

# Ứng dụng Hỗ trợ đọc tài liệu học thuật cho Người Khiếm Thị

Lê Duy Anh, Lê Nguyễn Minh Châu  
Nguyễn Tấn Hoàng, Huỳnh Cao Tuấn Kiệt  
Đinh Điền, Nguyễn Hồng Bửu Long

ldanh22@clc.fitus.edu.com, lnmchau22@clc.fitus.edu.com  
nthoang22@clc.fitus.edu.com, hctkiet22@clc.fitus.edu.com  
ddien@fit.hcmus.edu.vn, nhblong@fit.hcmus.edu.vn

*Trường Đại học Khoa học Tự nhiên, Đại học Quốc gia TP.HCM*

## Tóm tắt nội dung

Trong thế giới ngày nay, người khiếm thị gặp nhiều khó khăn trong việc tiếp cận sách học thuật, đặc biệt là ghi nhớ nội dung và quản lý thông tin trong các tài liệu ấy. Giải pháp này đề xuất một ứng dụng hỗ trợ thông minh tích hợp các tính năng: tóm tắt nội dung, tìm kiếm bằng giọng nói, và quản lý ghi chú hiệu quả thông qua việc sử dụng các công nghệ xử lý ngôn ngữ tự nhiên (NLP) và học máy (Machine Learning).

## 1 Đặt vấn đề

Người khiếm thị có niềm yêu thích lớn đối với việc đọc sách. Tuy nhiên, theo một thống kê từ Thư viện Quốc gia Việt Nam, phần lớn họ không thể duy trì thói quen này do nhiều rào cản như khó khăn trong việc di chuyển, thiếu tài liệu phù hợp, v.v., trong khi chỉ 1,9% số người được khảo sát cho biết không có hứng thú với việc đọc sách [1].

Trong bối cảnh đó, người khiếm thị phải đối mặt với nhiều khó khăn khi tiếp cận tài liệu học thuật. Một trong những thách thức lớn là sự thiếu vắng các nền tảng hỗ trợ chuyên biệt, khiến việc tiếp cận sách và tài liệu học tập trở nên hạn chế. Hơn nữa, việc ghi nhớ các nội dung dài qua hình thức âm thanh cũng gặp nhiều trở ngại [2]. Thêm vào đó, khả năng tạo và quản lý

ghi chú hiệu quả của người khiếm thị còn bị giới hạn, điều này làm cho việc học lâu dài trở nên khó khăn. Một số thiết bị đọc sách hiện có còn gặp khó khăn về tốc độ đọc, không linh hoạt, từ đó gây cản trở trong việc tiếp thu và hiểu các nội dung học thuật [2].

Người khiếm thị cũng phải đối mặt với khó khăn trong việc đọc đồng thời nhiều tài liệu học thuật. Khác với một người bình thường có thể dễ dàng đọc lướt qua nhiều nguồn thông tin, người khiếm thị buộc phải tiếp cận thông tin theo trình tự, dòng sau dòng, làm giảm đáng kể hiệu quả tiếp cận tri thức.

Đặc biệt, trong các lĩnh vực khoa học, công nghệ, kỹ thuật và toán học (STEM), người khiếm thị gặp nhiều trở ngại trong việc tiếp cận và hiểu các yếu tố trực quan như biểu đồ, bảng số liệu và biểu thức toán học. Những yếu tố này thường đòi hỏi khả năng quan sát thị giác và sự hỗ trợ đặc thù để chuyển đổi sang dạng dễ tiếp cận.

Do đó, việc nghiên cứu và phát triển một ứng dụng hỗ trợ thông minh là cấp thiết, nhằm đáp ứng nhu cầu học tập của người khiếm thị. Ứng dụng này cần giúp họ tiếp cận tri thức một cách dễ dàng hơn, đồng thời cải thiện trải nghiệm học tập thông qua các công nghệ hiện đại, góp phần xóa bỏ những rào cản hiện tại và nâng cao chất lượng học tập, nghiên cứu của nhóm đối tượng này.

## 2 Những nghiên cứu liên quan trước đó

Với sự phát triển mạnh mẽ của trí tuệ nhân tạo nói chung và học sâu nói riêng, rất nhiều giải pháp đã được nghiên cứu và áp dụng nhằm nâng cao chất lượng cuộc sống của những người khiếm thị. Một nghiên cứu đã đề xuất sử dụng mạng nơ-ron tích chập (CNN) để trích xuất các đặc trưng quan trọng từ hình ảnh chứa chữ viết. Sau đó, mạng Long Short-Term Memory (LSTM) được sử dụng để tạo ra mô tả văn bản cho các hình ảnh đã được phân tích. Mô tả văn bản này sau đó được chuyển đổi thành giọng nói thông qua API Text-To-Speech. Hệ thống này không chỉ hỗ trợ người khiếm thị trong việc nhận diện chữ viết, mà còn cung cấp khả năng hiểu nội dung các hình ảnh trong tài liệu in, từ đó giúp người khiếm thị tiếp cận thông tin một cách hiệu quả hơn [3]. Mặc dù các mô hình NLP hiện tại rất hiệu quả trong phân tích văn bản, nhưng chúng chủ yếu được thiết kế cho văn bản thông thường, vẫn không nhiều ứng dụng tối ưu cho sách học thuật.

### 3 Cách tiếp cận của giải pháp

Giải pháp đề xuất nhằm hỗ trợ người khiếm thị thông qua các tính năng tương tác giọng nói toàn diện, kết hợp với công nghệ xử lý ngôn ngữ tự nhiên (NLP) và các mô hình trí tuệ nhân tạo tiên tiến. Các tính năng chính được đề xuất để giải quyết những thách thức hiện tại như sau:

- **Hỗ trợ ghi nhớ nội dung:** Ứng dụng tích hợp mô hình hỗ trợ Question Answering, giúp trả lời các câu hỏi của người dùng dựa trên nội dung có sẵn trong tài liệu. Mô hình này có khả năng hiểu ngôn ngữ tự nhiên, trích xuất và cung cấp thông tin một cách chính xác, phù hợp. Ví dụ, ứng dụng có thể trả lời về nội dung trong các trang sách hoặc chương đã đọc trước đó. Ngoài ra, phần mềm còn hỗ trợ người dùng tiếp cận nội dung một cách chủ động bằng cách cung cấp các bản tóm tắt nội dung của chương, đoạn hoặc trang sách. Điều này giúp nâng cao hiệu quả xử lý thông tin trong quá trình học tập, đọc tài liệu của người khiếm thị.
- **Hỗ trợ ghi chú:** Sử dụng xử lý ngôn ngữ tự nhiên (NLP) để tự động tóm tắt và phân loại ghi chú dựa trên nội dung, giúp tổ chức thông tin một cách hiệu quả. Người dùng có thể dễ dàng truy vấn thông tin từ ghi chú chỉ bằng cách sử dụng từ khóa. Ngoài ra công cụ chuyển đổi từ giọng nói thành văn bản (STT) và từ văn bản thành giọng nói (TTS) cho phép người khiếm thị tạo ghi chú nhanh chóng thông qua giọng nói, sau đó tự động chuyển đổi thành văn bản để lưu trữ. Đồng thời, ứng dụng có thể đọc lại các ghi chú hoặc trả lời các câu hỏi liên quan đến nội dung đã ghi.
- **Cải thiện khả năng tiếp cận:** Hệ thống này được thiết kế nhằm giúp người khiếm thị tiếp cận lượng lớn tài liệu số một cách dễ dàng mà không yêu cầu định dạng đặc thù, từ đó tạo điều kiện thuận lợi cho việc đọc và nghiên cứu tài liệu. Bên cạnh đó, hệ thống tích hợp tính năng tìm kiếm thông minh, kết hợp với hỗ trợ giọng nói, giúp người khiếm thị tra cứu tài liệu nhanh chóng mà không cần sử dụng phương pháp truyền thống. Thêm vào đó, hệ thống còn hỗ trợ tra cứu thông tin từ các nguồn bên ngoài thông qua các mô hình ngôn ngữ lớn (LLMs), với khả năng tổng hợp và cung cấp thông tin từ nhiều nguồn tài liệu khác nhau, mang lại trải nghiệm tra cứu phong phú và tiện lợi hơn.
- **Xử lý các yếu tố trực quan:** Ứng dụng này sử dụng các mô hình hỗ trợ miêu tả hình ảnh dựa trên ngữ cảnh, giúp chuyển đổi các yếu tố trực quan trong tài liệu thành văn bản để tiếp cận. Bên cạnh đó,

ứng dụng còn hỗ trợ miêu tả bảng biểu, biểu đồ và các nội dung trực quan khác dưới dạng văn bản, từ đó giúp người khiếm thị dễ dàng nắm bắt thông tin phức tạp trong các lĩnh vực như khoa học, công nghệ, kỹ thuật và toán học (STEM). Nhờ vào tính năng này, người khiếm thị có thể tiếp cận và hiểu rõ các dữ liệu hình ảnh một cách thuận lợi và hiệu quả hơn.

## 4 Kết quả mong muốn đạt được

- **Nâng cao chất lượng cuộc sống:** Giúp người khiếm thị tiếp cận tri thức dễ dàng hơn, từ đó cải thiện khả năng học hỏi và phát triển bản thân.
- **Mở rộng cơ hội:** Mang lại những cơ hội mới cho người khiếm thị trong các lĩnh vực học thuật, giúp họ tham gia vào các hoạt động nghiên cứu và học tập một cách bình đẳng.
- **Cải thiện khả năng tự học:** Hỗ trợ người khiếm thị phát triển khả năng tự học thông qua việc sử dụng các công cụ học tập hiện đại.

## Tài liệu

- [1] ThS. Trần Thị Thanh Vân, “Nghiên cứu nhu cầu tin của người khiếm thị nhằm mở rộng công tác phục vụ người dùng tin tại các cơ quan thông tin, thư viện Việt Nam,” ĐH KHXH&NV, Đại học Quốc gia Hà Nội, 2018. Nghiên cứu nhu cầu tin của người khiếm thị
- [2] R. M. Jackson, *Audio-Supported Reading for Students Who Are Blind or Visually Impaired*, Wakefield, MA: National Center on Accessible Educational Materials, Jan. 2021. [Online]. Available: <https://aem.cast.org/binaries/content/assets/common/publications/aem/asr-blind-visuallyimpaired.pdf>.
- [3] Basit Qureshi *et al.*, “Deep Learning for Image Captioning and Text-to-Speech Systems,” *Electronics*, vol. 11, no. 20, p. 3335, Oct. 2022. [Online]. Available: <https://doi.org/10.3390/electronics11203335>.