5P. Nó chỉ tồn tại một vài kiểu shape được hỗ trợ.

* **Giải pháp đề xuất giải quyết các vấn đề trên hiệu quả hơn sẽ như sau:**

- Xây dựng các định dạng teamplate mẫu của H5P có thẩm mỹ hơn, sau đó sẽ kế thừa một vài định dạng ảnh và các văn bản trực tiếp tạo thành từng giao diện đẹp mắt hơn, tính tương tác cao hơn mà không cần tạo thủ công. Bởi khi làm thế, sẽ không mất công tính toán tỷ lệ giữa các đối tượng trong PowerPoint và H5P. Không phải tính toán lại vị trí các hình ảnh, văn bản sao cho phù hợp hơn trong H5P.

- Về vấn đề các loại định dạng hình ảnh, ta có thể tiến hành chuyển đổi định dạng các loại ảnh sao cho có định dạng phù hợp với tệp H5P.

- Về các ảnh bị dính nền trắng, không trong suốt. Chúng ta có thể nghiên cứu giải pháp, xây dựng kho ảnh trong suốt. Sử dụng AI nhận diện các ảnh có sẵn bị có nền trắng và tìm các hình ảnh tương tự trong kho ảnh. Từ đó đổi 2 ảnh này cho nhau. Ngoài ra cũng có thể nghiên cứu thuật toán chuyển ảnh có nền thành ảnh trong suốt. Kết hợp thuật toán làm mịn ảnh để không bị tình trạng ảnh sau khi chuyển vẫn còn nền trắng dạng răng cưa.

# KẾT LUẬN VÀ KIẾN NGHỊ

Đề tài nghiên cứu khoa học đã triển khai được 02 nội dung là:

1. Kế thừa toàn bộ nội dung có trong từng slide dưới dạng hình ảnh.
2. Chuyển đổi toàn bộ nội dung gồm các khối, hình ảnh, nền trang chiếu, văn bản, video.
3. Thêm được một số nội dung tương tác tự động

- Trong quá trình thực hiện đề tài, do có rất nhiều vấn đề phát sinh, số thành viên để hoàn thành các khối lượng công việc là khó hoàn thành hết, dẫn tới nhiều sai sót. Tuy nhiên đã tìm ra các giải pháp hỗ trợ phần nào các vấn đề phát sinh trong quá trình bóc tách nội dung từ PowerPoint sang H5P.

- Về khả năng mở rộng của đề tài, có thể ứng dụng đề tài này trong việc tạo các tệp tương tác trên H5P. Kế thừa các nội dung trong PowerPoint. Đặc biệt khi mở rộng đề tài hơn nữa có thể sử dụng AI để lọc ra các văn bản từ các tệp ebook, các tệp PDF tổng hợp tài liệu câu hỏi, lý thuyết. Từ đó làm tài nguyên cho hệ thống để có thể sinh ra các tệp H5P tự động, không mất công sức thêm thủ công bằng tay, nhằm giúp giáo viên có kho đề tài tương tác chất lượng cao, phù hợp với nhu cầu giảng dạy. Tiếp theo có thể xây dựng mô hình AI dựa vào khả năng hoàn thành bài giảng của người học, mà đề xuất ra các nội dung phù hợp với người học.

# TÀI LIỆU THAM KHẢO

**[1]. Cấu trúc tệp H5P:** [**https://H5P.org/specification/**](https://h5p.org/specification/)

**[2]. Định nghĩa gói H5P:** [**https://H5P.org/documentation/developers/json-file-definitions**](https://h5p.org/documentation/developers/json-file-definitions)

**[3]. Phần mở rộng tệp được phép:** [**https://H5P.org/allowed-file-extensions**](https://h5p.org/allowed-file-extensions)

**[4]. Python tao tác với tệp PowerPoint:** [**https://python-pptx.readthedocs.io/en/latest/**](https://python-pptx.readthedocs.io/en/latest/)

**[5]. Các nội dung, cấu trúc tệp PowerPoint:** [**https://learn.microsoft.com/en-us/office/open-xml/open-xml-sdk**](https://learn.microsoft.com/en-us/office/open-xml/open-xml-sdk)