

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG



Phí Đăng Khoa

ĐỀ ÁN THẠC SĨ KỸ THUẬT

(Theo định hướng ứng dụng đối với đề án tốt nghiệp)

HÀ NỘI - 2025

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG



Phí Đăng Khoa

NGHIÊN CỨU CÔNG CỤ TÓM TẮT VĂN BẢN VỀ MÔ TẢ LỖI KHI KIỂM THỬ KHẢ NĂNG TIẾP CẬN WEB

CHUYÊN NGÀNH : HỆ THỐNG THÔNG TIN

MÃ SỐ : 8.48.01.04 (Hệ thống thông tin)

ĐỀ ÁN THẠC SĨ KỸ THUẬT

(Theo định hướng ứng dụng)

NGƯỜI HƯỚNG DẪN KHOA HỌC:

1. PGS.TS ĐỖ TRUNG TUẤN

HÀ NỘI - 2025

LỜI CAM ĐOAN

Tôi xin cam đoan rằng nội dung trong đề án này là công trình nghiên cứu khoa học do chính tôi thực hiện. Đề án chưa từng được nộp hoặc sử dụng để bảo vệ học vị tại bất kỳ cơ sở giáo dục đại học hoặc sau đại học nào trước đây. Tất cả các tài liệu tham khảo được sử dụng trong đề án đều được trích dẫn rõ ràng và đầy đủ theo quy định.

Các kết quả nghiên cứu và thực nghiệm trình bày trong đề án là do tôi trực tiếp thực hiện dưới sự hướng dẫn của PGS.TS Đỗ Trung Tuấn. Dữ liệu và kết quả phân tích trong đề án được thu thập từ quá trình triển khai thực tế và tiến hành thống kê độc lập.

Trong quá trình thực hiện đề án, tôi có sử dụng một số công cụ trợ lý trí tuệ nhân tạo nhằm hỗ trợ tìm kiếm thông tin, kiểm tra chính tả, ngữ pháp, cải thiện cách diễn đạt, cũng như rà soát tính chuyên nghiệp trong văn bản. Tuy nhiên, toàn bộ nội dung phân tích, đánh giá và kết luận trong đề án là do tôi chủ động nghiên cứu và chịu trách nhiệm.

Tôi xin hoàn toàn chịu trách nhiệm về tính trung thực và nguyên bản của đề án này.

LỜI CẢM ƠN

Trước hết, tôi xin gửi lời cảm ơn chân thành đến các thầy cô giáo tại Học viện Công nghệ Bưu chính Viễn thông – những người đã tận tình giảng dạy, truyền đạt kiến thức và hướng dẫn tôi trong suốt quá trình học tập và nghiên cứu tại trường.

Đặc biệt, tôi xin bày tỏ lòng biết ơn sâu sắc và trân trọng đến PGS.TS Đỗ Trung Tuấn – người thầy không chỉ tận tâm hướng dẫn chuyên môn, định hướng đề tài mà còn là người đã khơi dậy trong tôi niềm đam mê nghiên cứu khoa học và tinh thần không ngừng học hỏi. Thầy luôn dành thời gian chia sẻ, giảng giải những kiến thức quý giá về khoa học, xã hội và nghệ thuật, giúp tôi không chỉ hoàn thiện đề án mà còn mở rộng tầm nhìn, nuôi dưỡng khát vọng ứng dụng công nghệ để giải quyết các vấn đề thực tiễn của xã hội.

Tôi cũng xin gửi lời cảm ơn sâu sắc tới PGS.TS Hoàng Xuân Dậu – người thầy đã định hướng tôi trên con đường phát triển học thuật và nghiên cứu. Sự định hướng và niềm tin mà thầy dành cho tôi là nguồn động viên to lớn để tôi quyết tâm theo đuổi con đường nghiên cứu và học tập này.

Bên cạnh đó, tôi xin cảm ơn vợ và con gái nhỏ là chỗ dựa tinh thần vững chắc của tôi trong suốt quá trình học tập và nghiên cứu. Giữa những áp lực từ việc vừa đi làm vừa học cao học, sự cảm thông và hỗ trợ từ gia đình đã tiếp thêm cho tôi sức mạnh để kiên trì theo đuổi hành trình này..

Mặc dù đã dành nhiều thời gian và tâm huyết để hoàn thiện đề án một cách toàn diện nhất có thể, tôi nhận thức rằng đề án vẫn còn những thiếu sót và hạn chế nhất định. Tôi rất mong nhận được những ý kiến đóng góp quý báu từ các thầy, cô để đề án có thể được hoàn thiện và phát triển hơn nữa.

Xin chân thành cảm ơn!

MỤC LỤC

LỜI CAM ĐOAN	i
LỜI CẢM ƠN	ii
MỤC LỤC	iii
Danh mục các ký hiệu, các chữ viết tắt	v
Danh mục các bảng	vi
Danh mục các hình	vii
MỞ ĐẦU	1
Lý do chọn đề tài	1
Mục tiêu, đối tượng và phạm vi nghiên cứu	2
Ý nghĩa khoa học và thực tiễn của đề tài	2
Cấu trúc đề án	3
CHƯƠNG I KHẢ NĂNG TIẾP CẬN WEB	5
1.1. Khả năng tiếp cận	5
1.2. Khả năng tiếp cận web	7
1.3. Kiểm thử khả năng tiếp cận web theo WCAG	9
1.4. Kết luận	11
CHƯƠNG II KỸ THUẬT TÓM TẮT VĂN BẢN	12
2.1. Tóm tắt văn bản	12
2.1.1. Khái niệm tóm tắt văn bản	12
2.1.2. Phân nhóm kỹ thuật tóm tắt văn bản	13
2.1.3. Kỹ thuật tóm tắt văn bản truyền thống	13
2.2. Tóm tắt văn bản dựa trên mô hình ngôn ngữ lớn (LLM)	15
2.2.1. Sự phát triển của mô hình ngôn ngữ lớn (LLM)	15
2.2.2. So sánh giữa kỹ thuật truyền thống và LLM	16
2.3. Chọn mô hình GPT để triển khai trong nghiên cứu	19
2.3.1. Lý do lựa chọn GPT	19
2.3.2. Minh chứng từ dự án thực tế	20
2.4. Kết luận	21

CHƯƠNG III THỰC NGHIỆM TÓM TẮT VĂN BẢN VỀ MÔ TẢ LỖI	23
3.1. Giới thiệu dự án đánh giá tính tiếp cận của của 20 cổng thông tin điện tử	23
3.1.1. Giới thiệu	23
3.1.2. Kết quả của dự án	26
3.2. Bộ dữ liệu để thực nghiệm trong nghiên cứu	27
3.2.1. Quy trình thu thập dữ liệu gốc	27
3.2.2. Mô tả và thống kê bộ dữ liệu gốc	30
3.2.3. Chuẩn bị bộ dữ liệu cho thực nghiệm	32
3.2.4. Thách thức tóm tắt thủ công	33
3.3. Áp dụng kỹ thuật tóm tắt văn bản bằng GPT trong tổng hợp mô tả lỗi	34
3.3.1. Phát triển quy trình kết hợp công cụ tóm tắt	34
3.3.2. Phát triển công cụ tóm tắt	36
3.3.3. Phương pháp đánh giá	43
3.3.4. Thực hiện đánh giá và kết quả	46
3.4 Kết luận	48
KẾT LUẬN VÀ KIẾN NGHỊ	49
Kết quả đã đạt được	49
Hạn chế và hướng phát triển	49
DANH MỤC TÀI LIỆU THAM KHẢO	51
PHỤ LỤC	53
Phụ lục 01: Mã nguồn công cụ tóm tắt văn bản mô tả lỗi dựa trên GPT	53
Phụ lục 02: Các bản tóm tắt và kết quả chấm điểm đầu ra công cụ tóm tắt dựa trên GPT	56
Phụ lục 03: Bộ dữ liệu gốc thu thập từ dự án UNDP	58

Danh mục các ký hiệu, các chữ viết tắt

Viết tắt	Tiếng Anh	Giải thích
PTIT	Posts and Telecommunications Institute of Technology	Học viện Công nghệ Bưu chính Viễn thông
WCAG	Web Content Accessibility Guidelines	Bộ hướng dẫn khả năng tiếp cận web
LSA	Latent Semantic Analysis	Phân tích ngữ nghĩa tiềm ẩn
BERT	Bidirectional Encoder Representations from Transformers	Mô hình mã hóa hai chiều hiểu ngữ cảnh từ cả hai phía
BART	Bidirectional and Auto- Regressive Transformers	Mô hình kết hợp mã hóa hai chiều và giải mã tự hồi quy
LLM	Large language model	Mô hình ngôn ngữ lớn
NLP	Natural Language Processing	Xử lý ngôn ngữ tự nhiên, là một lĩnh vực của AI tập trung vào việc giúp máy tính hiểu, diễn giải và tạo ra ngôn ngữ của con người
GPT	Generative Pre-trained Transformer	Là một mô hình ngôn ngữ lớn phát triển bởi OpenAI
Prompt	Prompt	Là đoạn nội dung đầu vào cung cấp cho một mô hình LLM như GPT, để nó hiểu yêu cầu và sinh ra câu trả lời mong muốn.
RAG	Retrieval-Augmented Generation	Kỹ thuật kết hợp giữa khả năng truy xuất thông tin từ các nguồn dữ liệu bên ngoài với sức mạnh tạo nội dung của mô hình LLM
ERP	Enterprise Resource Planning	Hệ thống hoạch định nguồn lực doanh nghiệp

Danh mục các bảng

Bảng I.1: Lịch sử các phiên bản chính của WCAG.....	9
Bảng II.1: So sánh kỹ thuật tóm tắt văn bản truyền thống và LLM	16
Bảng III.1: Danh sách 20 công thông tin được đánh giá trong dự án.....	24
Bảng III.2: Cấu trúc bảng và mẫu của 1 bản ghi trong bộ dữ liệu gốc từ file Excel chứa bộ dữ liệu kết quả đánh giá từ nhiều cộng tác viên.....	30
Bảng III.3: Cấu trúc bảng và mẫu của 1 bản ghi trong bộ dữ liệu chuẩn bị cho thực nghiệm.....	33
Bảng III.4: Ba cách sinh tóm tắt với GPT-4	38
Bảng III.5: Hướng dẫn chấm điểm thủ công kết quả tóm tắt bởi công cụ.....	44

Danh mục các hình

Hình I.1: Người cao tuổi gặp khó khăn tiếp cận nội dung trên điện thoại thông minh	5
Hình I.2: Tính năng Narrator có sẵn trong hệ điều hành Windows là "công cụ đọc màn hình" cho người khiếm thị	7
Hình I.3: Một người khiếm thị sử dụng "công cụ đọc màn hình" để đọc nội dung trên website	8
Hình II.1 OpenAI hướng dẫn kỹ thuật prompt trên mô hình GPT 4.1	16
Hình II.2: Giao diện sử dụng GPT 4.1 trên Microsoft Azure AI Portal	20
Hình II.3: Giao diện cấu hình GPT trên Microsoft Copilot Studio để thử nghiệm phát triển ứng dụng nhanh	21
Hình III.1 Bài đăng về buổi họp kỹ thuật giữa UNDP & Hội người mù Việt Nam "Tiếp cận cổng thông tin điện tử của các cơ quan nhà nước – nâng cao cơ hội hòa nhập cho NKT tại Việt Nam "	23
Hình III.2 Bài viết về bộ WCAG 1.0 trên website của W3C	25
Hình III.3: Giao diện Google Form đã tạo để cộng tác viên nhập kết quả đánh giá bởi cá nhân	29
Hình III.4: File excel File Excel dữ liệu thô, chứa toàn bộ kết quả kiểm thử, riêng lẻ theo từng cộng tác viên	29
Hình III.5: Phân bố bản ghi theo thể loại và nhóm kiểm thử	31
Hình III.6: Quy trình kiểm thử khả năng tiếp cận kết hợp Chuyên gia - Cộng đồng cộng tác viên người khiếm thị - Công cụ tóm tắt dựa trên GPT	35
Hình III.7 Kiến trúc giải pháp của Công cụ tóm tắt	36
Hình III.8 Bảng giá API mô hình GPT 4.1 từ Azure OpenAI	38

MỞ ĐẦU

Lý do chọn đề tài

Trong bối cảnh chuyển đổi số và bình dân học vụ số mạnh mẽ tại Việt Nam, khả năng tiếp cận của các ứng dụng số vẫn còn nhiều hạn chế, đặc biệt đối với người khuyết tật và người cao tuổi. Một đánh giá vào năm 2023 của Viện Nghiên cứu Chính sách và Phát triển Truyền thông (IPS) phối hợp cùng Chương trình Phát triển Liên Hợp Quốc (UNDP) cho thấy, trong số 63 cổng dịch vụ công trực tuyến cấp tỉnh được khảo sát, **không có cổng nào đáp ứng đầy đủ các tiêu chí cơ bản về khả năng tiếp cận**. Ngoài ra, một nghiên cứu công bố năm 2025 về khả năng tiếp cận số tại các quốc gia thuộc Nam Bán Cầu chỉ ra rằng **Việt Nam xếp thứ hai về số lượng lỗi khả năng tiếp cận** trung bình trên mỗi website, chỉ sau Bangladesh [1]. Việt Nam hiện chưa có quy định pháp lý cụ thể về khả năng tiếp cận, chỉ mới đề cập rải rác trong “Luật Người khuyết tật năm 2010”, “Chương trình Chuyển đổi số quốc gia đến năm 2025, định hướng đến năm 2030”, và Thông tư số 32/2017/TT-BTTTT [2].

Để cải thiện thực trạng này, cần tăng cường chính sách, nâng cao nhận thức của nhà phát triển và thực hiện kiểm thử thường xuyên để phát hiện lỗi. Tuy nhiên, thực tế cho thấy các **công cụ kiểm thử tự động hiện nay chỉ phát hiện được khoảng 30–40% lỗi**, phần lớn lỗi còn lại – đặc biệt là các lỗi về ngữ nghĩa, bối cảnh và trải nghiệm thực tế – **chỉ có thể được phát hiện thông qua kiểm thử thủ công** bởi con người, cụ thể là người dùng thực tế như người khiếm thị.

Tuy vậy, quá trình kiểm thử thủ công tại Việt Nam vẫn gặp nhiều hạn chế. Các kiểm thử viên chuyên nghiệp thường không phải là người khuyết tật nên khó phát hiện được các lỗi thực sự ảnh hưởng đến người dùng khuyết tật. Trong khi đó, **người khiếm thị lại thiếu kỹ năng kiểm thử chuyên môn**, chưa được đào tạo bài bản, dẫn đến việc kiểm thử không thể diễn ra thường xuyên và thiếu tính hệ thống. Ngoài ra, tại Việt Nam hiện nay, cơ hội nghề nghiệp và đào tạo cho người khuyết tật tham gia kiểm thử phần mềm gần như không có, họ thường chỉ được tham gia dưới dạng cộng tác viên ngắn hạn trong các dự án thử nghiệm.

Trước những thách thức đó, việc nghiên cứu và phát triển các giải pháp hỗ trợ kiểm thử khả năng tiếp cận, đặc biệt là phù hợp với kiểm thử viên là người khiếm thị, trở thành nhu cầu cấp thiết và mang ý nghĩa thực tiễn sâu sắc.

Mục tiêu, đối tượng và phạm vi nghiên cứu

Mục tiêu của đề án là tập trung nghiên cứu và phát triển **một quy trình kiểm thử khả năng truy cập phù hợp cho người khiếm thị và một công cụ hỗ trợ tóm tắt văn bản mô tả lỗi** kiểm thử từ nhiều cộng tác viên. Kết quả đầu ra là một quy trình kiểm thử có tính thực tiễn cao, dễ áp dụng, phù hợp với nguồn lực hạn chế của cộng tác viên là người khiếm thị, cùng với một công cụ tự động tổng hợp báo cáo lỗi từ nhiều kết quả kiểm thử. Giải pháp này nhằm hỗ trợ đơn vị chủ quản website dễ dàng tiếp nhận báo cáo, từ đó thực hiện các cải thiện cần thiết để nâng cao khả năng tiếp cận cho người khuyết tật.

Đối tượng nghiên cứu là các kết quả kiểm thử khả năng truy cập do cộng tác viên khiếm thị thực hiện, dựa trên hướng dẫn Web Content Accessibility Guidelines (WCAG), tập trung vào các lỗi ảnh hưởng đến khả năng tiếp cận thông tin và trải nghiệm người dùng trên website.

Phạm vi nghiên cứu là tập trung vào hoạt động kiểm thử và tổng hợp hơn **3600 kết quả kiểm thử khả năng truy cập** từ các cộng tác viên khiếm thị, của **20 website công thông tin** (xem danh sách tại Bảng III.1) trong dự án hợp tác giữa Chương trình Phát triển Liên Hợp Quốc (UNDP) và Hội Người mù Việt Nam.

Ý nghĩa khoa học và thực tiễn của đề tài

Đề tài có ý nghĩa khoa học khi góp phần bổ sung vào nghiên cứu về **thúc đẩy khả năng tiếp cận số tại Việt Nam**, đặc biệt tập trung vào vai trò của người khuyết tật – cụ thể là người khiếm thị – trong việc kiểm thử khả năng truy cập các website dịch vụ công. Việc phát triển quy trình kiểm thử khả năng truy cập chuyên biệt dành cho cộng tác viên khiếm thị, cùng với công cụ hỗ trợ tự động tổng hợp báo cáo lỗi, là một hướng tiếp cận mới. Hướng tiếp cận này không chỉ giúp phát huy vai trò của người khuyết tật trong quá trình phát hiện lỗi mà còn tận dụng chính trải nghiệm thực tế của họ để phát hiện những vấn đề mà công cụ tự động không thể

phát hiện được, từ đó nâng cao chất lượng kiểm thử và tăng tính ứng dụng trong thực tế.

Về mặt thực tiễn, đề tài mang lại giá trị ứng dụng rõ rệt khi **giúp các đơn vị quản lý website dịch vụ công dễ dàng tiếp nhận các báo cáo lỗi khả năng tiếp cận** được tổng hợp rõ ràng, mạch lạc và dễ hiểu, từ đó hỗ trợ quá trình cải thiện khả năng tiếp cận của website một cách hiệu quả. Bên cạnh đó, giải pháp được đề xuất phù hợp với bối cảnh nguồn lực hạn chế của các dự án có sự tham gia của cộng tác viên khiếm thị, giúp tiết kiệm thời gian, công sức trong quá trình tổng hợp và báo cáo. Ngoài ra, đề tài còn góp phần **thúc đẩy sự tham gia và tạo cơ hội nghề nghiệp cho người khuyết tật trong lĩnh vực kiểm thử phần mềm** – một lĩnh vực hiện nay còn rất ít chương trình đào tạo chính quy và cơ hội việc làm dành riêng cho họ tại Việt Nam.

Cấu trúc đề án

Đề án được tổ chức thành 5 phần với các nội dung chính sau:

Phần Mở đầu giới thiệu lý do chọn đề tài, mục tiêu, đối tượng và phạm vi nghiên cứu, ý nghĩa khoa học và thực tiễn của đề tài. Phần này làm rõ động lực nghiên cứu, các thách thức thực tiễn trong kiểm thử khả năng tiếp cận web và nhu cầu phát triển công cụ tóm tắt hỗ trợ cộng tác viên khiếm thị trong bối cảnh tại Việt Nam.

Chương I giới thiệu các khái niệm nền tảng về khả năng tiếp cận số và khả năng tiếp cận web, phân tích thực trạng và thách thức kiểm thử khả năng tiếp cận tại Việt Nam, giới thiệu các tiêu chuẩn kiểm thử phổ biến (WCAG) và vai trò của người khiếm thị trong quy trình kiểm thử.

Chương II tập trung vào các kỹ thuật tóm tắt văn bản, từ các phương pháp truyền thống như TextRank, LSA, BERT, BART đến các phương pháp hiện đại dựa trên mô hình ngôn ngữ lớn (LLM) như GPT. Phần này cũng so sánh các kỹ thuật, phân tích lý do lựa chọn GPT làm nền tảng triển khai trong nghiên cứu.

Chương III trình bày quá trình thực nghiệm kiểm thử khả năng tiếp cận web với 20 cổng thông tin điện tử. Đề xuất quy trình và tích hợp công cụ tóm tắt vào hệ thống kiểm thử thực tế, nhằm giảm tải công việc thủ công và nâng cao chất lượng

báo cáo. Phần này cũng trình bày phương pháp đánh giá chất lượng tóm tắt và kết quả thực nghiệm.

Phần Kết luận tổng kết các kết quả đạt được, nêu bật đóng góp chính của nghiên cứu, đồng thời chỉ ra các hạn chế và đề xuất hướng phát triển trong tương lai như mở rộng công cụ cho các nền tảng khác ngoài web, cải thiện độ chính xác bằng kỹ thuật RAG, và ứng dụng trong các lĩnh vực kiểm thử khả năng tiếp cận khác.

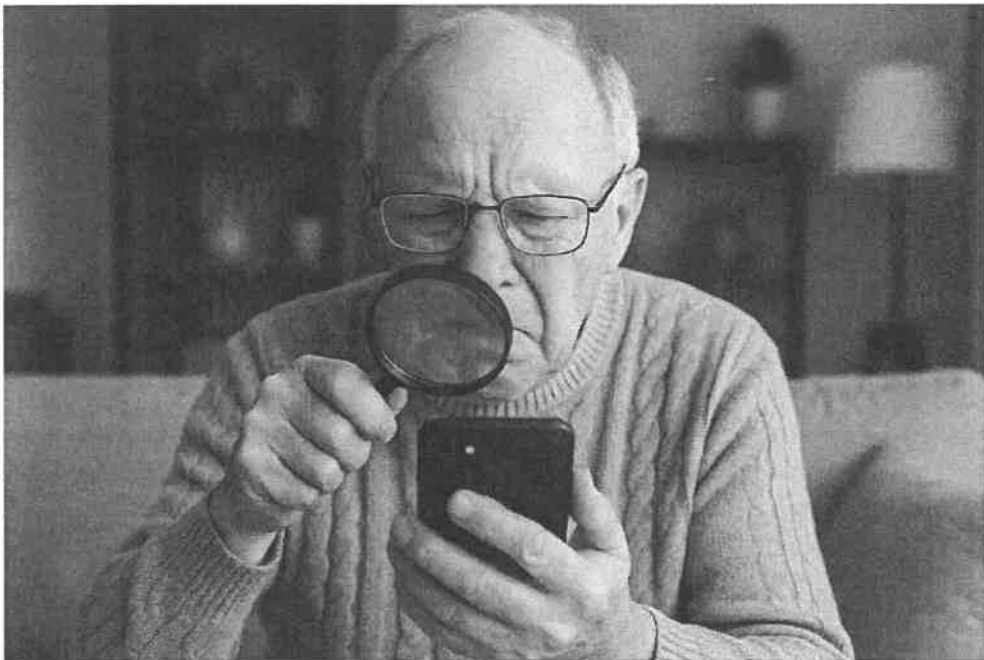
CHƯƠNG I

KHẢ NĂNG TIẾP CẬN WEB

1.1. Khả năng tiếp cận

Khả năng tiếp cận (**accessibility**) là khái niệm đề cập đến mức độ mà một hệ thống, sản phẩm, dịch vụ hoặc môi trường có thể được sử dụng bởi tất cả mọi người, bao gồm cả những người có khiếm khuyết về thể chất, giác quan, nhận thức hoặc điều kiện sử dụng đặc biệt. Trong thế giới số ngày nay, khả năng tiếp cận trở thành một yêu cầu thiết yếu, đảm bảo mọi người đều có thể tiếp cận thông tin, thực hiện tác vụ và tham gia đầy đủ vào xã hội số.

Khả năng tiếp cận không chỉ phục vụ người khuyết tật mà còn mang lại lợi ích cho nhiều nhóm người dùng khác như: người lớn tuổi, người làm việc trong môi trường hạn chế ánh sáng, không gian. Vì vậy, nhiều quốc gia và tập đoàn công nghệ đã xem đây là ưu tiên chiến lược và bắt buộc pháp lý trong thiết kế sản phẩm và dịch vụ.



Hình I.1: Người cao tuổi gặp khó khăn tiếp cận nội dung trên điện thoại thông minh

Tại các quốc gia phát triển như Mỹ và Nhật Bản, khả năng tiếp cận không chỉ được coi là một chuẩn mực đạo đức trong thiết kế và cung cấp dịch vụ, mà còn là

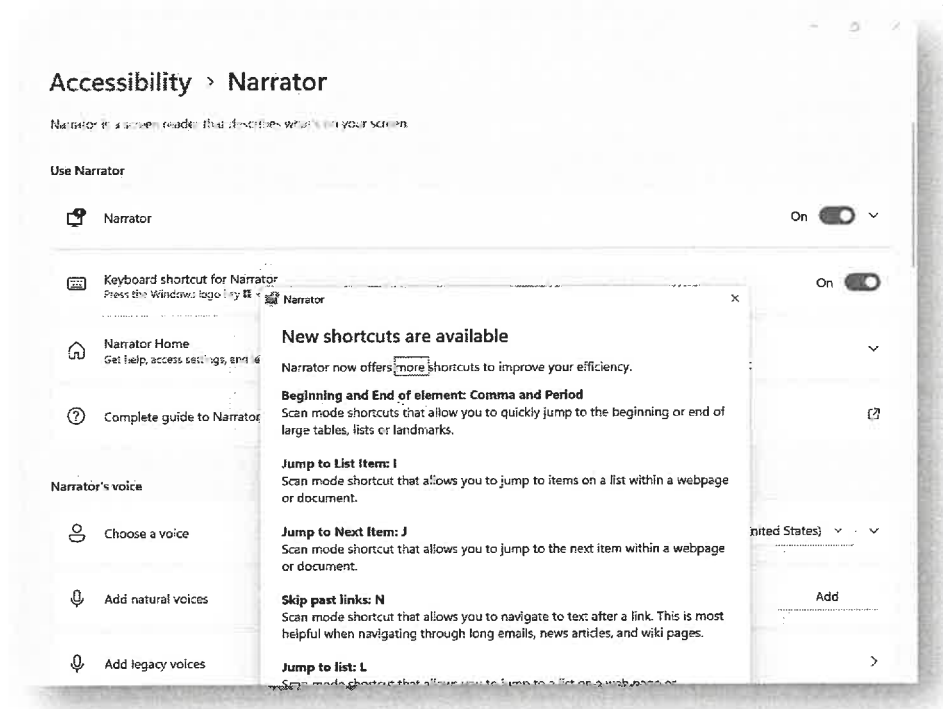
một nghĩa vụ pháp lý bắt buộc. Việc không tuân thủ có thể dẫn đến chế tài xử phạt, đặc biệt áp dụng đối với các tổ chức công quyền, cơ sở giáo dục, bệnh viện và các doanh nghiệp lớn:

- Mỹ: Là quốc gia đi đầu trong lĩnh vực này, Mỹ đã ban hành Đạo luật Người khuyết tật (ADA - Americans with Disabilities Act) từ năm 1990. Trong lĩnh vực công nghệ số, Section 508 thuộc Luật Phục hồi chức năng (Rehabilitation Act) yêu cầu tất cả các hệ thống công nghệ thông tin của cơ quan chính phủ phải đảm bảo khả năng tiếp cận cho người khuyết tật. Bộ tiêu chuẩn tham chiếu chính là WCAG 2.0 Level AA.
- Nhật Bản: Quốc gia này đã ban hành Luật Xúc tiến Chính sách Người khuyết tật (Act on Promotion of Measures for Persons with Disabilities) và có hướng dẫn kỹ thuật quốc gia JIS X 8341-3, vốn được nội địa hóa từ tiêu chuẩn WCAG. Các cơ quan nhà nước và doanh nghiệp đều được khuyến nghị tuân thủ để tạo ra môi trường số hòa nhập.

Tại các tập đoàn công nghệ hàng đầu như Microsoft, Apple, Google, Adobe, “khả năng tiếp cận” không còn là tính năng phụ mà đã trở thành một yêu cầu mặc định trong chiến lược phát triển sản phẩm:

- Microsoft: Xây dựng nguyên tắc "Inclusive Design" (Thiết kế toàn diện), cung cấp công cụ hỗ trợ như Narrator (trình đọc màn hình), Eye Control, Accessibility Insights (dành cho nhà phát triển kiểm thử tự động). Microsoft cũng phát hành AI for Accessibility Grant – quỹ tài trợ cho các dự án đổi mới sáng tạo vì người khuyết tật.
- Apple: Tích hợp sâu các tính năng như VoiceOver, Magnifier, Sound Recognition, AssistiveTouch trong toàn bộ hệ điều hành macOS, iOS. Apple tuyên bố rằng mọi sản phẩm của họ đều phải có khả năng truy cập cho người khuyết tật "ngay từ ngày đầu tiên".
- Google: Cung cấp TalkBack, Live Caption, Lookout, đồng thời xây dựng nền tảng Android Accessibility Suite. Ngoài ra, Google cũng phát triển các

công cụ mã nguồn mở để hỗ trợ cộng đồng nhà phát triển xây dựng sản phẩm dễ tiếp cận.



Hình 1.2: Tính năng Narrator có sẵn trong hệ điều hành Windows là "công cụ đọc màn hình" cho người khiếm thị

Tại Việt Nam, Luật Người khuyết tật 2010 [3] đã công nhận quyền được tiếp cận công nghệ của người khuyết tật. Tuy nhiên, trong thực tế, việc triển khai vẫn còn hạn chế do thiếu nguồn lực kỹ thuật, quy định cụ thể và nhận thức xã hội chưa cao.

1.2. Khả năng tiếp cận web

Khả năng tiếp cận web (**web accessibility**) là một phần mở rộng quan trọng và mang tính thực tiễn cao trong khái niệm tổng quát về khả năng tiếp cận đã trình bày ở mục trước. Nếu như khả năng tiếp cận đề cập đến việc tạo ra các sản phẩm và dịch vụ mà tất cả mọi người – bao gồm cả người khuyết tật – có thể sử dụng bình đẳng, thì khả năng tiếp cận web cụ thể hóa điều đó thông qua việc thiết kế và phát triển các website, ứng dụng và nội dung trực tuyến dễ tiếp cận, dễ hiểu và dễ thao tác cho mọi đối tượng người dùng.

Theo tổ chức W3C (World Wide Web Consortium), một website có khả năng tiếp cận tốt cần thỏa mãn các yếu tố POUR [4], viết tắt của:

- **Perceivable** (Có thể cảm nhận được): Hình ảnh có văn bản thay thế (alt text), video có phụ đề.
- **Operable** (Có thể vận hành được): Có thể điều khiển bằng bàn phím, chuột, giọng nói hoặc công nghệ hỗ trợ.
- **Understandable** (Có thể hiểu được): Cấu trúc rõ ràng, biểu mẫu được gắn nhãn đầy đủ.
- **Robust** (Ổn định, tương thích được): Hoạt động tốt với trình đọc màn hình, phần mềm phóng to, thiết bị hỗ trợ khác.



Hình 1.3: Một người khiếm thị sử dụng "công cụ đọc màn hình" để đọc nội dung trên website

Tại Việt Nam, khả năng tiếp cận web vẫn là một vấn đề tồn tại nhiều thách thức, đặc biệt trong khu vực công. Nhiều cổng thông tin điện tử của cơ quan nhà nước chưa đáp ứng được các tiêu chí tiếp cận cơ bản như công cụ tìm kiếm nội dung, tương thích trình đọc màn hình cho người khiếm thị [5].

1.3. Kiểm thử khả năng tiếp cận web theo WCAG

Khả năng tiếp cận web cần được đánh giá và kiểm thử một cách hệ thống, dựa trên các tiêu chuẩn quốc tế đã được công nhận, nhằm đảm bảo tính hiệu quả, đồng nhất và có thể cải thiện liên tục. Trong bối cảnh đó, Web Content Accessibility Guidelines (WCAG) [6] bộ hướng dẫn kỹ thuật do tổ chức World Wide Web Consortium (W3C) phát triển, đã trở thành tiêu chuẩn toàn cầu phổ biến nhất trong việc đánh giá và cải thiện khả năng tiếp cận của website và nội dung số.

Bảng I.1: Lịch sử các phiên bản chính của WCAG

Phiên bản	Năm phát hành	Mô tả
WCAG 1.0	1999	Là phiên bản đầu tiên, tập trung vào các quy tắc kỹ thuật cụ thể như gắn văn bản thay thế (alt text) cho hình ảnh hoặc sử dụng HTML hợp lệ.
WCAG 2.0	2008	Đưa ra bốn nguyên tắc cốt lõi POUR, độc lập nền tảng và trở thành tiêu chuẩn phổ biến nhất toàn cầu (được ISO công nhận năm 2012)
WCAG 2.1	2018	Bổ sung các tiêu chí hỗ trợ người dùng trên thiết bị di động và người rối loạn nhận thức.
WCAG 2.2	2023	Là phiên bản mới nhất tại thời điểm viết đề án, mở rộng tiêu chí để hỗ trợ tốt hơn người dùng có rối loạn chức năng nhận thức và sử dụng thiết bị cảm ứng

Dưới đây là thông tin về một tiêu chí cụ thể được sử dụng trong phần thực nghiệm của đề tài là **Page Titled 2.4.2 thuộc WCAG 2.1**. Có thể dễ nhận thấy việc đáp ứng tiêu chí 2.4.2 này không đơn giản về mặt ngôn ngữ lập trình, có thể <title> trong HTML, mà còn về mặt nội dung cần rõ ràng và đúng mục đích trang.

Mục tiêu	Mỗi trang web có một tiêu đề có ý nghĩa
-----------------	---

Cần làm gì	Cung cấp tiêu đề mô tả rõ ràng cho trang bằng công nghệ phù hợp
Tại sao lại quan trọng	Tiêu đề trang giúp người dùng nhận biết và phân biệt các trang khác nhau
Ví dụ	An HTML web page The descriptive title of an HTML web page is marked up with the <title> element so that it will be displayed in the title bar of the user agent. A document collection The title of Understanding WCAG 2.2 is "Understanding WCAG 2.2". • The Introduction to Understanding WCAG page has the title "Introduction to Understanding WCAG". • Major sections of the document collection are pages titled "Understanding Guideline X" and "Understanding Success Criterion X." • Appendix A has the title "Glossary." • Appendix B has the title "Acknowledgements." • Appendix C has the title "References." A web application A banking application lets users inspect their bank accounts, view past statements, and perform transactions. The web application dynamically generates titles for each web page, e.g., "Bank XYZ, accounts for Alex Smith" "Bank XYZ, December 2005 statement for Account 1234-5678".

Việc kiểm thử khả năng tiếp cận của website theo chuẩn WCAG có thể được thực hiện thông qua hai phương pháp chính là kiểm thử tự động và kiểm thử thủ công.

- Kiểm thử **tự động** sử dụng các công cụ như Axe Accessibility, Google Lighthouse, và Accessibility Insights để nhanh chóng phát hiện các lỗi kỹ

thuật phổ biến như thiếu văn bản thay thế (alt text) cho hình ảnh, tiêu đề trùng lặp, hoặc lỗi về độ tương phản màu sắc. Những công cụ này giúp tiết kiệm thời gian và công sức trong việc phát hiện các vấn đề cơ bản về khả năng tiếp cận.

- Tuy nhiên, kiểm thử **thủ công** vẫn đóng vai trò không thể thay thế trong việc phát hiện các lỗi phức tạp mà công cụ tự động không thể nhận diện. Ví dụ, **chỉ có con người mới có thể đánh giá** liệu văn bản thay thế cho hình ảnh có phù hợp với ngữ cảnh hay không, hoặc xác định xem trình tự điều hướng trên trang có hợp lý và dễ hiểu đối với người dùng sử dụng trình đọc màn hình hay không. Kiểm thử thủ công thường được thực hiện bởi các chuyên gia hoặc người dùng thực tế, đặc biệt là người khuyết tật. Chuyên gia về khả năng tiếp cận Marcy Sutton [7] đã nhấn mạnh rằng: *“Công cụ kiểm thử tự động là một khởi đầu tuyệt vời, nhưng chúng **chỉ phát hiện được khoảng 30% vấn đề**. Kiểm thử thủ công là cần thiết để đảm bảo khả năng tiếp cận đầy đủ.”*

Do đó, sự kết hợp giữa kiểm thử tự động và thủ công là phương pháp hiệu quả nhất để đảm bảo khả năng tiếp cận toàn diện cho một website.

1.4. Kết luận

Chương I đã giới thiệu những khái niệm cơ bản nhất về khả năng tiếp cận, từ khái niệm tổng quát đến khả năng tiếp cận trên website. Khả năng tiếp cận web không chỉ dành riêng cho người khuyết tật mà còn giúp ích cho nhiều đối tượng khác, như người lớn tuổi hay người dùng trong môi trường hạn chế.

Kiểm thử khả năng tiếp cận web, đặc biệt theo chuẩn WCAG, là phương pháp quan trọng để đảm bảo website thân thiện với mọi người dùng. Tuy nhiên, thực tế cho thấy các công cụ kiểm thử tự động hiện tại chưa thể thay thế hoàn toàn kiểm thử thủ công, đặc biệt là với các lỗi ngữ cảnh và tính dễ hiểu. Điều này nhấn mạnh vai trò thiết yếu của người kiểm thử là người khuyết tật, đồng thời đặt ra nhu cầu phát triển các công cụ hỗ trợ tổng hợp kết quả kiểm thử một cách tự động, chính xác và dễ sử dụng – đây cũng là hướng nghiên cứu chính của đề án.

CHƯƠNG II

KỸ THUẬT TÓM TẮT VĂN BẢN

2.1. Tóm tắt văn bản

2.1.1. Khái niệm tóm tắt văn bản

Tóm tắt văn bản là quá trình xử lý và rút gọn nội dung từ một hoặc nhiều tài liệu gốc thành một phiên bản ngắn gọn hơn, trong khi vẫn giữ lại các ý chính và thông tin quan trọng nhất của văn bản ban đầu. Mục tiêu chính của tóm tắt là giúp người đọc nắm bắt nhanh những thông tin cốt lõi mà không cần đọc toàn bộ nội dung chi tiết, đặc biệt hữu ích trong những trường hợp khối lượng dữ liệu lớn hoặc khi cần tiết kiệm thời gian tổng hợp.

Tóm tắt văn bản không chỉ đơn thuần là cắt gọn câu chữ, mà còn đòi hỏi khả năng hiểu và phân tích nội dung để xác định phần thông tin nào là quan trọng, phần nào có thể loại bỏ hoặc giản lược mà không làm mất ý nghĩa. Một bản tóm tắt chất lượng cần đảm bảo tính chính xác, phản ánh đầy đủ nội dung cốt lõi, đồng thời phải rõ ràng, súc tích và dễ đọc, ngay cả với người không chuyên.

Trong lĩnh vực kiểm thử phần mềm thủ công, hoặc cụ thể kiểm thử khả năng tiếp cận website, tóm tắt văn bản được áp dụng để tổng hợp các lỗi được báo cáo bởi nhiều kiểm thử viên. Ví dụ, một nhóm kiểm thử viên khiếm thị có thể thực hiện đánh giá độc lập cùng một trang web và ghi lại mô tả lỗi theo từng tiêu chí cụ thể trong chuẩn WCAG. Những mô tả này thường rất chi tiết, đa dạng về cách diễn đạt và có thể trùng lặp nội dung ở nhiều mức độ khác nhau. Việc tổng hợp các lỗi này thành một báo cáo tóm tắt rõ ràng, mạch lạc và dễ đọc là rất cần thiết để giúp các bên liên quan như quản trị viên website hoặc nhà phát triển dễ dàng hiểu và xử lý vấn đề. Tóm tắt mô tả lỗi trong bối cảnh này không chỉ giúp tiết kiệm thời gian tổng hợp thủ công mà còn đảm bảo các thông tin quan trọng được làm nổi bật, loại bỏ trùng lặp, và chuẩn hóa cách diễn đạt, từ đó nâng cao hiệu quả trao đổi thông tin trong dự án.

2.1.2. Phân nhóm kỹ thuật tóm tắt văn bản

Các kỹ thuật tóm tắt văn bản được chia thành hai nhóm chính là tóm tắt trích xuất và tóm tắt khái quát, cụ thể:

Tóm tắt trích xuất (Extractive Summarization): Phương pháp này chọn và trích xuất các câu hoặc cụm từ quan trọng nhất từ văn bản gốc để tạo thành bản tóm tắt. Nó không tạo nội dung mới mà chỉ chọn các câu, từ có sẵn. Một số phương pháp tiêu biểu bao gồm TextRank, Phân tích ngữ nghĩa tiềm ẩn (LSA) và các mô hình trích xuất dựa trên BERT. Ví dụ, trong quá trình tổng hợp lỗi từ nhiều kiểm thử viên khiếm thị về một website, tóm tắt trích xuất sẽ lấy trực tiếp các câu quan trọng như "Thiếu alt text cho hình ảnh logo", "Tiêu đề trang không rõ nghĩa" và ghép lại thành bản tóm tắt. Phương pháp này nhanh, dễ thực hiện, và hạn chế rủi ro diễn giải sai ý gốc, nhưng bản tóm tắt có thể rời rạc, thiếu mạch lạc và khó đọc.

Tóm tắt khái quát (Abstractive Summarization): Thay vì sao chép nguyên văn câu, phương pháp sáng tạo sẽ viết lại nội dung bằng ngôn ngữ mới, cô đọng và súc tích hơn. Phương pháp này thường cần đến các mô hình học sâu có khả năng hiểu cấu trúc ngôn ngữ, như BART, mô hình Transformer. Vẫn trong tình huống ví dụ bên trên, tóm tắt khái quát có thể diễn đạt lại thành câu như: "Website thiếu văn bản thay thế cho hình ảnh và tiêu đề trang không phản ánh nội dung, gây khó khăn cho người khiếm thị." Phương pháp này giúp tạo bản tóm tắt dễ đọc, rõ nghĩa hơn cho người không chuyên, nhưng đòi hỏi công cụ AI mạnh mẽ và cẩn thận kiểm soát đầu ra để tránh sai sót hoặc mất chi tiết quan trọng.

2.1.3. Kỹ thuật tóm tắt văn bản truyền thống

Trong lĩnh vực tóm tắt văn bản truyền thống, một số phương pháp nổi bật thường được áp dụng là TextRank, LSA, BERT thuộc nhóm tóm tắt trích xuất, và BART thuộc nhóm tóm tắt khái quát.

TextRank là một thuật toán dựa trên ý tưởng xếp hạng đồ thị, được phát triển dựa trên nguyên lý của PageRank. Mỗi câu trong văn bản được coi là một nút trong đồ thị, kết nối với các câu khác dựa trên mức độ tương đồng về từ vựng. Những câu được đánh giá là quan trọng nhất (có trọng số cao nhất) sẽ được chọn

vào bản tóm tắt. TextRank không yêu cầu dữ liệu huấn luyện sẵn, có thể áp dụng trực tiếp lên văn bản gốc, phù hợp với các bài toán cần xử lý nhanh mà không cần nhiều tài nguyên. Yêu cầu tính toán thấp, không cần GPU, có thể chạy trên máy tính cá nhân. Tuy nhiên, TextRank chỉ dựa vào sự xuất hiện của từ khóa mà không hiểu được ngữ cảnh sâu, dễ tạo ra bản tóm tắt rời rạc, thiếu mạch lạc.

LSA (Latent Semantic Analysis) sử dụng phân tích ngữ nghĩa tiềm ẩn để tìm các câu quan trọng trong văn bản. LSA xây dựng ma trận từ-văn bản và phân rã giá trị riêng (SVD) để phát hiện các chủ đề ẩn trong nội dung. Những câu liên quan đến các chủ đề chính sẽ được chọn vào bản tóm tắt. Phương pháp này không cần dữ liệu gán nhãn nhưng yêu cầu dữ liệu đầu vào phải đủ lớn, đa dạng về chủ đề để đảm bảo tính chính xác. Về tính toán, LSA đòi hỏi năng lực xử lý cao hơn TextRank, đặc biệt với dữ liệu lớn. Cần chuẩn bị dữ liệu đầu vào như loại bỏ từ dừng, chuẩn hóa văn bản, nhưng không cần huấn luyện mô hình phức tạp.

BERT (Bidirectional Encoder Representations from Transformers), là kỹ thuật hiện đại hơn nhưng vẫn thuộc hướng tóm tắt trích xuất, có khả năng hiểu ngữ cảnh sâu, giúp xác định các câu quan trọng dựa trên mối quan hệ giữa các câu và ý nghĩa toàn cục của văn bản. Phương pháp này yêu cầu dữ liệu huấn luyện được gán nhãn rõ ràng (ví dụ: câu nào quan trọng, câu nào không), hoặc cần tinh chỉnh (fine-tuning) trên tập dữ liệu đặc thù để đạt kết quả tốt nhất. Yêu cầu về năng lực tính toán cao hơn nhiều: cần GPU, bộ nhớ lớn và công cụ lập trình phức tạp như Python, PyTorch hoặc TensorFlow. Ngoài ra, quá trình huấn luyện, kiểm thử và vận hành cũng đòi hỏi công sức đáng kể, từ thiết lập môi trường, chuẩn bị dữ liệu, đến đánh giá chất lượng đầu ra.

BART (Bidirectional and Auto-Regressive Transformers) là một kỹ thuật nổi bật thuộc nhóm tóm tắt khái quát, được phát triển dựa trên kiến trúc Transformer, kết hợp điểm mạnh của cả hai mô hình mã hóa hai chiều (Bidirectional Encoder) và sinh văn bản tự hồi tiếp (Auto-Regressive Decoder). BART hoạt động theo cơ chế huấn luyện "denoising autoencoder", tức là làm hỏng (corrupt) văn bản đầu vào và yêu cầu mô hình phục hồi lại phiên bản gốc, giúp mô hình học được cách tái cấu

trúc câu và diễn đạt lại nội dung. Trong bài toán tóm tắt, BART có khả năng tự sinh câu mới, không đơn thuần là trích xuất câu từ văn bản gốc. Nhờ vậy, BART tạo ra các bản tóm tắt ngắn gọn, mạch lạc và dễ đọc hơn, đặc biệt hữu ích trong các báo cáo chuyên môn hoặc các văn bản cần truyền tải ý chính rõ ràng. Ví dụ, khi tổng hợp báo cáo lỗi kiểm thử từ nhiều người khiếm thị, BART có thể diễn giải nội dung rời rạc của nhiều người thành một đoạn văn súc tích, rõ nghĩa như: "Website thiếu văn bản thay thế cho hình ảnh, tiêu đề không phản ánh đúng nội dung, và các nút bấm không được gán nhãn rõ ràng, gây khó khăn cho người khiếm thị." Tuy nhiên, kỹ thuật này yêu cầu dữ liệu huấn luyện lớn và chất lượng cao để mô hình học được cách diễn đạt tự nhiên và chính xác. BART cũng đòi hỏi năng lực tính toán cao – cần GPU mạnh, dung lượng bộ nhớ lớn và thời gian huấn luyện lâu hơn so với các phương pháp trích xuất. Ngoài ra, để vận hành hiệu quả, người dùng cần am hiểu kỹ thuật xử lý ngôn ngữ tự nhiên (NLP), quản lý mô hình, và kiểm soát chất lượng đầu ra để tránh lỗi "bịa thông tin" – vấn đề thường gặp của các mô hình khái quát.

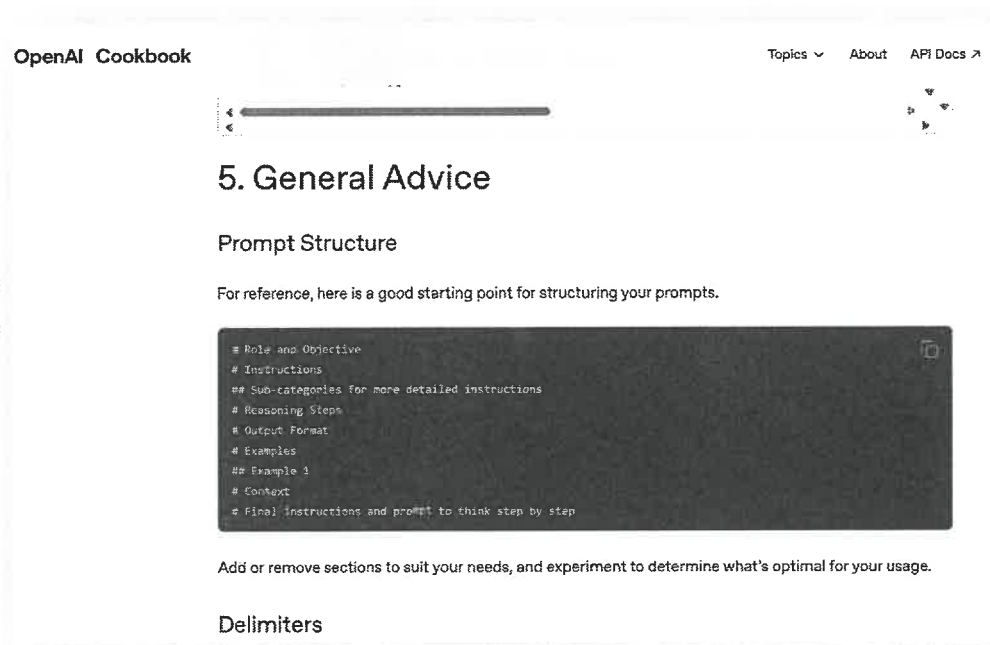
2.2. Tóm tắt văn bản dựa trên mô hình ngôn ngữ lớn (LLM)

2.2.1. Sự phát triển của mô hình ngôn ngữ lớn (LLM)

Trong những năm gần đây, sự phát triển vượt bậc của các mô hình ngôn ngữ lớn (Large Language Models – LLM) đã mở ra bước ngoặt quan trọng trong lĩnh vực xử lý ngôn ngữ tự nhiên (NLP) nói chung và tóm tắt văn bản nói riêng. Khác với các kỹ thuật truyền thống vốn chỉ dựa vào thống kê đơn giản hoặc học sâu trên quy mô nhỏ, các LLM được huấn luyện trên tập dữ liệu khổng lồ với hàng trăm tỷ tham số, cho phép chúng hiểu ngữ cảnh sâu hơn, nắm bắt các mối quan hệ phức tạp trong ngôn ngữ và sinh ra các văn bản mới một cách tự nhiên như con người viết.

Các mô hình nổi bật như GPT-4 (OpenAI), Claude (Anthropic), Gemini (Google) đều có khả năng tóm tắt văn bản theo phong cách linh hoạt, điều chỉnh độ dài, mức độ chi tiết và giọng văn theo yêu cầu đầu vào (prompt). Không chỉ dừng ở tóm tắt một tài liệu, LLM còn hỗ trợ tóm tắt đa tài liệu, tóm tắt hội thoại, hoặc tóm tắt dữ liệu bán cấu trúc và phi cấu trúc – điều mà các kỹ thuật truyền thống khó đạt được.

Đặc biệt, LLM có thể hoạt động tốt ở chế độ **zero-shot** (không cần huấn luyện thêm) hoặc **few-shot** (học từ vài ví dụ minh họa), giảm đáng kể công sức chuẩn bị dữ liệu gán nhãn so với các phương pháp trước đây. Điều này mang lại tiềm năng ứng dụng cao cho các lĩnh vực yêu cầu tổng hợp nội dung từ nhiều nguồn – chẳng hạn như tóm tắt mô tả lỗi từ nhiều kiểm thử viên khiếm thị trong bài toán đánh giá khả năng tiếp cận website.



Hình II.1 OpenAI hướng dẫn kỹ thuật prompt trên mô hình GPT 4.1

2.2.2. So sánh giữa kỹ thuật truyền thống và LLM

Phần lớn các nghiên cứu tóm tắt hiện nay sử dụng các mô hình LLM làm nền tảng [8], tận dụng prompt engineering như zero-shot, few-shot để định hướng tóm tắt theo phong cách, độ dài, thay vì cách tiếp cận tinh chỉnh trước đây. Bảng so sánh về tóm tắt văn bản sử dụng **kỹ thuật truyền thống** (TextRank, LSA, BERT, BART) và **kỹ thuật dựa trên LLM** (GPT-4, Claude, Gemini):

Bảng II.1: So sánh kỹ thuật tóm tắt văn bản truyền thống và LLM

Tiêu chí	Kỹ thuật truyền thống	Kỹ thuật dựa trên LLM
Khả năng hiểu ngữ cảnh và thông tin ngầm	- TextRank/LSA: Phụ thuộc vào mối quan hệ bề mặt (tần suất, đồng xuất hiện) giữa các từ và câu, không hiểu được ngữ cảnh sâu. - BERT/BART: Có khả năng mã hóa ngữ cảnh cục bộ (contextualized embeddings), nhưng giới hạn trong phạm vi ngắn và khó hiểu được mối liên kết ngữ nghĩa phức tạp giữa nhiều đoạn văn.	- GPT-4, Claude, Gemini: Có khả năng xử lý ngữ cảnh phức tạp trên quy mô lớn (long-context understanding), nhờ vào cơ chế attention đa tầng và huấn luyện trên tập dữ liệu lớn. [9] [10] - Hiểu được các mối quan hệ logic ẩn, thông tin ngầm (implied information), các tham chiếu gián tiếp (coreference), và ngữ cảnh xuyên suốt trong văn bản dài.
Tính linh hoạt và tùy biến	Cứng nhắc, ít hỗ trợ điều chỉnh theo mục tiêu người dùng: - TextRank, LSA: Gần như không hỗ trợ điều chỉnh kết quả theo ý người dùng, chủ yếu dựa trên tính toán thống kê, đồng xuất hiện từ khóa. - BERT, BART: Có thể tinh chỉnh qua fine-tuning nhưng yêu cầu kỹ thuật phức tạp và tốn chi phí.	Linh hoạt nhờ điều chỉnh qua prompting: - Có thể dễ dàng tạo tóm tắt ngắn, dài, theo phong cách trang trọng/thân thiện, theo định dạng JSON, Markdown... - Có thể yêu cầu trả lời theo vai trò cụ thể (giảng viên, nhà báo, luật sư...) [9].
Chất lượng	- Không được thiết kế tối ưu hóa cho khả năng sinh ngôn	- Thiết kế tối ưu cho khả năng sinh ngôn ngữ tự nhiên nhờ được huấn luyện trên

Tiêu chí	Kỹ thuật truyền thống	Kỹ thuật dựa trên LLM
ngôn ngữ đầu ra	ngữ tự nhiên (NLG) mà thường phục vụ cho các tác vụ khác như classification, QA nên đầu ra mang tính rời rạc, thiếu tính liên mạch	tập dữ liệu lớn và đa dạng, vậy nên đầu ra tự nhiên, mạch lạc, gần giống con người viết.
Khả năng tích hợp vào hệ thống người dùng	Thường phải tự triển khai đòi hỏi kỹ năng lập trình, đào tạo & finetune mô hình để tích hợp vào hệ thống thực tế. - Các mô hình truyền thống nếu tự triển khai thường yêu cầu phần cứng thấp, ví dụ TextRank, LSA có thể chạy trực tiếp trên máy tính cá nhân mà không cần GPU, BERT & BART đòi hỏi phần cứng cao hơn cần GPU. - Tuy nhiên các mô hình này thiếu dịch vụ thương mại để thuê hoặc mua mô hình được dựng sẵn API.	Thường sẽ thuê mua API dịch vụ thương mại sẵn có để tích hợp nhanh chóng và dễ dàng vào hệ thống thực tế. - Các mô hình LLM nếu tự triển khai sẽ yêu cầu phần cứng rất cao, ví dụ: cụm server đa GPU cao cấp NVIDIA A100, H100. Khó khả thi để cá nhân hay doanh nghiệp nhỏ ở cả góc độ tài chính đầu tư cho phần cứng lẫn phức tạp khi triển khai. - Tuy nhiên, mô hình LLM lại rất sẵn sàng API dịch vụ thương mại trên hầu hết nhà cung cấp Cloud (OpenAI, Microsoft, Google ...). Hỗ trợ nhiều nền tảng và ngôn ngữ lập trình, đi kèm tài liệu kỹ thuật và cộng đồng hỗ trợ rộng rãi.
Khả năng giải thích	Dễ giải thích và độ minh bạch cao: - TextRank, LSA: Dễ giải thích hơn nhờ tính chất dựa trên thuật toán thống kê (ví	- LLM như GPT-4, Claude, Gemini là các mô hình hộp đen (black-box) và độ phức tạp cao với hàng trăm tỷ tham số nên hạn chế tiếp cận thông tin cũng như khó giải thích chi tiết tại sao mô hình

Tiêu chí	Kỹ thuật truyền thống	Kỹ thuật dựa trên LLM
	dự: dựa trên tần suất xuất hiện từ, đồng xuất hiện, hoặc vector hóa). Người dùng có thể truy vết cách thuật toán tính trọng số và chọn câu. - BERT, BART: Khả năng giải thích khó hơn do là mô hình neural network nhiều lớp, nhưng vẫn có thể sử dụng kỹ thuật attention visualization để giải thích một phần.	đưa ra một tóm tắt cụ thể.

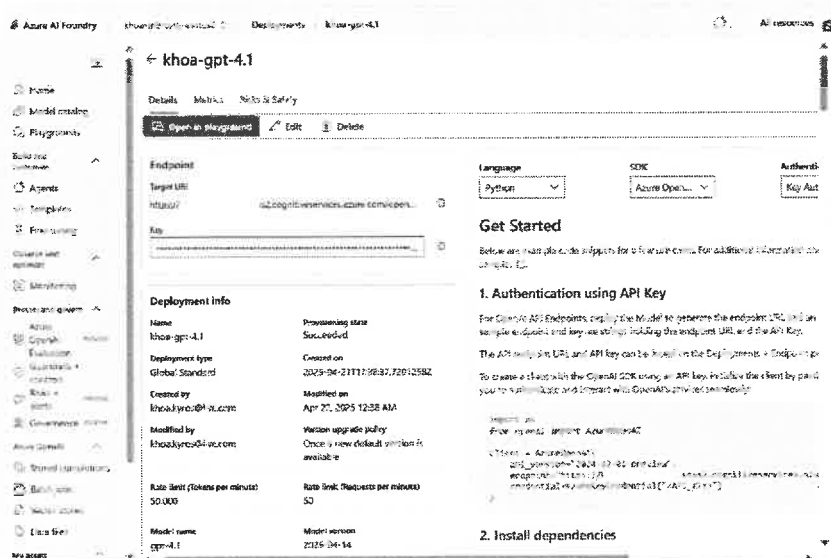
Từ bảng so sánh trên, có thể thấy rõ sự khác biệt căn bản giữa các kỹ thuật tóm tắt truyền thống và kỹ thuật dựa trên mô hình ngôn ngữ lớn (LLM). Trong khi các kỹ thuật truyền thống đơn giản hơn, dễ giải thích và yêu cầu phần cứng thấp, chúng lại hạn chế về khả năng hiểu ngữ cảnh sâu, tính linh hoạt, chất lượng ngôn ngữ đầu ra, và khả năng mở rộng ứng dụng. Ngược lại, các mô hình LLM như GPT-4, Claude, Gemini vượt trội ở khả năng sinh ngôn ngữ tự nhiên, hiểu ngữ cảnh phức tạp và linh hoạt theo yêu cầu người dùng, nhưng lại đối mặt với thách thức về giải thích kết quả và yêu cầu phần cứng rất cao nếu không thuê mua API thương mại.

2.3. Chọn mô hình GPT để triển khai trong nghiên cứu

2.3.1. Lý do lựa chọn GPT

Các mô hình ngôn ngữ lớn (LLM) phổ biến hiện nay như GPT-4 (OpenAI), Claude (Anthropic), Gemini (Google) đều cho thấy khả năng tóm tắt văn bản tốt. Tuy nhiên, trong phạm vi nghiên cứu này, GPT, cụ thể là **GPT-4.1**, được lựa chọn làm mô hình LLM nền tảng dựa trên các lý do sau:

- **Được hỗ trợ tốt bởi hệ sinh thái công cụ API và cộng đồng sử dụng:** GPT là mô hình phổ biến nhất trong cộng đồng nghiên cứu và ứng dụng, với hệ sinh thái phát triển mạnh mẽ. Các API và SDK đa ngôn ngữ, cùng khả năng tích hợp linh hoạt với các nền tảng như Microsoft Copilot Studio, Google Colab Platform giúp việc thử nghiệm, xây dựng và triển khai mô hình trở nên dễ dàng và tiết kiệm thời gian.
- **Tính sẵn sàng cao:** GPT được cung cấp thông qua dịch vụ OpenAI API và Azure Open AI với chính sách giá rõ ràng và khả năng mở rộng cao. Ngoài ra, dự án đã có sẵn tài khoản Azure với 150 đô Mỹ/ tháng miễn phí, nên phù hợp sử dụng cho truy cập mô hình GPT để nghiên cứu ngay.

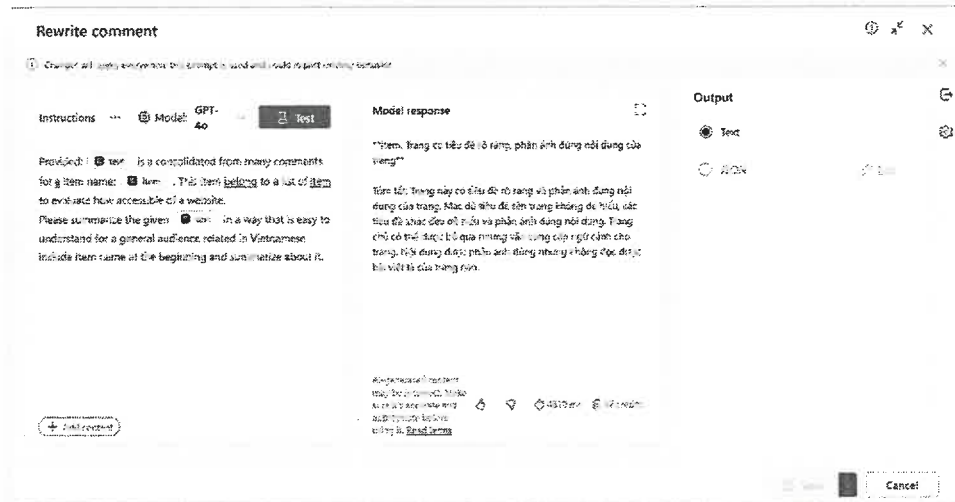


Hình II.2: Giao diện sử dụng GPT 4.1 trên Microsoft Azure AI Portal

2.3.2. Minh chứng từ dự án thực tế

Kinh nghiệm thực tiễn từ dự án đánh giá khả năng tiếp cận gần đây với UNDP là minh chứng quan trọng cho việc lựa chọn GPT làm mô hình chủ đạo. Trong dự án này, nhóm nghiên cứu đã tiến hành phát triển nhanh công cụ trên Microsoft Copilot Studio và kết nối GPT để thử nghiệm tóm tắt kết quả đánh giá lỗi khả năng tiếp cận do cộng tác viên khiếm thị thực hiện trên nhiều website. Mặc dù chỉ là bản thử nghiệm nhanh nhưng tạo kết quả khá ấn tượng về độ chính xác và sự

mạch lạc của bản tóm tắt do GPT sinh ra. Kết quả này khẳng định tính khả thi của việc lựa chọn GPT để nghiên cứu chuyên sâu hơn trong đề án này.



Hình II.3: Giao diện cấu hình GPT trên Microsoft Copilot Studio để thử nghiệm phát triển ứng dụng nhanh

2.4. Kết luận

Chương II đã trình bày tổng quan về các kỹ thuật tóm tắt văn bản, từ các phương pháp truyền thống như TextRank, LSA, BERT, BART đến các kỹ thuật hiện đại dựa trên mô hình ngôn ngữ lớn (LLM) như GPT-4.

Phân tích so sánh cho thấy, trong khi các kỹ thuật truyền thống có ưu điểm về tính đơn giản, dễ triển khai và yêu cầu tài nguyên thấp, chúng vẫn tồn tại hạn chế lớn về khả năng hiểu ngữ cảnh, chất lượng ngôn ngữ đầu ra và tính linh hoạt trong điều chỉnh theo yêu cầu cụ thể của người dùng. Ngược lại, các mô hình LLM, điển hình là GPT-4, thể hiện ưu thế vượt trội nhờ khả năng hiểu ngữ cảnh phức tạp, sinh ngôn ngữ tự nhiên và đáp ứng linh hoạt các yêu cầu tóm tắt theo nhiều định dạng, phong cách khác nhau. Dù còn tồn tại thách thức về yêu cầu phần cứng cao và tính giải thích hạn chế, việc tích hợp mô hình LLM thông qua các dịch vụ API thương mại giúp giảm đáng kể rào cản kỹ thuật cho các nhóm nghiên cứu và phát triển ứng dụng nhỏ.

Việc lựa chọn GPT-4 làm mô hình tóm tắt chính trong nghiên cứu này ngoài khả năng tóm tắt văn bản vượt trội, còn là hệ sinh thái công cụ API và cộng đồng sử dụng và tính sẵn sàng cao, có sẵn tài khoản và môi trường để nghiên cứu. Kết quả thử nghiệm bước đầu trên dự án đã triển khai đánh giá khả năng tiếp cận giúp củng cố tính khả thi của việc áp dụng GPT-4 trong nghiên cứu của đề án này.

CHƯƠNG III

THỰC NGHIỆM TÓM TẮT VĂN BẢN VỀ MÔ TẢ LỖI

3.1. Giới thiệu dự án đánh giá tính tiếp cận của của 20 cổng thông tin điện tử

3.1.1. Giới thiệu

Trong năm 2023 và 2024, Chương trình Phát triển Liên Hợp Quốc (UNDP) và Viện Nghiên cứu Chính sách và Phát triển Truyền thông (IPS) đã thực hiện hai “Đánh giá 63 Cổng Dịch vụ công cấp tỉnh từ góc độ người dùng”. Theo hai báo cáo này, 63 cổng dịch vụ công cấp tỉnh đều mắc lỗi kỹ thuật, gây khó khăn cho người khuyết tật nhìn dùng trình đọc màn hình. Khó khăn trong sử dụng dịch vụ công trực tuyến cũng được người khuyết tật nhìn phản ánh trong “Đánh giá thường niên mức độ hòa nhập người khuyết tật trong quản trị địa phương năm 2024” do UNDP và Viện Nghiên cứu Phát triển Mekong (MDRI) đồng thực hiện.

<https://hnmvn.vn/quan-he-quoc-te/tiep-can-cong-thong-tin-dien-tu-cua-cac-co-quan-nha-nuoc-nang-cao-co-hoi-hoa-nhap-cho-n>

Sáng 9/4/2025, Hội Người mù Việt Nam phối hợp với Chương trình Phát triển Liên Hợp Quốc (UNDP) tổ chức cuộc họp kỹ thuật “Tiếp cận cổng thông tin điện tử của các cơ quan nhà nước – nâng cao cơ hội hòa nhập cho NKT tại Việt Nam”.

Tham dự sự kiện trực tiếp tại tòa nhà Xanh Một Liên Hợp Quốc có bà Mette Møglestue - Phó Đại sứ Na Uy tại Việt Nam, ông Patrick Haverman - Phó Trưởng đại diện thường trú UNDP tại Việt Nam, ông Nguyễn Hồng Ngọc - Phó Vụ trưởng Vụ Văn hóa và Xã hội - Ủy ban Văn hóa và Xã hội của Quốc hội, ông Tạ Ngọc Trí - Phó trưởng Ban thường trực Ban Chỉ đạo giáo dục trẻ khuyết tật và trẻ em có hoàn cảnh khó khăn - Bộ Giáo dục và Đào tạo, ông Cù Kim Long - Phó Giám đốc Trung tâm Công nghệ thông tin - Bộ Khoa học và Công nghệ, đại diện đơn vị thuộc Bộ Tư pháp, Bộ Tài chính, Bộ Xây dựng, Bộ Nội vụ, Thư viện Quốc hội, các tổ chức của và vì NKT, các viện nghiên cứu, tổ chức về công nghệ. Ngoài ra, nhiều đại biểu tham dự và tham vấn trực tuyến qua Zoom.



Hình III.1 Bài đăng về buổi họp kỹ thuật giữa UNDP & Hội người Mù Việt Nam "Tiếp cận cổng thông tin điện tử của các cơ quan nhà nước – nâng cao cơ hội hòa nhập cho NKT tại Việt Nam "

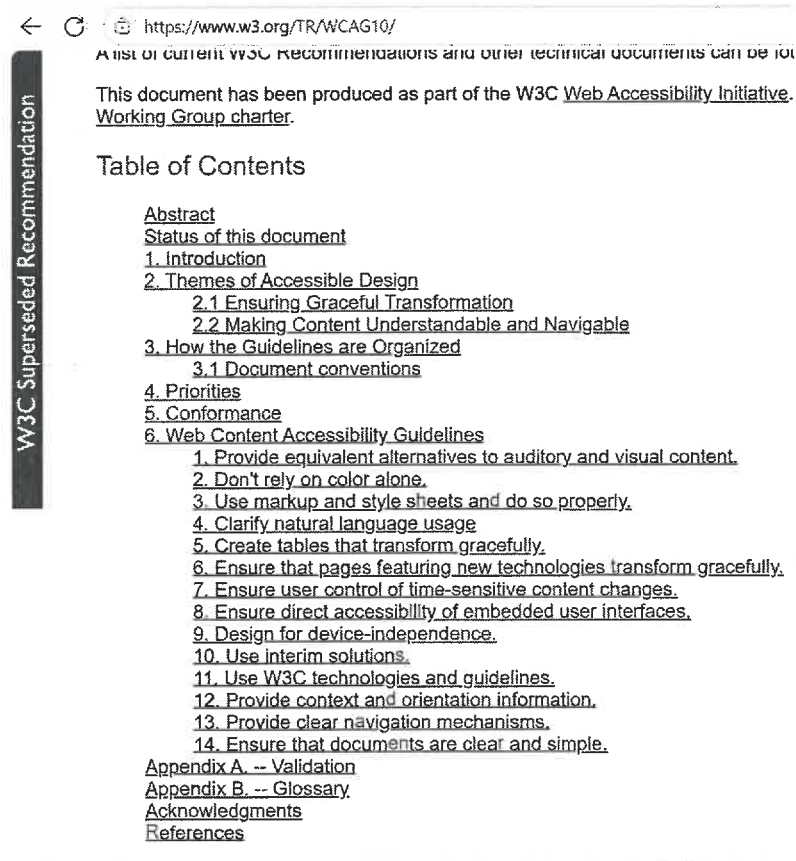
Kế thừa kết quả của các nghiên cứu trên, UNDP và Hội Người mù Việt Nam triển khai đánh giá tính tiếp cận của 20 cổng thông tin điện tử [11] của 18 bộ, văn phòng Chính phủ và Quốc hội để tiếp tục đóng góp dữ liệu làm căn cứ cho việc cải thiện các cổng dịch vụ công trực tuyến và cổng thông tin điện tử của các cơ quan Nhà nước. Nghiên cứu cũng góp phần thúc đẩy việc thực thi quyền tiếp cận thông tin của người khuyết tật được quy định trong Điều 9 và Điều 21 của Công ước Quốc tế về Quyền của Người khuyết tật (CRPD), và đã được nội luật hóa trong các văn bản quy phạm pháp luật của Việt Nam .

Bảng III.1: Danh sách 20 cổng thông tin được đánh giá trong dự án

STT	Cơ quan	Địa chỉ Website
1	Bộ Công an	https://bocongan.gov.vn/
2	Bộ Công Thương	https://moit.gov.vn/
3	Bộ Giáo dục và Đào tạo	https://moet.gov.vn/
4	Bộ Giao thông vận tải	https://mt.gov.vn/vn
5	Bộ Kế hoạch và Đầu tư	https://www.mpi.gov.vn
6	Bộ Khoa học và Công nghệ	https://www.most.gov.vn/vn/
7	Bộ Lao động - Thương binh và Xã hội	https://www.molisa.gov.vn/
8	Bộ Ngoại giao	https://www.mofa.gov.vn/vi/
9	Bộ Nội vụ	https://moha.gov.vn/
10	Bộ Nông nghiệp và Phát triển nông thôn	https://mard.gov.vn/
11	Bộ Quốc phòng	https://www.mod.gov.vn/
12	Bộ Tài chính	https://www.mof.gov.vn/
13	Bộ Tài nguyên và Môi trường	https://www.monre.gov.vn/
14	Bộ Thông tin và Truyền thông	https://mic.gov.vn/
15	Bộ Tư pháp	https://www.moj.gov.vn/

STT	Cơ quan	Địa chỉ Website
16	Bộ Văn hóa, Thể thao và Du lịch	https://bvhttdl.gov.vn/
17	Bộ Xây dựng	https://moc.gov.vn/vn/
18	Bộ Y tế	https://moh.gov.vn/
19	Chính phủ	https://chinhphu.vn/
20	Quốc hội (dự thảo luật)	https://duthaoonline.quochoi.vn/

Nghiên cứu do một nhóm cộng tác viên gồm 04 người khiếm thị dùng trình đọc màn hình và 02 người không dùng trình đọc màn hình tiến hành đánh giá trên 20 cổng thông tin điện tử bên trên từ tháng 8 đến tháng 10/2024. 20 cổng thông tin điện tử này nằm trong các nhóm trang web phải tuân thủ tối thiểu bộ Hướng dẫn Tiếp cận nội dung web (WCAG 1.0) và khuyến khích đáp ứng WCAG 2.0, và chưa được đánh giá trong các nghiên cứu của UNDP trước đây.



Hình III.2 Bài viết về bộ WCAG 1.0 trên website của W3C

Bộ tiêu chí dùng để đánh giá các cổng thông tin điện tử trong nghiên cứu này được chọn lọc và kết hợp giữa WCAG 1.0 và 2.1 AA do một số tiêu chí trong WCAG 1.0 không còn áp dụng với các nền tảng web hiện nay. Trước khi tiến hành khảo sát, các nghiên cứu viên được tập huấn về bộ tiêu chí kết hợp này để đảm bảo từng người hiểu đúng từng tiêu chí.

3.1.2. Kết quả của dự án

Kết quả đánh giá cho thấy rằng, mặc dù đã có những nỗ lực đáng kể trong việc cải thiện tính tiếp cận, nhiều cổng thông tin vẫn tồn tại các rào cản đối với người sử dụng công nghệ hỗ trợ. Một số trang web đạt mức độ tiếp cận tốt ở một số tiêu chí nhất định, nhưng vẫn chưa đảm bảo tiếp cận toàn diện theo các tiêu chuẩn quốc tế như WCAG 1.0 hoặc 2.1 AA.

Những thách thức chủ yếu bao gồm: thiết kế giao diện chưa thân thiện với người dùng trình đọc màn hình và thông tin chưa được cung cấp bằng các phương thức khác nhau (hình ảnh, màu sắc, mô tả...) cho đa dạng người dùng. Ngoài ra, sự nhất quán trong việc triển khai các tiêu chí tiếp cận giữa các cổng thông tin cũng chưa cao, dẫn đến trải nghiệm không đồng đều cho người dùng.

Từ những phát hiện này, nghiên cứu nhấn mạnh sự cần thiết của việc cải thiện chính sách và kỹ thuật để tăng cường tính tiếp cận của các cổng thông tin điện tử, đảm bảo mọi người dân, đặc biệt là người khuyết tật, có thể tiếp cận thông tin từ các cơ quan Chính phủ và Quốc hội một cách bình đẳng.

Ngoài ra, dự án đã thu thập được hơn 3.600 đánh giá chi tiết từ các cộng tác viên khiếm thị, phản ánh mức độ đáp ứng tiêu chí và các lỗi cụ thể trên từng trang/tiêu chí của các cổng thông tin. Tuy nhiên, kết quả này mới chỉ dừng lại ở mức dữ liệu thô, chưa được thực hiện tóm tắt, tổng hợp chi tiết cho từng website và từng tiêu chí cụ thể. Việc tổng hợp cuối cùng chỉ dừng ở mức kết luận định tính về các nhóm lỗi phổ biến, mà chưa có các báo cáo chi tiết cho từng trang, từng tiêu chí. Nguyên nhân chính đến từ hạn chế nguồn lực và thiếu chuyên môn trong việc biên tập, tóm tắt báo cáo, cụ thể các cộng tác viên khiếm thị có thể thực hiện kiểm thử và ghi nhận lỗi, nhưng chưa được đào tạo bài bản về kỹ năng

biên tập, trong khi các chuyên gia kiểm thử lại không có thời gian hoặc không trực tiếp tham gia vào công đoạn tổng hợp này.

3.2. Bộ dữ liệu để thực nghiệm trong nghiên cứu

3.2.1. Quy trình thu thập dữ liệu gốc

Trong khuôn khổ dự án hợp tác giữa UNDP và Hội người mù Việt Nam, nhóm dự án đã tổ chức việc kiểm thử và thu thập kết quả kiểm thử của 20 websites là cổng thông tin điện tử như sau:

Bước	Nội dung công việc	Thực hiện bởi	Kết quả đầu ra
1	Chọn 4–5 trang con (page) cho mỗi website trong 20 website cổng thông tin điện tử được chỉ định. Lựa chọn trang con thuộc loại: • Trang chủ • Tin tức - sự kiện • Chế độ chính sách • Bài viết / Điểm báo	Trưởng nhóm	Danh sách trang con cần kiểm thử cho từng website.
2	Chọn các tiêu chí kiểm thử WCAG phù hợp & ưu tiên cho người khiếm thị và viết mô tả chi tiết cách đánh giá từng tiêu chí.	Chuyên gia	Danh sách 49 tiêu chí kiểm thử và hướng dẫn cách thực hiện cho từng tiêu chí. Ví dụ: • Trang có mồi tiêu đề (banner) (thuộc WCAG phiên bản 1.0, tiêu chí 12.3) • Trang có tiêu đề rõ ràng, mô tả đúng tên trang web (thuộc

			WCAG phiên bản 2.1, tiêu chí 2.4.2)
3	Tạo Google Form để cộng tác viên khiếm thị nhập kết quả kiểm thử cho từng trang con.	Chuyên gia	Trang web trên Google Form chứa các trường thông tin để nhập về website, trang con, môi trường kiểm thử ... và kết quả kiểm thử ứng với từng tiêu chí, bao gồm điểm số (1-5) và mô tả lỗi Ví dụ ở Hình III.3 Hình III.3: Giao diện Google Form đã tạo để cộng tác viên nhập kết quả đánh giá bởi cá nhân
4	04 cộng tác viên khiếm thị sử dụng công cụ đọc màn hình NVDA / JAWS và thực hiện kiểm thử độc lập với từng trang con ở Bước 1, cho các tiêu chí ở Bước 2. Sau đó nhập kết quả lên Google Form đã tạo ở Bước 3.	Từng cộng tác viên	Dữ liệu kiểm thử được điền vào Google Form đã tạo ở bước trên
5	Xuất kết quả trên Google Form thành file Excel duy nhất	Chuyên gia	File Excel chứa kết quả kiểm thử khả năng tiếp cận. Kết quả riêng lẻ theo từng cộng tác viên, chưa được tổng hợp Ví ở Hình III.4

Đánh giá mức độ tiếp cận của cổng thông tin điện tử (dành cho cán bộ khảo sát dùng trình đọc màn hình)

khoabcvrt@gmail.com Switch account

Your email will be recorded when you submit this form

* Indicates required question

I. Thông tin chung

1. Họ tên cán bộ khảo sát *

Your answer

2. website được đánh giá *

Choose

3. Tên trang và đường link cụ thể (ví dụ: trang chủ + đường link) *

Your answer

4. Tên trình đọc màn hình và phiên bản (ví dụ: Jaws 2024) *

Your answer

Hình III.3: Giao diện Google Form đã tạo để cộng tác viên nhập kết quả đánh giá bởi cá nhân

1. Họ tên cán bộ khảo sát	2. website được đánh giá	3. Tên trang và đường link cụ thể (ví dụ: trang chủ + đường link)	4. Tên trình đọc màn hình và phiên bản (ví dụ: Jaws 2024)
Trần Bình Minh	1. Bộ Quốc phòng	Trang chủ: https://mod.gov.vn/ NVDA 202 Google C	5 Dễ hiểu
Nguyễn Tuấn Tú	1. Bộ Quốc phòng	Tin tức - sự kiện https://mod.gov.vn/ NVDA 202 Google C	4 Dễ hiểu
Trần Bình Minh	1. Bộ Quốc phòng	Tin tức - sự kiện https://mod.gov.vn/ NVDA 202 Google C	4 Dễ hiểu nhưng không d
Nguyễn Tuấn Tú	1. Bộ Quốc phòng	Chế độ chính sách: https://mod.gov.vn/ NVDA 202 Google C	3 Đọc đúng là mục Chế d
Trần Bình Minh	1. Bộ Quốc phòng	Bài viết: https://mod.gov.vn/ NVDA 202 Google C	5 Hoàn Tiểu đề trên, trình duyệt 1 = không Tiểu đề trên trình duyệt 5 = hoàn 1 Tiểu đề tr 1 =
Nguyễn Tuấn Tú	1. Bộ Quốc phòng	Bài viết: https://mod.gov.vn/ NVDA 202 Google C	5 Hoàn Tiểu đề trên, trình duyệt 1 = không Tiểu đề trên trình duyệt 5 = hoàn 1 Tiểu đề tr 1 =
Trần Bình Minh	1. Bộ Quốc phòng	Trang chủ: https://bocongan.gov.vn/ NVDA 202 Google C	5 Hoàn Tiểu đề trên, trình duyệt 1 = không Tiểu đề trên trình duyệt 5 = hoàn 1 Tiểu đề tr 1 =
Nguyễn Tuấn Tú	2. Bộ Công an	Trang chủ: https://bocongan.gov.vn/ NVDA 202 Google C	5 Hoàn Tiểu đề trên, trình duyệt 1 = không Tiểu đề trên trình duyệt 5 = hoàn 1 Tiểu đề tr 1 =
Trần Bình Minh	2. Bộ Công an	Phổ biến giáo dục pháp luật: ht NVDA 202 Chrome 1 5 = Hoàn Tiểu đề trên, trình duyệt 1 = không Tiểu đề trên trình duyệt 5 = hoàn 1 Tiểu đề tr 1 =	3 Có thể hiểu
Nguyễn Tuấn Tú	2. Bộ Công an	Phổ biến giáo dục pháp luật: ht NVDA 202 Chrome p	3 Có thể hiểu
Trần Bình Minh	2. Bộ Công an	Hỏi đáp: https://bocongan.gov.vn/ NVDA 202 Chrome P	5 Dễ hiểu
Nguyễn Tuấn Tú	2. Bộ Công an	Bài viết: https://bocongan.gov.vn/ NVDA 202 Chrome 1 5 = Hoàn Tiểu đề trên, trình duyệt 1 = không Tiểu đề trên trình duyệt 5 = hoàn 1 Tiểu đề tr 1 =	3 Đúng nhưng không đọc
Trần Bình Minh	3. Bộ Ngoại giao	Trang chủ: https://mofa.gov.vn/ NVDA 202 Chrome 1 5 = Hoàn Tiểu đề trên, trình duyệt 1 = không Tiểu đề trên trình duyệt 5 = hoàn 1 Tiểu đề tr 1 =	3 Chỉ đọc là homepage kí 1 = không Đọc là homepage và cũn 1 = không Không có Kh

Hình III.4: File excel File Excel dữ liệu thô, chứa toàn bộ kết quả kiểm thử, riêng lẻ theo từng cộng tác viên

3.2.2. Mô tả và thống kê bộ dữ liệu gốc

Mỗi bản ghi (dòng) trong file Excel chứa kết quả kiểm thử khả năng tiếp cận tương ứng với 1 trang con (page) trên 1 website cụ thể, được kiểm thử bởi 1 cộng tác viên khiếm thị, xem Bảng III.2.

Cụ thể, mỗi bản ghi sẽ bao gồm các thông tin chính sau:

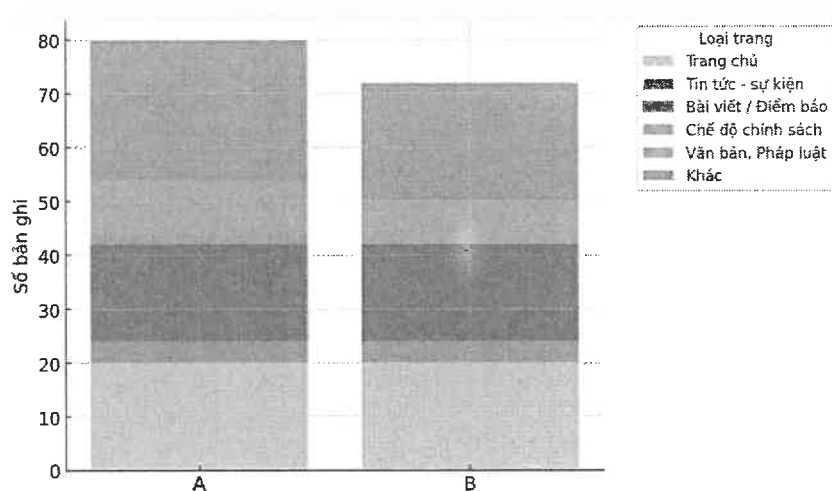
- Website, Trang con được đánh giá và Người đánh giá: Đây là ba trường thông tin cơ bản, giúp xác định rõ ràng bản ghi này thuộc về website nào, trang con nào, và cộng tác viên nào là người thực hiện kiểm thử.
- Kết quả đánh giá cho 49 tiêu chí kiểm thử. Mỗi tiêu chí được ghi lại dưới dạng cặp thông tin:
 - Điểm số: Thang điểm từ 1 đến 5, thể hiện mức độ đạt yêu cầu của tiêu chí đó trên trang web (1 là kém nhất, 5 là tốt nhất).
 - Mô tả lỗi: Phần ghi chú của cộng tác viên về lỗi hoặc vấn đề phát hiện được (nếu có)

Bảng III.2: Cấu trúc bảng và mẫu của 1 bản ghi trong bộ dữ liệu gốc từ file Excel chứa bộ dữ liệu kết quả đánh giá từ nhiều cộng tác viên

#	Website đánh giá	Trang con được đánh giá	Người đánh giá	Tiêu chí 1: Điểm	Tiêu chí 1: Mô tả lỗi	...	Tiêu chí 49: Điểm	Tiêu chí 49: Mô tả lỗi
1	Boquocphong.vn	Trang chủ	Tester 01	5	Tiêu đề trên trình duyệt phản ánh đúng tên bài viết	...	2	Tiêu đề trang là chế độ chính sách trong khi tiêu đề đầu tiên của trang lại là văn bản hướng dẫn. Điều này có thể gây khó khăn cho người dùng trong việc định hướng và tìm kiếm nội dung trong trang.

			Tester 02	5	Phản ánh đúng nội dung, nhưng không đọc bài viết là của trang nào.	... 1	Tiêu đề trang không đúng, ghi là Web Content Viewer tiêu đề cấp 2
		Tin tức sự kiện	Tester 01	...	...		...
			Tester 02	...	...		...
...		...					
20	Quốc hội (dự thảo luật)	...					
...	...	...					

Các cộng tác viên khiếm thị đã chia thành hai nhóm nhỏ. Nhóm A, gồm Trần Bình Minh và Nguyễn Tuấn Tú, đã kiểm thử websites thứ tự từ 1 đến 10, đóng góp 80 bản ghi. Nhóm B, gồm Đặng Mạnh Cường và Võ Cẩm Giang, đã kiểm thử websites từ 11 đến 20, đóng góp 72 bản ghi. Biểu đồ Hình III.5 tổng số bản ghi từng nhóm tester và cách chúng phân bố theo loại trang:



Hình III.5: Phân bố bản ghi theo thể loại và nhóm kiểm thử

3.2.3. Chuẩn bị bộ dữ liệu cho thực nghiệm

Trong quy trình thu thập dữ liệu gốc, do thiết kế chưa tốt ở bước tạo Google Form yêu cầu cộng tác viên khiếm thị phải nhập thủ công một số trường thông tin đã được chuẩn hóa trước đó, ví dụ như tên trang con, nên đã phát sinh tình trạng dữ liệu không đồng nhất – như lỗi chính tả, viết tắt, hoặc viết hoa/thường không nhất quán. Điều này dẫn đến sai lệch dữ liệu cần xử lý ở bước sau. Vì vậy, tôi đã thực hiện một bước xử lý và làm sạch dữ liệu, bao gồm:

- Chuẩn hóa thông tin: Đồng bộ lại các trường dữ liệu có sự khác biệt trong cách nhập, ví dụ tên trang con được cộng tác viên nhập không nhất quán, giúp đảm bảo tính đồng nhất và chính xác của dữ liệu.
- Xử lý dữ liệu thiếu, trùng lặp hoặc không hợp lệ: Rà soát toàn bộ tập dữ liệu để phát hiện và xử lý các giá trị bị thiếu hoặc lỗi phát sinh trong quá trình nhập liệu.

Trong tương lai, bước này hoàn toàn có thể được loại bỏ bằng cách cải tiến biểu mẫu Google Form: chuẩn bị sẵn danh sách dữ liệu chuẩn hóa để cộng tác viên lựa chọn (thay vì nhập tay), từ đó giảm thiểu sai sót và tiết kiệm thời gian xử lý sau thu thập.

Sau khi chuẩn hóa, tôi đã thực hiện **chuyển đổi dữ liệu sang cấu trúc bảng mới**, xem Bảng III.3. Cụ thể mỗi bản ghi mới đại diện cho kết quả kiểm thử khả năng tiếp cận của 1 tiêu chí cụ thể trên 1 trang con của 1 website, và có điểm số kèm mô tả lỗi của toàn bộ cộng tác viên.

#	Website đánh giá	Trang con được đánh giá	Tiêu chí đánh giá	Tổng hợp đánh giá gốc của các testers
1	Boquocphong.vn	Trang chủ: https://mod.gov.vn/	T1. Trang có tiêu đề rõ ràng, phản ánh đúng nội dung của trang (thuộc WCAG phiên bản 2.1, tiêu chí 2.4.2).	1. "Nguyễn Tuấn Tú": "5 điểm", "VÌ đây là trang chủ của công nên từ trang chủ có thể được bỏ qua nhưng vẫn cung cấp được ngữ cảnh cho

				trang", 2. "Trần Bình Minh": "5 điểm", "Dễ hiểu"
...	...	...	...	...
367 5	Quốc hội (dự thảo luật)	Trang chủ http://duthaonline.quochoi.vn/Pages/trangchu.aspx	O3 Có cảnh báo rõ ràng khi bắt kỳ thao tác nào làm chuyển trang (thuộc WCAG phiên bản 1.0, tiêu chí 7.5). Mô tả cụ thể lỗi nếu có liên quan tiêu chí đánh giá	1. "Đặng Mạnh Cường": "3 điểm", "Không có kiểu cảnh báo này", 2. "Võ Cẩm Giang": "2 điểm", "Trang không có bất kì dấu hiệu nào cho biết đang chuyển sang trang mới, Thậm chí tiêu đề trang vẫn luôn không đổi, làm cho người dùng không biết là có tải được trang hay chưa."

Bảng III.3: Cấu trúc bảng và mẫu của 1 bản ghi trong bộ dữ liệu chuẩn bị cho thực nghiệm

Kết quả cuối cùng sau xử lý, bộ dữ liệu đã được chuyển đổi và làm sạch gồm

3 675 bản ghi, sẵn sàng cho bước tóm tắt tiếp theo.

3.2.4. Thách thức tóm tắt thủ công

Sau khi thu thập dữ liệu kiểm thử riêng lẻ bên trên, một yêu cầu quan trọng tiếp theo là thực hiện **tóm tắt văn bản mô tả lỗi** của nhiều cộng tác viên (N người) thành một văn bản đánh giá chung **cho từng bản ghi** (mỗi trang con – mỗi tiêu chí).

Tuy nhiên, thực tế từ dự án kiểm thử khả năng tiếp cận 20 cổng thông tin do UNDP triển khai cho thấy, nhóm dự án **đã không thể thực hiện được công việc tóm tắt chi tiết này** cho từng bản ghi. Thay vào đó, báo cáo cuối cùng chỉ dừng ở phần kết luận tổng quan do chuyên gia viết, không có phần tóm tắt lỗi chi tiết cho từng trang con – tiêu chí. Nguyên nhân chính đến từ những khó khăn do khối lượng công việc và hạn chế nguồn lực thực tế:

- **Khối lượng công việc quá lớn:** Với 3675 bản ghi, ước tính thời gian thực hiện cho mỗi bản ghi khoảng 60 giây để đọc nội dung mô tả lỗi từ các cộng

tác viên và 60 giây để viết phần tóm tắt. Tổng cộng, công việc này cần khoảng 122 giờ làm việc liên tục, và với nguồn lực thời gian hạn chế của cộng tác viên khoảng 2 giờ/ ngày thì **cần tới 2-3 tháng làm việc bổ sung thêm** cho việc tóm tắt chi tiết này.

- **Kỹ năng hạn chế của cộng tác viên:** Các cộng tác viên là người khiếm thị, được đào tạo ngắn hạn để thực hiện kiểm thử, nhưng chưa được đào tạo kỹ năng biên tập báo cáo chuyên nghiệp. Vậy nên khó khăn không chỉ đến từ cần thêm thời gian làm việc, mà còn cả năng lực để đáp ứng yêu cầu chất lượng tóm tắt văn bản chuyên môn sao cho mạch lạc, rõ ràng, và nhất quán giữa các bản ghi.

Đây cũng là lý do cần thiết phải nghiên cứu và phát triển quy trình và công cụ tóm tắt văn bản mô tả lỗi, nhằm phục vụ các bài toán thực tế như kiểm thử khả năng tiếp cận website bởi người khiếm thị.

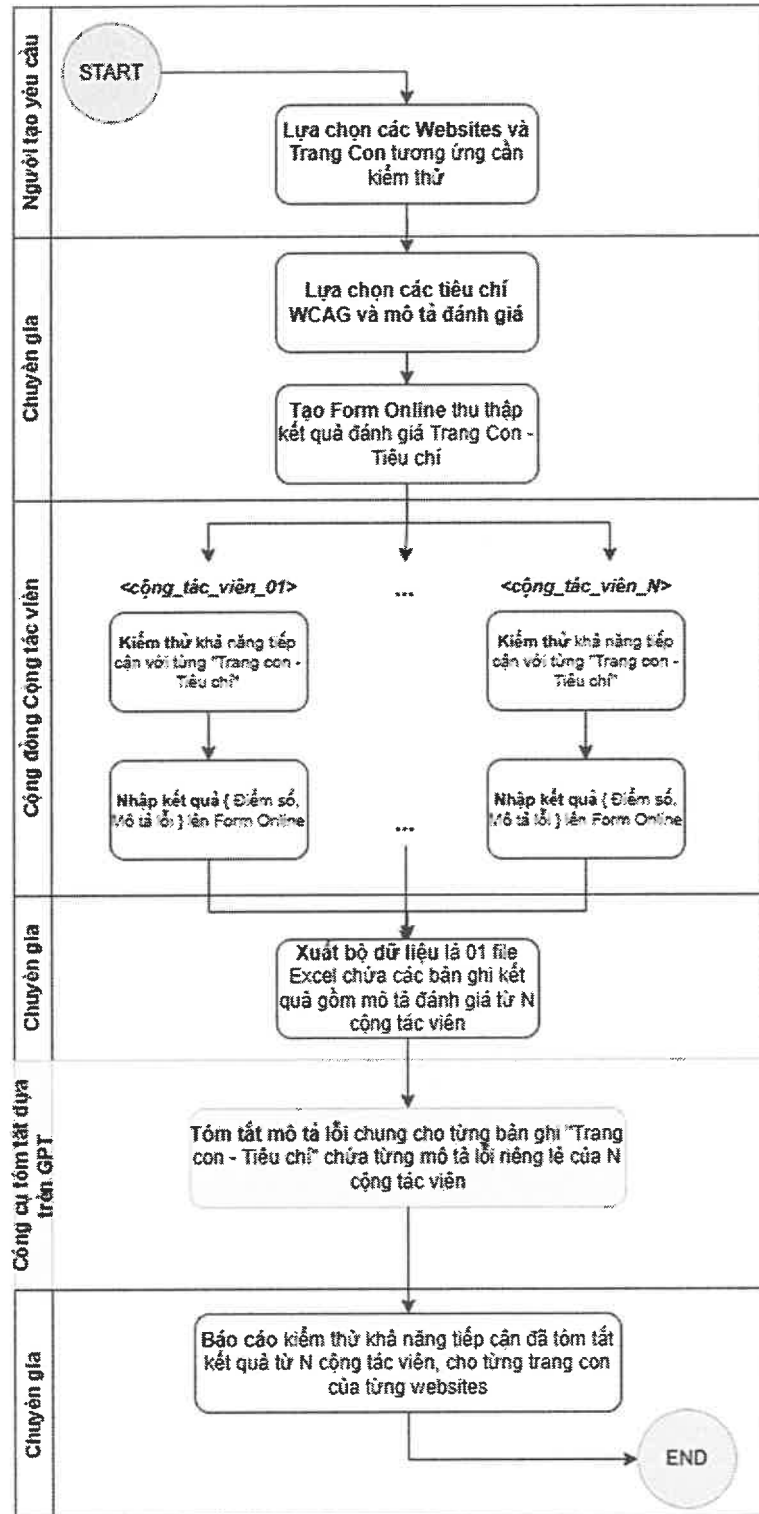
3.3. Áp dụng kỹ thuật tóm tắt văn bản bằng GPT trong tổng hợp mô tả lỗi

3.3.1. Phát triển quy trình kết hợp công cụ tóm tắt

Thông qua **kết hợp giữa kinh nghiệm thực tế** trong dự án đã triển khai với UNDP và **nghiên cứu về các kỹ thuật** tóm tắt văn bản tự động, đề án này **đề xuất một quy trình kiểm thử khả năng tiếp cận website** bao gồm ba trụ cột chính:

1. **Chuyên gia về khả năng tiếp cận (accessibility):** Đóng vai trò lựa chọn và xác định các tiêu chí kiểm thử WCAG, đồng thời xây dựng hướng dẫn chi tiết giúp đảm bảo cộng tác viên khiếm thị có thể thực hiện đánh giá một cách nhất quán.
2. **Cộng đồng cộng tác viên người khiếm thị:** Là lực lượng người khiếm thị thực hiện kiểm thử trên từng trang con – tiêu chí của các website, sau đó nhập dữ liệu vào biểu mẫu online (Google Form) đã được chuẩn hóa.
3. **Công cụ tóm tắt mô tả lỗi tự động sử dụng GPT-4:** Là giải pháp quan trọng nhằm xử lý thách thức thực tế về khối lượng công việc và đảm bảo chất lượng tóm tắt tương đương chuyên gia có kỹ năng biên tập. Công cụ này thực hiện tóm tắt mô tả lỗi từ nhiều cộng tác viên (N người) về một bản tóm tắt chung duy nhất cho từng trang con – tiêu chí.

Quy trình được minh họa rõ ràng trong sơ đồ luồng sau, thể hiện sự phối hợp giữa các vai trò:



Hình III.6: Quy trình kiểm thử khả năng tiếp cận kết hợp Chuyên gia - Cộng đồng cộng tác viên người khiếm thị - Công cụ tóm tắt dựa trên GPT

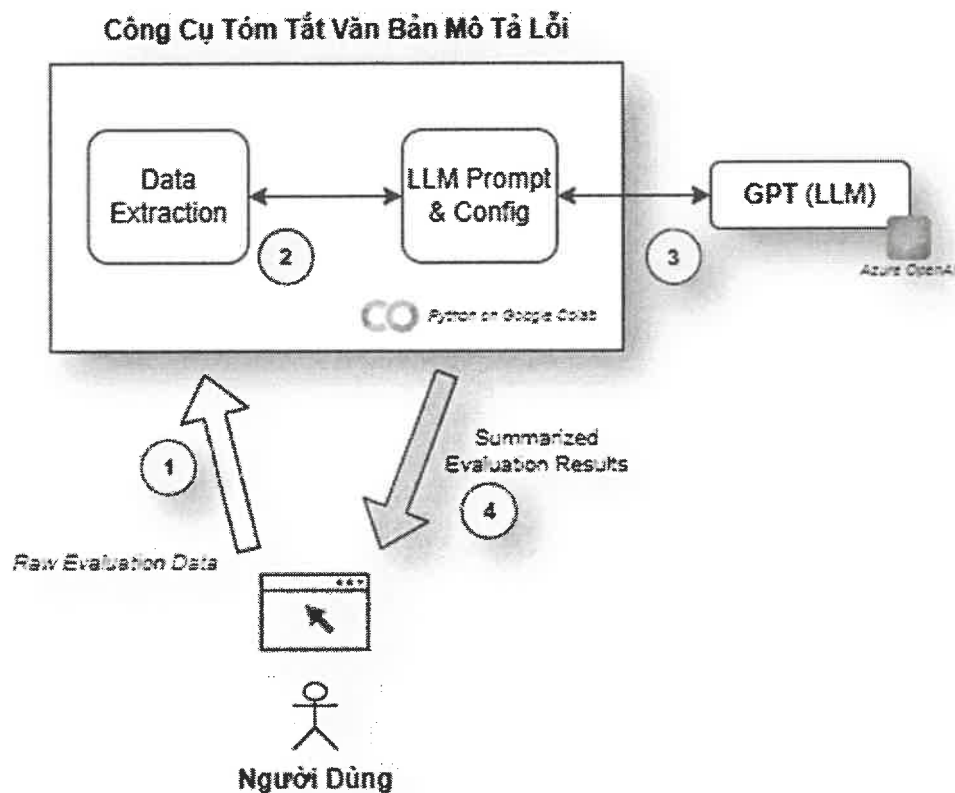
Trong đó, “Công cụ tóm tắt” trong quy trình trên là một thành phần cốt lõi, được nghiên cứu phát triển chi tiết trong đề án và sẽ được mô tả cụ thể ở các phần tiếp theo.

3.3.2. Phát triển công cụ tóm tắt

Xây dựng một công cụ tóm tắt, là một thành phần trong quy trình bên trên, để tóm tắt văn bản lỗi kiểm thử khả năng tiếp cận từ kết quả đánh giá của nhiều cộng tác viên khiếm thị (N người) cho mỗi trang con – tiêu chí. Hướng tới mục tiêu:

- Tiết kiệm thời gian và công sức biên tập.
- Tạo ra bản tóm tắt lỗi có chất lượng đồng đều, rõ ràng, dễ hiểu cho đơn vị chủ quản website

Mô tả kiến trúc giải pháp:



Hình III.7 Kiến trúc giải pháp của Công cụ tóm tắt

Công cụ tóm tắt là một ứng dụng web được phát triển bằng ngôn ngữ lập trình Python, chạy trên môi trường điện toán đám mây Google Colab. Mô hình GPT được

sử dụng thông qua API thương mại cung cấp bởi Azure OpenAI, phục vụ cho việc tóm tắt văn bản.

Quy trình hoạt động:

1. Người dùng truy cập công cụ trên web và tải lên (upload) tệp Excel chứa dữ liệu gốc – bao gồm các mô tả lỗi kiểm thử khả năng tiếp cận được thực hiện bởi nhiều cộng tác viên cho từng trang web và tiêu chí kiểm thử.
2. Công cụ sẽ tự động trích xuất thông tin từ tệp dữ liệu, bao gồm: mô tả lỗi của từng cộng tác viên, trang web được đánh giá, và tiêu chí kiểm thử.
3. Công cụ gửi dữ liệu đã trích xuất cùng mẫu prompt được cấu hình sẵn đến API mô hình GPT (Azure OpenAI) để thực hiện tóm tắt. GPT sẽ sinh ra bản tóm tắt chính xác nhưng dễ hiểu cho mỗi trang web/tiêu chí dựa trên thông tin đầu vào.
4. Người dùng tải về (download) tệp Excel mới, trong đó mỗi bản ghi chứa thông tin về trang web, tiêu chí và bản tóm tắt lỗi đã được tự động tạo bởi GPT.

Ví dụ minh họa cho đầu vào và đầu ra của công cụ cho phần tóm tắt mô tả lỗi:

- Dữ liệu mô tả lỗi ban đầu:

```
{
  "Cộng tác viên 01": "Trang có tải lại khi thực hiện tìm kiếm nhưng kết quả xuất hiện không được thể hiện rõ ràng về mức độ liên quan đến từ khóa",
  "Cộng tác viên 02": "Không thể tìm kiếm bất kỳ thứ gì"
}
```

- Dữ liệu mô tả lỗi sau khi được công cụ tóm tắt:

Lỗi nghiêm trọng: Chức năng tìm kiếm không hỗ trợ tìm kiếm linh hoạt (có dấu, không dấu, thiếu chữ, sai chính tả). Kết quả tìm kiếm không rõ ràng hoặc không xuất hiện, gây khó khăn cho người dùng khi sử dụng chức năng này.

Chi phí hạ tầng:

- Công cụ là một ứng dụng web, chạy trên Google Colab ở tier miễn phí do đòi hỏi ít tài nguyên và không cần chạy liên tục.
- Mô hình GPT phiên bản GPT-4.1 được sử dụng thông qua API thương mại. Tại thời điểm kiểm tra (31/05/2025) có bảng giá tính theo token và quy đổi khoảng 1000 - 2000 Việt Nam đồng/ cho 1 bản ghi tóm tắt độ dài như ví dụ bên trên.

Model	Pricing (1M Tokens)	Pricing with Batch API (1M Tokens)
GPT-4.1-2025-04-14	Input: \$2 Cached input: \$0.50 Output: \$8	N/A
GPT-4.1-mini-2025-04-14	Input: \$0.40 Cached input: \$0.10 Output: \$1.60	N/A
GPT-4.1-nano-2025-04-14	Input: \$0.10 Cached input: \$0.03 Output: \$0.40	N/A

Hình III.8 Bảng giá API mô hình GPT 4.1 từ Azure OpenAI

Các phương án tinh chỉnh kỹ thuật tóm tắt trên GPT:

Công cụ được phát triển trên nền tảng LLM lựa chọn là GPT-4, nhờ khả năng xử lý ngôn ngữ tự nhiên vượt trội cũng như các ưu điểm khác đã mô tả ở 2.3.1. Lý do lựa chọn GPT. Nhằm tìm ra cách tối ưu nhất để ứng dụng GPT-4 vào công cụ, đề án thử nghiệm 3 phương án prompt trên cùng bộ dữ liệu, cụ thể

Bảng III.4: Ba cách sinh tóm tắt với GPT-4

Cách	Chi tiết Prompt	Ví dụ Prompt minh họa
Prompt Zero-shot	1: Đơn giản, baseline để đo lường khả năng tự nhiên của GPT: GPT chỉ nhận	Prompt mẫu: Bạn sẽ nhận dữ liệu đầu vào như sau:

	dữ liệu comment và tóm tắt trực tiếp. Tính điểm trung bình từ score (nếu hợp lệ, 1–5).	<pre>{ "url": "https://mod.gov.vn/", "page": "Trang chủ", "criterion": "T1. Trang có tiêu đề rõ ràng, phản ánh đúng nội dung của trang (thuộc WCAG phiên bản 2.1, tiêu chí 2.4.2) ", "evaluations": [{"score": 2, "comment": "Thiếu nhãn alt cho hình ảnh"}, {"score": 4, "comment": "Liên kết không có mô tả rõ ràng" }] }</pre> Yêu cầu: Tóm tắt mô tả lỗi kiểm thử khả năng tiếp cận web (web accessibility) và tính điểm trung bình từ các kiểm thử viên từ dữ liệu đầu vào.
Prompt 2: Zero-shot with Instructional Role Prompting	Mở rộng hơn prompt 1 bằng cách bổ sung ngữ cảnh và vai trò: GPT được yêu cầu đóng vai "Kiểm thử viên trưởng".	Prompt mẫu: Bạn sẽ nhận dữ liệu đầu vào như sau: <pre>{ "url": "https://mod.gov.vn/",</pre>

	Hướng dẫn sử dụng thêm ngữ cảnh trang web và mô tả tiêu chí WCAG để làm rõ bản tóm tắt.	<pre> "page": "Trang chủ", "criteria": "T1. Trang có tiêu đề rõ ràng, phản ánh đúng nội dung của trang (thuộc WCAG phiên bản 2.1, tiêu chí 2.4.2) ", "evaluations": [{"score": 2, "comment": "Thiếu nhãn alt cho hình ảnh"}, {"score": 4, "comment": "Liên kết không có mô tả rõ ràng" }] } </pre> Yêu cầu: Đóng vai là kiểm thử viên trưởng cần tạo bảng tổng hợp kết quả kiểm thử để gửi cho đơn vị chủ quản website đọc và hiểu họ đang gặp lỗi web accessibility nào. Tóm tắt mô tả lỗi kiểm thử khả năng tiếp cận web (web accessibility) cho 1 tiêu chí cụ thể từ các kiểm thử viên riêng lẻ và tính điểm trung bình từ các kiểm thử viên từ dữ liệu đầu vào.
Prompt 3: Few-	Mở rộng hơn prompt	Prompt mẫu:

shot with Role Prompting	2 bằng cách thêm ví dụ: • Bổ sung 2 ví dụ rõ ràng về cách tóm tắt chuẩn/tệ • Hướng dẫn phong cách tóm tắt: sử dụng “ngữ cảnh lỗi”, “ví dụ cụ thể”, “kết luận rõ”	Bạn sẽ nhận dữ liệu đầu vào như sau: <pre>{ "url": "https://mod.gov.vn/", "page": "Trang chủ", "criterion": "T1. Trang có tiêu đề rõ ràng, phản ánh đúng nội dung của trang (thuộc WCAG phiên bản 2.1, tiêu chí 2.4.2) ", "evaluations": [{"score": 2, "comment": "Thiếu nhãn alt cho hình ảnh"}, {"score": 4, "comment": "Liên kết không có mô tả rõ ràng" }] }</pre> Yêu cầu: Đóng vai là kiểm thử viên trưởng cần tạo bảng tổng hợp kết quả kiểm thử để gửi cho đơn vị chủ quản website đọc và hiểu họ đang gặp lỗi web accessibility nào. Tóm tắt mô tả lỗi kiểm thử khả năng tiếp cận web (web accessibility) cho 1 tiêu chí cụ thể từ các kiểm thử
--	--	--

viên riêng lẻ và tính điểm trung bình từ các kiểm thử viên từ dữ liệu đầu vào.

Ví dụ về kết quả tóm tắt mô tả lỗi kiểm thử:A

Ví dụ 01: Dữ liệu đầu vào là: "Tiêu đề trang là chế độ chính sách trong khi tiêu đề đầu tiên của trang lại là văn bản hướng dẫn. Điều này có thể gây khó khăn cho người dùng trong việc định hướng và tìm kiếm nội dung trong trang." . Kết quả tóm tắt là: "Lỗi tiếp cận (accessibility) liên quan đến sự không nhất quán giữa tiêu đề trang và tiêu đề đầu tiên của nội dung. Ví dụ cụ thể, tiêu đề trang hiển thị là "chế độ chính sách", nhưng tiêu đề cấp 1 lại là "văn bản hướng dẫn". Sự khác biệt này có thể gây nhầm lẫn cho người dùng, ảnh hưởng đến khả năng định hướng và tìm kiếm thông tin."

Ví dụ 02: Dữ liệu đầu vào là: "Dễ hiểu. Vì đây là trang chủ nên không nhất

		thiết phải có từ "trang chủ" trên tiêu đề của trang. Thay vào đó, trang nêu rõ tên là "cổng thông tin điện tử bộ công an" là đã đủ.". Kết quả tóm tắt là: "Không có lỗi. Tuy nhiên, có góp ý rằng trang chủ không nhất thiết phải có từ "Trang chủ" trong tiêu đề. Thay vào đó, chỉ cần hiển thị rõ ràng tên trang là "Cổng thông tin điện tử Bộ Công an" cũng đủ để giúp người dùng nhận diện nội dung trang mà không gây trùng lặp hoặc dư thừa thông tin."
--	--	--

3.3.3. Phương pháp đánh giá

Để đánh giá chất lượng đầu ra của công cụ tóm tắt lỗi dựa trên GPT, đề án sử dụng phương pháp **đánh giá thủ công bởi chuyên gia**. Cụ thể, chuyên gia sẽ chấm điểm từng bản tóm tắt được sinh ra từ 3 loại prompt khác nhau (Zero-shot, Zero-shot with Role Prompting, Few-shot with Role Prompting) trên cùng một bản ghi dữ liệu.

Mỗi bản tóm tắt sẽ được chuyên gia chấm điểm riêng cho 4 tiêu chí mô tả ở

Bảng III.5, theo thang điểm từ 1 (rất kém) đến 5 (rất tốt):

Bảng III.5: Hướng dẫn chấm điểm thủ công kết quả tóm tắt bởi công cụ

Tiêu chí	Câu hỏi trọng tâm	Yếu tố kiểm tra	Ví dụ minh họa
1. Đầy đủ thông tin	<i>Tóm tắt đã bao quát mọi điểm quan trọng trong bản gốc chưa?</i>	• Liệt kê đủ lỗi chính • Nhắc hậu quả/vi phạm WCAG (nếu có) • Không bỏ sót quan điểm khác biệt giữa các tester	Bản gốc nêu 3 lỗi → tóm tắt cũng đề cập 3 lỗi.
2. Chính xác	<i>Nội dung có đúng sự thật và không xuyên tạc?</i>	• Diễn giải đúng mức độ nghiêm trọng • Dùng đúng mã WCAG, số lượng, trạng thái lỗi • Không thêm/đổi chi tiết	Bản gốc nói “5/11 phần tử không đọc được” → tóm tắt ghi “5/11”, không thành “một số”.
3. Dễ hiểu	<i>Người không chuyên có thể đọc một lần và hiểu?</i>	• Câu ngắn ≤ 25 từ • Thuật ngữ giải thích/viết đầy đủ (WCAG 13.1...) • Tránh rút gọn khó đoán (“btn”, “cbx”) • Logic mạch lạc (Có-Vấn đề → Hậu quả)	“Liên kết ‘xem thêm’ dẫn sang trang khác, gây nhầm lẫn cho người dùng” — dễ hiểu hơn “Link alias mismatch”.
4. Ngắn gọn	<i>Thông tin dư thừa tối thiểu nhưng vẫn giữ ý cốt lõi?</i>	• ≤ 120 từ hoặc ≤ 3 câu • Không lặp (không nhắc WCAG 13.1 hai lần) • Bỏ mệnh đề “như đã nói ở trên” • Gộp ý tương tự	Thay vì 5 câu rời rạc → gộp 2-3 câu đầy đủ nghĩa.

Kết quả đánh giá không chỉ phục vụ mục đích so sánh hiệu quả của 3 loại prompt, mà còn trả lời câu hỏi quan trọng hơn: Liệu chất lượng bản **tóm tắt sinh ra từ GPT có đủ tốt để thay thế công việc tóm tắt thủ công của chuyên gia** hay không.

3.3.4. Thực hiện đánh giá và kết quả

Do khối lượng dữ liệu đầu vào rất lớn – với **3.675 bản ghi**, mỗi bản ghi có 3 kết quả tóm tắt từ 3 cách prompt khác nhau, tổng cộng hơn **11.000 bản tóm tắt cần đánh giá** – việc chấm điểm toàn bộ dataset là không khả thi về nguồn lực. Vì vậy, đề án áp dụng phương pháp lấy mẫu phân tầng 10%, tương ứng 368 bản ghi, để thực hiện đánh giá thủ công.

Việc lấy mẫu phân tầng giúp giảm tải công việc đánh giá, nhưng vẫn đảm bảo tính đại diện và khách quan của kết quả. Cách lấy mẫu đảm bảo tính đại diện, theo các tiêu chí phân tầng sau:

- Bao phủ toàn bộ 20 website được khảo sát.
- Đảm bảo đủ các loại trang con: Trang chủ, Tin tức – Sự kiện, Bài viết/Điểm báo, Chế độ chính sách, và các loại khác.

Tiếp theo, thực hiện đánh giá bằng chấm điểm với **368 bản ghi**, cho ra kết quả chi tiết ở

*Phụ lục 02: Các bản tóm tắt và kết quả chấm điểm đầu ra công cụ tóm tắt dựa trên GPT, và **trung bình điểm** như sau:*

Tiêu chí	Prompt 1	Prompt 2	Prompt 3
1. Đầy đủ thông tin	4,896	4,847	4,869
2. Chính xác	4,153	4,297	4,777
3. Dễ hiểu	4,624	4,986	4,989
4. Ngắn gọn	4,698	4,965	4,937
<i>Toàn bộ tiêu chí</i>	<i>4,593</i>	<i>4,774</i>	<i>4,893</i>

Đánh giá điểm cho 3 cách prompt:

- Prompt 1: Điểm cao nhất ở tiêu chí "Đầy đủ thông tin" (4.90), thấp nhất ở tiêu chí "Chính xác" (4.15) do diễn đạt chưa rõ ràng, thiếu bối cảnh. Phù hợp nhất cho yêu cầu cần độ đầy đủ thông tin cao. Ví dụ một tóm tắt mô tả lỗi bởi Prompt 1: *Trang có tải lại khi thực hiện tìm kiếm nhưng kết quả xuất hiện không được thể hiện rõ ràng về mức độ liên quan đến từ khóa. Không thể tìm kiếm bất kỳ thứ gì.*
- Prompt 2: Điểm rất cao ở tiêu chí "Dễ hiểu" (4.99) và "Ngắn gọn" (4.96), nhưng thấp hơn một chút về tiêu chí "Chính xác" (4.30). Phù hợp nhất khi cần một bản tóm tắt dễ hiểu, ngắn gọn và xúc tích. Ví dụ một tóm tắt cùng mô tả lỗi trên bởi Prompt 2: *Chức năng tìm kiếm không hỗ trợ tìm kiếm linh hoạt (có dấu, không dấu, thiếu chữ, sai chính tả), kết quả tìm kiếm không rõ ràng hoặc không xuất hiện.*
- Prompt 3: Điểm cao đồng đều ở cả 4 tiêu chí (đặc biệt "Dễ hiểu": 4.99, "Ngắn gọn": 4.94). Cân bằng tốt giữa tất cả các tiêu chí và hiệu quả tổng hợp cao. Ví dụ một tóm tắt cùng mô tả lỗi trên bởi Prompt 3: *Lỗi nghiêm trọng: Chức năng tìm kiếm không hỗ trợ tìm kiếm linh hoạt (có dấu, không dấu, thiếu chữ, sai chính tả). Kết quả tìm kiếm không rõ ràng hoặc không xuất hiện, gây khó khăn cho người dùng khi sử dụng chức năng này.*

Đánh giá chung dựa trên mục tiêu đánh giá:

1. Các bản tóm tắt do GPT sinh ra đạt điểm đánh giá chuyên gia rất cao (đều trên mức 4/5 ở các tiêu chí).
2. Chất lượng các bản tóm tắt từ GPT, đặc biệt là Prompt 3, thể hiện khả năng thay thế tốt cho việc tóm tắt thủ công, với ưu điểm nổi bật ở sự dễ hiểu, ngắn gọn, và độ chính xác tốt.
3. **Prompt 3 được đánh giá là cách hiệu quả nhất**, cân đối được tất cả các yêu cầu của một bản tóm tắt tốt và tiềm năng cao để sử dụng thay thế cho công việc thủ công.

Như vậy, công cụ tóm tắt dựa trên GPT đủ tốt để thay thế công việc tóm tắt mô tả lỗi khả năng tiếp cận thủ công của chuyên gia, và **lựa chọn cách Prompt 3 (Few-shot with Role Prompting)** cho công cụ này.

3.4 Kết luận

Chương III đã trình bày thực nghiệm chi tiết về nghiên cứu và phát triển quy trình kiểm thử kết hợp với công cụ tóm tắt văn bản dựa trên mô hình GPT để tổng hợp mô tả lỗi khả năng tiếp cận web. Quy trình này bao gồm ba trụ cột chính: (1) Chuyên gia xây dựng bộ tiêu chí và hướng dẫn, (2) Cộng đồng người khiếm thị trực tiếp kiểm thử và thu thập dữ liệu, và (3) Công cụ tóm tắt tự động bằng GPT để tổng hợp kết quả từ nhiều người đánh giá.

Kết quả thực nghiệm khẳng định tính khả thi và hiệu quả của quy trình đề xuất. Công cụ tóm tắt văn bản sử dụng GPT, đặc biệt với phương án Prompt 3 (Few-shot with Role Prompting), đạt chất lượng vượt trội (điểm trung bình trên 4.89/5), đảm bảo dễ hiểu, ngắn gọn và độ chính xác cao. Tuy nhiên, việc áp dụng GPT cũng cần lưu ý kiểm soát độ chính xác để tránh sinh ra các lỗi không có thật.

Tổng thể, sự kết hợp giữa quy trình và công cụ đã giải quyết hiệu quả thách thức về khối lượng công việc lớn và năng lực biên tập hạn chế của cộng tác viên khiếm thị, mở ra tiềm năng ứng dụng rộng rãi trong các dự án tương tự về đánh giá khả năng tiếp cận số trong tương lai.

KẾT LUẬN VÀ KIẾN NGHỊ

Kết quả đã đạt được

Đề án đã hoàn thành việc nghiên cứu và phát triển một **quy trình kiểm thử khả năng tiếp cận web** dành cho cộng tác viên khiếm thị, kết hợp cùng **công cụ tóm tắt mô tả lỗi** tự động sử dụng GPT-4. Quy trình gồm ba trụ cột chính: (1) chuyên gia xây dựng tiêu chí và hướng dẫn kiểm thử; (2) cộng đồng cộng tác viên người khiếm thị thực hiện kiểm thử thực tế và thu thập dữ liệu; (3) công cụ tóm tắt tự động bằng GPT để tổng hợp kết quả kiểm thử từ nhiều người đánh giá

Kết quả thực nghiệm từ dự án đã triển khai cũng như công cụ tóm tắt mới trên bộ dữ liệu từ 20 website dịch vụ công cho thấy:

- **Quy trình phối hợp chặt chẽ** giữa các vai trò, giúp đảm bảo chất lượng đánh giá, tính nhất quán và dễ dàng áp dụng thực tế trong điều kiện hạn chế nguồn lực.
- **Công cụ tóm tắt lỗi dựa trên GPT**, đặc biệt với cách prompt Few-shot with Role Prompting (Prompt 3), **cho kết quả cao** về độ đầy đủ thông tin, chính xác, dễ hiểu và ngắn gọn (điểm trung bình trên 4.89/5 theo đánh giá chuyên gia).
- **Hiệu quả tổng hợp báo cáo được cải thiện** rõ rệt, giảm đáng kể thời gian và công sức so với phương pháp thủ công, giúp tiết kiệm hàng trăm giờ lao động mà vẫn đảm bảo chất lượng báo cáo đầu ra ở mức chuyên nghiệp.

Đề án đã khẳng định tiềm năng ứng dụng mô hình LLM nói chung, hay GPT-4 nói riêng, trong bài toán tóm tắt lỗi kiểm thử khả năng tiếp cận và khả năng thay thế quy trình tóm tắt thủ công trước đây, góp phần nâng cao hiệu quả kiểm thử và thúc đẩy cơ hội nghề nghiệp cho người khiếm thị.

Hạn chế và hướng phát triển

Bên cạnh những kết quả đạt được, nghiên cứu vẫn còn một số hạn chế và cần tiếp tục hoàn thiện:

- **Giới hạn về phạm vi đánh giá công cụ:** Hiện tại mới thực hiện đánh giá thủ công bởi chuyên gia trên khoảng 10% tập dữ liệu. Cần mở rộng đánh giá

toàn bộ các bản ghi còn lại, đồng thời bổ sung nhiều chuyên gia tham gia để đa dạng góc nhìn và hoàn thiện các tiêu chí chấm điểm.

- **Cải thiện prompt:** Kết quả đánh giá có thể được cải thiện hơn nữa bằng cách tinh chỉnh prompt dựa trên phản hồi từ nhiều chuyên gia và dự án thực tế.
- **Tinh chỉnh cấu hình GPT:** Ngoài các kỹ thuật prompt, thì mô hình GPT cho phép cài đặt các thông số cấu hình như temperature, top_p, max_tokens. Ví dụ với thông số temperature, nó cho phép cài đặt giá trị từ 0.0 tới 2.0. Giá trị cao hơn như 0.8 sẽ làm cho đầu ra ngẫu nhiên hơn, trong khi giá trị thấp hơn như 0.2 sẽ làm đầu ra tập trung và có tính quyết định hơn.
- **Tăng cường khả năng giải thích của công cụ:** Trong tương lai, có thể kết hợp kỹ thuật Retrieval-Augmented Generation (RAG) với kho ví dụ thực tế về lỗi WCAG để giúp công cụ tóm tắt không chỉ tạo nội dung mạch lạc mà còn giải thích rõ ràng lý do và minh họa bằng ví dụ cụ thể cho từng lỗi phát hiện.
- **Mở rộng áp dụng:** Quy trình và công cụ có thể được mở rộng theo hướng: Sử dụng các LLM khác như Claude (Anthropic), Gemini (Google) hay LLM mã mở mới nổi DeepSeek; Áp dụng cho các lĩnh vực kiểm thử khả năng tiếp cận khác như với ứng dụng chạy trên Windows, hay thiết bị di động Android, iOS.

DANH MỤC TÀI LIỆU THAM KHẢO

- [1] C.-A. S. M. V. Y. Z. Masudul Hasan Masud Bhuiyan, "Non-Western Perspectives on Web Inclusivity: A Study of Accessibility Practices in the Global South," arxiv, 2025.
- [2] Bộ Thông Tin và Truyền Thông, "Thông Tư: Quy định về việc cung cấp dịch vụ công trực tuyến và bảo đảm khả năng truy cập thuận tiện đối với trang thông tin điện tử hoặc cổng thông tin điện tử của cơ quan nhà nước," Bộ Thông Tin và Truyền Thông, Việt Nam, 2017.
- [3] Quốc Hội, "Luật Người Khuyết Tật," Quốc Hội, Việt Nam, 2010.
- [4] W3C, "Accessibility Principles," [Online]. Available: <https://www.w3.org/WAI/fundamentals/accessibility-principles/>. [Accessed 24 05 2025].
- [5] United Nations Development Programme (UNDP) Viet Nam, "First Review of Accessibility and User Friendliness of 63 Provincial E-service Portals," UNDP Viet Nam, Hà Nội, 2023.
- [6] W3C, "Web Content Accessibility Guidelines (WCAG) 2.1," W3C, 06 05 2025. [Online]. Available: <https://www.w3.org/TR/WCAG21/>.
- [7] Saef Iqbal, "How to Perform a Web Accessibility Audit: Step-by-Step Guide," Pivotal Accessibility, 13 11 2024. [Online]. Available: <https://www.pivotalaccessibility.com/2024/11/how-to-perform-a-web-accessibility-audit-step-by-step-guide/>.
- [8] P. S. Y. J. Z. Haopeng Zhang, "A Systematic Survey of Text Summarization: From Statistical Methods to Large Language Models," arXiv, 2024.
- [9] OpenAI, "GPT-4 Technical Report v6," OpenAI, 2024.
- [10] Anthropic, "Introducing the next generation of Claude," Anthropic, 4 3 2024. [Online]. Available: <https://www.anthropic.com/news/claude-3-family>.

- [11] UNDP & VBA, "Tiếp cận công thông tin điện tử của các cơ quan nhà nước – nâng cao cơ hội hòa nhập cho NKT tại Việt Nam," Hội người mù Việt Nam, 9 4 2025. [Online]. Available: <https://hnmvn.vn/quan-he-quoc-te/tiep-can-cong-thong-tin-dien-tu-cua-cac-co-quan-nha-nuoc-%E2%80%93-nang-cao-co-hoi-hoa-nhap-cho-nkt-tai-viet-nam>.

PHỤ LỤC

Phụ lục 01: Mã nguồn công cụ tóm tắt văn bản mô tả lỗi dựa trên GPT

```
# Import thư viện cần thiết
import pandas as pd
import json
from openai import AzureOpenAI

# Hàm đọc file JSON chứa dữ liệu kiểm thử
def read_json_file(file_path):
    '''Đọc file JSON chứa dữ liệu kiểm thử'''
    with open(file_path, 'r', encoding='utf-8') as file:
        data = json.load(file)
    return data

# === TẠO PROMPT ===

# Cách 1: Zero-shot Prompt - Tóm tắt cơ bản, không cung cấp ví dụ mẫu
def create_prompt_1(record):
    '''Tạo prompt dạng Zero-shot'''
    evaluations = "\n".join([f"- {e['score']} điểm: {e['comment']}" for e in
record['evaluations']])
    prompt = f'''Bạn sẽ nhận dữ liệu sau:
URL: {record['url']}
Trang: {record['page']}
Tiêu chí: {record['criterion']}
Đánh giá:
{evaluations}
Yêu cầu: Tóm tắt mô tả lỗi kiểm thử web và tính điểm trung bình từ các kiểm
thử viên.'''
    return prompt

# Cách 2: Zero-shot with Role Prompting - Thêm vai trò kiểm thử viên trưởng
và hướng dẫn rõ hơn
def create_prompt_2(record):
    '''Tạo prompt dạng Zero-shot có vai trò kiểm thử viên trưởng'''
    evaluations = "\n".join([f"- {e['score']} điểm: {e['comment']}" for e in
record['evaluations']])
```

```

    prompt = f'''Bạn là kiểm thử viên trưởng. Hãy tổng hợp kết quả kiểm thử
từ các cộng tác viên:
URL: {record['url']}
Trang: {record['page']}
Tiêu chí: {record['criterion']}
Đánh giá:
{evaluations}
Yêu cầu: Tạo bảng tổng hợp kết quả kiểm thử để gửi cho đơn vị chủ quản
website đọc và hiểu họ đang gặp lỗi web accessibility nào. Tóm tắt mô tả lỗi
kiểm thử khả năng tiếp cận web (web accessibility) cho 1 tiêu chí cụ thể từ
các kiểm thử viên riêng lẻ và tính điểm trung bình từ các kiểm thử viên từ
dữ liệu đầu vào.'''
    return prompt

```

Cách 3: Few-shot with Role Prompting - Cung cấp ví dụ mẫu để GPT tham khảo trước khi tóm tắt

```

def create_prompt_3(record):
    '''Tạo prompt dạng Few-shot with Role Prompting'''
    evaluations = "\n".join([f"- {e['score']} điểm: {e['comment']}" for e in
record['evaluations']])
    prompt = f'''Bạn là kiểm thử viên trưởng. Đây là 2 ví dụ mẫu:
Ví dụ 1: Lỗi tiêu đề trang không nhất quán, gây nhầm lẫn cho người dùng.
Ví dụ 2: Lỗi thiếu alt text cho hình ảnh logo, gây khó khăn cho người khiếm
thị.

```

Dữ liệu cần xử lý:

```

URL: {record['url']}
Trang: {record['page']}
Tiêu chí: {record['criterion']}
Đánh giá:
{evaluations}
Yêu cầu: Tạo bảng tổng hợp kết quả kiểm thử để gửi cho đơn vị chủ quản
website đọc và hiểu họ đang gặp lỗi web accessibility nào. Tóm tắt mô tả lỗi
kiểm thử khả năng tiếp cận web (web accessibility) cho 1 tiêu chí cụ thể từ
các kiểm thử viên riêng lẻ và tính điểm trung bình từ các kiểm thử viên từ
dữ liệu đầu vào.'''
    return prompt

```

=== GPT API ===

```

# Hàm gửi prompt tới GPT (Azure OpenAI) và nhận kết quả tóm tắt
def generate_summary(prompt, api_key, endpoint, deployment_name):
    '''Gửi prompt tới GPT API và nhận kết quả tóm tắt'''
    client = AzureOpenAI(
        api_version="2025-01-01-preview",

```

```
        azure_endpoint="https://your_azure_openai_modelname.cognitiveservice  
s.azure.com/",  
        api_key="YOUR_AZURE_OPENAI_KEY_HERE"  
    )  
    response = client.chat.completions.create(  
        model=deployment_name,  
        messages=[{"role": "user", "content": prompt}],  
        max_tokens=500  
    )  
    return response.choices[0].message.content.strip()
```

Phụ lục 02: Các bản tóm tắt và kết quả chấm điểm đầu ra công cụ tóm tắt dựa trên GPT

Phụ lục này trình bày kết quả đánh giá thủ công của chuyên gia đối với ba phương án prompt được sử dụng trong công cụ tóm tắt mô tả lỗi kiểm thử khả năng tiếp cận web. Mỗi bản tóm tắt được chấm theo bốn tiêu chí:

- Tiêu chí 1: Đầy đủ thông tin
- Tiêu chí 2: Chính xác
- Tiêu chí 3: Dễ hiểu
- Tiêu chí 4: Ngắn gọn

Kết quả được chấm điểm trên tập mẫu 10% phân tầng dữ liệu gốc, tương ứng 368 bản ghi, đảm bảo đại diện cho 20 website và các loại trang con. Mẫu minh họa bảng dữ liệu với 2 bản ghi:

2. website được đánh giá	3. Tên trang và đường link cụ thể	Tiêu chí đánh giá	Tổng hợp đánh giá gốc của các testers	Tóm Tắt: Cách_1	Tóm Tắt: Cách_2	Tóm Tắt: Cách_3	H_C1_Tiền chỉ_1	H_C1_Tiền chỉ_2	H_C1_Tiền chỉ_3	H_C1_Tiền chỉ_4	H_C2_Tiền chỉ_1	H_C2_Tiền chỉ_2	H_C2_Tiền chỉ_3	H_C2_Tiền chỉ_4	H_C3_Tiền chỉ_1	H_C3_Tiền chỉ_2	H_C3_Tiền chỉ_3	H_C3_Tiền chỉ_4
1. Bộ Quốc phòng	Bài viết: https://mod.gov.vn/vn/loi-dung/sa-cdcs/sa-cdcs-tt/doan-dai-bieu-quan-uy-trung-uong-bo-quoc-phong-vao-lang-vieng-chu-tich-ho-chi-minh-va-dang-huong-tuong-niem-cac-anh-hung-liet-si	H2 Tiêu đề (heading) được gắn đúng trình tự logic (thuộc WCAG phiên bản 1.0, tiêu chí 3.5). Mô tả cụ thể lỗi nếu có liên quan tiêu chí đánh giá	{ "Nguyễn Tuấn Tú": "Về thứ tự di chuyển thì các tiêu đề của trang đáp ứng đúng mục đích. Đúng mục đích.", "Trần Bình Minh": "Đúng mục đích"} }	Về thứ tự di chuyển thì các tiêu đề của trang đáp ứng đúng mục đích. Đúng mục đích.	Thứ tự di chuyển giữa các tiêu đề hợp lý, đáp ứng đúng mục đích cấu trúc nội dung, giúp người dùng dễ dàng điều hướng.	Không có lỗi nghiêm trọng. Thứ tự tiêu đề hợp lý, đáp ứng mục đích định hướng nội dung cho người dùng.	5	5	4	4	5	5	5	5	5	5	5	
1. Bộ Quốc phòng	Bài viết: https://mod.gov.vn/vn/loi-dung/sa-cdcs/sa-cdcs-tt/doan-dai-bieu-quan-uy-trung-uong-bo-quoc-phong-vao-lang-vieng-chu-tich-ho-chi-minh-va-dang-huong-tuong-niem-cac-anh-hung-liet-si	S3 Kết quả tìm kiếm có được thể hiện một cách dễ tiếp cận hay không (1.0 13.7). Mô tả cụ thể lỗi nếu có liên quan tiêu chí đánh giá	{ "Nguyễn Tuấn Tú": "Trang không thể hiện kết quả tìm kiếm", "Trần Bình Minh": "Kết quả chủ yếu ra kết quả là Http/1.1 Service Unavaila ble hoặc không ra kết quả." } }	Trang không thể hiện kết quả tìm kiếm. Kết quả chủ yếu ra kết quả là Http/1.1 Service Unavaila ble hoặc không ra kết quả.	Kết quả tìm kiếm không được hiển thị hoặc trả về lỗi (Http/1.1 Service Unavailabl e), không hỗ trợ người dùng tiếp cận kết quả.	Lỗi nghiêm m trọng: Kết quả tìm kiếm không hiển thị hoặc trả về lỗi, không hỗ trợ người dùng tiếp cận thông tin	5	4	5	4	5	4	5	5	4	5	5	

						cán thiết.													
--	--	--	--	--	--	---------------	--	--	--	--	--	--	--	--	--	--	--	--	--

Danh sách đầy đủ các bản ghi tóm tắt và kết quả chấm điểm được nộp kèm đề án dưới dạng file Excel.

Phụ lục 03: Bộ dữ liệu gốc thu thập từ dự án UNDP

Trong khuôn khổ dự án hợp tác giữa UNDP và Hội người mù Việt Nam, nhóm dự án đã tổ chức việc kiểm thử và thu thập kết quả kiểm thử của 20 websites. Toàn bộ kết quả kiểm thử của từng cộng tác viên được nhập trên Google Form, nên có thể xuất ra thành 1 file Excel chung. Mỗi bản ghi (dòng) trong file Excel chứa kết quả kiểm thử khả năng tiếp cận tương ứng với 1 trang con (page) trên 1 website cụ thể, được kiểm thử bởi 1 cộng tác viên khiếm thị. Cụ thể như minh họa dưới, lưu ý do hơn 100 cột nên đã xoay (cột -> hàng) để thuận tiện trình bày minh họa trong tài liệu này:

Email Address	ntuantu29@gmail.com	camgianghnmquangngai@gmail.com
1. Họ tên cán bộ khảo sát	Nguyễn Tuấn Tú	Võ Cẩm Giang
2. website được đánh giá	1. Bộ Quốc phòng	11. Bộ Xây dựng
3. Tên trang và đường link cụ thể (ví dụ: trang chủ + đường link)	Trang chủ: https://mod.gov.vn/	Trang chủ https://moc.gov.vn/vn/
4. Tên trình đọc màn hình và phiên bản (ví dụ: Jaws 2024)	NVDA 2024.1	Jaws 2024
5. Tên trình duyệt web và phiên bản (ví dụ: Google Chrome 127.0.6533.127)	Chrome 127	Google Chrome Version 127.0.6533.101
T1. Trang có tiêu đề rõ ràng, phản ánh đúng nội dung của trang (2.1 2.4.2)	5 = Hoàn toàn dễ tiếp cận	5 = Hoàn toàn dễ tiếp cận
T1. Mô tả cụ thể lỗi đó:	Vì đây là trang chủ của cổng nên từ trang chủ có thể được bỏ qua nhưng vẫn cung cấp được ngữ cảnh cho trang	Không có lỗi

T2. Trang có tiêu đề rõ ràng, mô tả đúng tên trang web (2.1 2.4.2)	4	5 = hoàn toàn dễ tiếp cận
T2 Mô tả cụ thể lỗi đó:	Tiêu đề trang phản ánh đúng nội dung của cổng bộ quốc phòng nhưng có thể được cải thiện nếu chữ thông tin điện tử (TTĐT) được viết đầy đủ thay vì viết tắt	Không có lỗi
T3. Tiêu đề trang trùng hoặc tương tự với tiêu đề cấp 1 của trang (2.1 2.4.2)	Không áp dụng	Không áp dụng
T3 Mô tả cụ thể lỗi đó:	Do đây là trang chủ nên trang không cần có tiêu đề 1 để phản ánh nội dung trang	Trang không có tiêu đề cấp 1, tiêu đề đầu tiên là tiêu đề cấp 2 và là tên của một bài viết.
D1 Trang có móc tiêu đề (banner) (1.0 - 12.3)	1 = không dễ tiếp cận chút nào	5 = hoàn toàn dễ tiếp cận
D1 Mô tả cụ thể lỗi đó:	Trang có 2 vùng chính: IBM WebSphere và Complementary Content và một vùng dành cho thanh điều hướng. Tuy nhiên, vùng dành cho thanh điều hướng lại được chứa bên trong vùng IBM WebSphere. Các vùng này không được chia chức năng rõ ràng. Có vẻ toàn bộ nội dung trang, bao gồm cả phần thông tin thêm ở cuối trang đều nằm trong vùng này. Bản thân vùng Complementary Content được đặt ở cuối trang nhưng không có nội dung nào bên trong.	Không có lỗi
D2 Trang có móc chân trang (contentinfo) (1.0 - 12.3)	1 = không dễ tiếp cận chút nào	1 = không dễ tiếp cận chút nào
D2 Mô tả cụ thể lỗi đó:	Trang có 2 vùng chính: IBM WebSphere và Complementary Content và một vùng dành cho thanh điều hướng. Tuy nhiên, vùng dành cho thanh điều hướng lại được chứa bên trong vùng IBM WebSphere. Các vùng này không được chia chức năng rõ ràng. Có vẻ toàn bộ nội dung trang, bao gồm cả phần thông tin thêm ở cuối trang đều nằm trong vùng này. Bản thân vùng Complementary Content được đặt ở cuối trang nhưng không có nội dung nào bên trong.	Trang chỉ có móc tiêu đề (banner) mà không có Main và Contentinfo.
D3 Trang có một móc chính (main) (1.0 - 12.3)	1 = không dễ tiếp cận chút nào	1 = không dễ tiếp cận chút nào

D3 Mô tả cụ thể lỗi đó:	Trang có 2 vùng chính: IBM WebSphere và Complementary Content và một vùng dành cho thanh điều hướng. Tuy nhiên, vùng dành cho thanh điều hướng lại được chứa bên trong vùng IBM WebSphere. Các vùng này không được chia chức năng rõ ràng. Có vẻ toàn bộ nội dung trang, bao gồm cả phần thông tin thêm ở cuối trang đều nằm trong vùng này. Bản thân vùng Complementary Content được đặt ở cuối trang nhưng không có nội dung nào bên trong.	Trang chỉ có mốc tiêu đề (banner) mà không có Main và Contentinfo.
D4 Bất kỳ mốc nào cũng không được chứa bên trong một mốc cấp cao nhất như Main (1.0 - 13.3)	1 = không dễ tiếp cận chút nào	Không áp dụng
D4 Mô tả cụ thể lỗi đó:	Vùng dành cho thanh điều hướng nằm bên trong vùng mang tên IBM WebSphere.	Trang chỉ có banner mà không có mốc nào khác.
D5 Thành phần tương tác đầu tiên của trang là một link cho phép đi đến nội dung chính của trang (1.0 - 13.3)	1 = không dễ tiếp cận chút nào	1 = không dễ tiếp cận chút nào
D5 Mô tả cụ thể lỗi đó:	Link này không hề tồn tại	Trang không có liên kết cho phép đi đến nội dung chính.
D6 Các mốc sắp xếp theo trình tự hợp lý (1.0 - 12.3)	1 = không dễ tiếp cận chút nào	1 = không dễ tiếp cận chút nào
D6 Mô tả cụ thể lỗi đó:	Mọi thông tin của trang đều nằm trong vùng IBM WebSphere	Trang chỉ có một mốc tiêu đề.
L1 Các link đều có tên cụ thể, có thể hiểu được ngoài ngữ cảnh (1.0 - 13.1)	2	2
L1 Mô tả cụ thể lỗi đó:	Link để dẫn tới các bài viết đều được đính kèm với tiêu đề của bài nên có thể hiểu. Tuy nhiên, rất nhiều link được nhúng vào các logo và hình minh hoạt đều không được cấu trúc rõ ràng, chỉ đính kèm tên trang web mà nó sẽ liên kết tới. Nhóm link còn lại đều là một chuỗi ký tự không có nghĩa.	Có nhiều link không có tên hoặc có nhưng không rõ ràng, không thể hiểu được ngoài ngữ cảnh như: Rss, XDNET, Tintuc/mimage00320240813211608, Tintuc/m08069, Index...
L2 Các link đều kích hoạt được, và mở ra đúng nội dung mong đợi (1.0 - 13.1)	2	3

home mod.gov.vn TIN TỨC - SỰ KIẾN QĐND VIỆT NAM QUỐC PHÒNG VIỆT NAM QUẢN LÝ - CHỈ ĐẠO HÀNH CHÍNH CÔNG CHẾ ĐỘ - CHÍNH SÁCH THÔNG TIN - DỊCH VỤ English Thứ Bảy, 17/8/2024	Liên kết ở đầu trang "vn" khi kích hoạt vào không có gì hiển thị; Các liên kết như: giới thiệu, tin tức, cải cách hành chính, WebForm_DoPostBackWithOptions(new WebForm_PostBackOptions("ctl00\$ctl43\$lkSearch1", "", true, "", "", false, true)),... không thể kích hoạt được bằng bàn phím; Liên kết Pages/default có thể kích hoạt được nhưng vì nhấn cho liên kết không rõ ràng và khi mở link ra chữ ở liên kết này bị lỗi nhiều, làm cho người dùng rất khó để hiểu được nội dung.
bo-quoc-phong-xet-duyet-diem-c... Bộ Quốc phòng xét duyệt điểm chuẩn tuyển sinh quân sự năm 2024 (Bqp.vn) - Sáng 16/8, tại Hà Nội, Ban Tuyển sinh quân sự (TSQS) Bộ Quốc phòng tổ chức Hội nghị Xét duyệt điểm chuẩn TSQS năm 2024, Thượng tướng Nguyễn Tân Cương, Ủy viên Trung ương Đảng, Ủy viên Thường vụ Quân ủy Trung ương, Tổng Tham mưu trưởng QĐND Việt Nam, Thứ trưởng Bộ Quốc phòng, Trưởng Ban TSQS Bộ Quốc phòng chủ trì hội nghị. Bộ Quốc phòng trả lời cử tri tỉnh Phú Yên phản ánh chế độ cho người tham gia chiến trường Campuchia còn thấp Bộ Quốc phòng trả lời cử tri tỉnh Bình Định kiến nghị chế độ hằng tháng cho người tham gia chiến trường Campuchia đã được chế độ một lần Lễ kỷ niệm lần thứ 78 Ngày Độc lập nước Cộng hòa Ấn Độ Bộ Quốc phòng trả lời cử tri tỉnh Bắc Kạn kiến nghị huấn luyện nghĩa vụ quân sự từ 1978-1982 được bảo hiểm y tế theo Quyết định 62/2011/QĐ-TTg Bộ Quốc phòng trả lời cử tri tỉnh Bắc Giang kiến nghị người hoàn thành nghĩa vụ quân sự từ 30/4/1975 - 31/12/1989 được chế độ theo Quyết định 62/2011/QĐ-TTg Bộ Quốc phòng tặng Bằng khen cho 3 tập thể và 8 cá nhân có thành tích xuất sắc đột xuất trong tìm kiếm cứu nạn, khắc phục hậu quả thiên tai detail current=true&urile=wcm:... Tổng cục Chính trị sơ kết công tác phối hợp xây dựng, hoạt động khu vực phòng thủ detail current=true&urile=wcm:... Bộ Quốc phòng - Bộ Công an gặp mặt kỷ niệm 79 năm Ngày truyền thống Công an nhân dân detail current=true&urile=wcm:... Đẩy nhanh tiến độ triển khai Kết luận của Ban Thường vụ Quân ủy Trung ương và Ban Thường vụ Thành ủy Thành phố Hồ Chí Minh ĐỐI NGOẠI QUỐC PHÒNG detail current=true&urile=wcm:... Lễ kỷ niệm lần thứ 78 Ngày Độc lập nước Cộng hòa Ấn Độ (Bqp.vn) - Tối 16/8, tại Hà Nội, Đại sứ quán Ấn Độ tại Việt Nam long trọng tổ chức Lễ kỷ niệm lần thứ 78 Ngày Độc lập nước Cộng hòa Ấn Độ (15/8/1947 - 15/8/2024). Đại tướng Phan Văn Giang tiếp Tư lệnh Lực lượng Quốc phòng Ma-lai-xi-a Tăng cường hợp tác giữa Cảnh sát biển Việt Nam và Cảnh sát biển Ấn Độ	

L3 Các link cùng phải đến cùng trang (1.0 - 13.1)	2	Không áp dụng
L3 Mô tả cụ thể lỗi đó:	Nhiều link bài viết được đính kèm vào hình ảnh đại diện cho bài nhưng lại không được đính kèm với tiêu đề bài. Thay vào đó, link được gắn với một chuỗi ký tự vô nghĩa. Các link này chỉ được hiểu có cùng liên kết với nội dung bài viết vì đứng bên cạnh các link bài viết.	Không tìm thấy link cùng tên
H1 Tiêu đề (heading) được sử dụng đúng mục tạo cấu trúc nội dung,	1 = không dễ tiếp cận chút nào	5 = hoàn toàn dễ tiếp cận
H1 Mô tả cụ thể lỗi đó:	Trang chỉ có các tiêu đề được dùng để gắn với tên các bài viết thay vì dùng để tạo ra các vùng riêng biệt cho trang.	Tiêu đề là tên của các bài viết.
H2 Tiêu đề (heading) được gắn đúng trình tự logic (1.0 - 3.5)	1 = không dễ tiếp cận chút nào	2
H2 Mô tả cụ thể lỗi đó:	Trang có một H1 là tiêu đề bài viết và sau đó là các H4 và H5 cũng là tiêu đề bài viết. Những tiêu đề này không thực sự tạo ra được ý nghĩa gì về bố cục khi các bài viết không có dấu hiệu liên quan đến nhau.	Trang chỉ có Heading level 2 và 3 mà không có Heading level 1. Heading cũng đồng thời là Link và các bài viết liên tục luân phiên thay đổi nên mỗi lúc nhấn chữ "H" để di chuyển nhanh đến các Heading lại là một tiêu đề của một bài viết khác nhau.
I1 Có thể điều hướng đến các danh sách trên bảng phím tắt L (1.0 -	1 = không dễ tiếp cận chút nào	5 = hoàn toàn dễ tiếp cận
I1 Mô tả cụ thể lỗi đó:	Trang có một số danh sách và đều có thể di chuyển tới. Tuy nhiên, các danh sách không được gắn tiêu đề phù hợp. Thay vào đó, trình đọc màn hình sẽ mặc định đọc mục đầu tiên của danh sách, thường là một link.	Không có lỗi
I2 Tất cả các danh sách (list) được chính xác (1.0 - 3.6)	4	5 = hoàn toàn dễ tiếp cận
I2 Mô tả cụ thể lỗi đó:	Mục trong các danh sách đều có thể được đọc nhưng nó không tạo được ngữ cảnh để người dùng hiểu. Nhiều danh sách chỉ chứa tên bài mà không có tiêu đề phân chia để mô tả mục đích tạo danh sách của trang.	Chưa thấy sai gì.

B1 Có thể điều hướng đến các nút bằng phím tắt B (1.0 - 12.4)	4	Không áp dụng
B1 Mô tả cụ thể lỗi đó:	Trang không có quá nhiều nút, chỉ có một nút tìm kiếm nằm kế bên khung tìm kiếm và các nút điều khiển một video được nhúng trong trang.	Không thấy có nút nào trên trang chủ.
B2 Các nút được gắn nhãn có ý nghĩa (1.0 - 12.4)	4	Không áp dụng
B2 Mô tả cụ thể lỗi đó:	Nút tìm kiếm không có nhãn còn các nút điều khiển video thì khá tiếp cận nhưng được ghi bằng tiếng Anh.	Trang không có nút.
B3 Có thể kích hoạt vào các nút bằng phím khoảng trắng hoặc enter (1.0 - 12.4)	5 = hoàn toàn dễ tiếp cận	Không áp dụng
B3 Mô tả cụ thể lỗi đó:	Hoàn toàn có thể dùng bàn phím để kích hoạt các nút này.	Trang không có nút.
B4 Các thành phần định nghĩa là "nút" một cách hợp lý (1.0 - 12.4)	5 = hoàn toàn dễ tiếp cận	Không áp dụng
B4 Mô tả cụ thể lỗi đó:	Nút tìm kiếm trên trang được định dạng là nút hợp lý.	Trang không có nút.
R1 Có thể điều hướng tới từng nút radio trên trang bằng phím tắt R (với NVDA)	Không áp dụng	Không áp dụng
R1 Mô tả cụ thể lỗi đó:	Trang không có nút radio	Không thấy có nút Radio nào.
R2 Tất cả nút radio được gắn nhãn có ý nghĩa (1.0 - 12.4)	Không áp dụng	Không áp dụng

R2 Mô tả cụ thể lỗi đó:	Trang không có nút radio	Không thấy có nút Radio nào.
R3 Có thể chọn từng nút radio bằng phím khoảng trắng (1.0 - 12.4)	Không áp dụng	Không áp dụng
R3 Mô tả cụ thể lỗi đó:	Trang không có nút radio	Không thấy có nút Radio nào.
R4 Trình đọc màn hình thông báo trạng thái của nút radio là được chọn	Không áp dụng	Không áp dụng
R4 Mô tả cụ thể lỗi đó:	Trang không có nút radio	Không thấy có nút Radio nào.
R5 Tất cả các thành phần được định nghĩa là "Nút radio" một cách hợp lý	Không áp dụng	Không áp dụng
R5 Mô tả cụ thể lỗi đó:	Trang chỉ liệt kê thông tin, không có thành phần nào cần được định dạng là nút radio	Không thấy có nút Radio nào.
C1 Có thể điều hướng đến các combobox trên trang bằng phím tắt C (1.0 -	Không áp dụng	5 = hoàn toàn dễ tiếp cận
C1 Mô tả cụ thể lỗi đó:	Trang không có combobox	Không có lỗi
C2 Có thể di chuyển đến và chọn bất kỳ tùy chọn nào trong combobox (1.0 - 12.4)	Không áp dụng	5 = hoàn toàn dễ tiếp cận
C2 Mô tả cụ thể lỗi đó:	Trang không có combobox	Không có lỗi

C3 Mọi tùy chọn trong combobox đều được đọc chính xác (1.0 - 12.4)	Không áp dụng	5 = hoàn toàn dễ tiếp cận
C3 Mô tả cụ thể lỗi đó:	Trang không có combobox	Không có lỗi
C4 Các combobox đều được gắn nhãn hợp lý (1.0 - 12.4)	Không áp dụng	5 = hoàn toàn dễ tiếp cận
C4 Mô tả cụ thể lỗi đó:	trang không có combobox	Ở trang chủ có một ComboBox có nhãn rõ ràng là "Liên kết website."
X1 Có thể điều hướng tới mọi hộp kiểm (checkbox) trên trang bằng phím	Không áp dụng	Không áp dụng
X1 Mô tả cụ thể lỗi đó:	Trang không có checkbox	Không có CheckBox nào.
X2 Có thể chọn hoặc bỏ chọn các checkbox bằng phím khoảng trắng (1.0 - 12.4)	Không áp dụng	Không áp dụng
X2 Mô tả cụ thể lỗi đó:	Trang không có checkbox	Không có CheckBox nào.
X3 Trình đọc màn hình thông báo kiểm tra có được chọn hay không	Không áp dụng	Không áp dụng
X3 Mô tả cụ thể lỗi đó:	Trang không có checkbox	Không có CheckBox nào.
X4 Các checkbox đều được gắn nhãn có nghĩa (1.0 - 12.4)	Không áp dụng	Không áp dụng

X4 Mô tả cụ thể lỗi đó:	Trang không có checkbox	Không có CheckBox nào.
E1 Có thể điều hướng tới các hộp soạn thảo trên trang bằng phím tắt E (1.0 -	5 = hoàn toàn dễ tiếp cận	5 = hoàn toàn dễ tiếp cận
E1 Mô tả cụ thể lỗi đó:	Trang có một khung tìm kiếm và có thể di chuyển tới khung này bằng phím tắt.	Không có lỗi
E2 Có thể nhập liệu vào các hộp soạn thảo (1.0 - 12.4)	5 = hoàn toàn dễ tiếp cận	5 = hoàn toàn dễ tiếp cận
E2 Mô tả cụ thể lỗi đó:	Có thể nhập liệu vào khung tìm kiếm này một cách dễ dàng.	Không có lỗi
E3 Các hộp soạn thảo có nhãn rõ ràng thể hiện chức năng của chúng (1.0 -	5 = hoàn toàn dễ tiếp cận	5 = hoàn toàn dễ tiếp cận
E3 Mô tả cụ thể lỗi đó:	Khung tìm kiếm này được trình đọc màn hình phát âm "tìm kiếm nội dung" khi di chuyển tới nó.	Không có lỗi
A1 Có thể điều hướng tới mọi bảng trên trang bằng phím tắt T (1.0 - 5.1)	5 = hoàn toàn dễ tiếp cận	5 = hoàn toàn dễ tiếp cận
A1 Mô tả cụ thể lỗi đó:	Trang có một bảng liệt kê các văn bản và ngày phát hành.	Không có lỗi
A2 Tất cả các tiêu đề bảng được chỉnh xác (1.0 - 5.1)	5 = hoàn toàn dễ tiếp cận	1 = không dễ tiếp cận chút nào
A2 Mô tả cụ thể lỗi đó:	Bảng được đặt tiêu đề ở hàng đầu tiên phù hợp	Bảng chỉ được đọc với tiêu đề là "Một số kí hiệu"

A3 Tiêu đề cột và dòng được nhắc lại cho từng ô (1.0 - 5.1)	5 = hoàn toàn dễ tiếp cận	1 = không dễ tiếp cận chút nào
A3 Mô tả cụ thể lỗi đó:	Tiêu đề của mỗi cột được đọc lại khi di chuyển giữa các ô	Khi di chuyển qua từng ô thì chỉ đọc được nội dung của ô đó mà không rõ tiêu đề cột đó là gì.
A4 Bảng không được sử dụng cho mục đích định dạng (1.0 - 5.3)	5 = hoàn toàn dễ tiếp cận	5 = hoàn toàn dễ tiếp cận
A4 Mô tả cụ thể lỗi đó:	Bảng này được dùng để liệt kê thông tin.	Không có lỗi
N1 Trang có thanh điều hướng đồng nhất với các trang khác (1.0 13.5)	5 = hoàn toàn dễ tiếp cận	4
N1 Mô tả cụ thể lỗi đó:	Thanh điều hướng của trang luôn đứng ở vị trí đầu trang khi đi vào các trang con hoặc vào các bài đăng trong mục tin tức.	Hầu hết các trang khác đều có thanh điều hướng đồng nhất với trang chủ, chỉ có liên kết "Dịch vụ công trực tuyến" là không đồng nhất. Khi mở liên kết "Dịch vụ công trực tuyến" thì tiêu đề trang vẫn là Bộ Xây dựng.
N2 Trình độ đọc màn hình có thể đọc được nhãn của các mục trong thanh	4	5 = hoàn toàn dễ tiếp cận
N2 Mô tả cụ thể lỗi đó:	Tất cả các link đều có thể tiếp cận trừ mục trang chủ. Mục trang chủ chỉ được gắn nhãn là tên trang web mod.gov.vn	Không có lỗi.
G1 Có thể tiếp cận văn bản thay thế của các hình ảnh trên trang	1 = không dễ tiếp cận chút nào	2
G1 Mô tả cụ thể lỗi đó:	Trang sử dụng nhiều hình ảnh minh họa cho các bài viết nhưng lại không định dạng thuoojc tính đúng cho các hình ảnh này. Khi dùng G để tìm kiếm thì trang không hề có bất kỳ hình ảnh nào.	Các hình ảnh trên trang đa phần được đính kèm một liên kết, liên kết đó có nhãn cũng không thực sự rõ ràng và trước khi đọc nhãn của liên kết thì trình đọc màn hình chỉ nhận diện và đọc là "hình ảnh không có nhãn".

S1 Thanh tìm kiếm xuất trong trang và có thể được di chuyển đến	5 = hoàn toàn dễ tiếp cận	5 = hoàn toàn dễ tiếp cận
S1 Mô tả cụ thể lỗi đó:	Hoàn toàn tiếp cận được với thanh tìm kiếm	Không có lỗi
S2 Có thể tìm kiếm nhiều bằng cách (có không dấu, thiếu chữ, sai	1 = không dễ tiếp cận chút nào	2
S2 Mô tả cụ thể lỗi đó:	Không thể tìm kiếm bất kỳ nội dung nào khi gõ có dấu, không dấu hoặc gõ ký tự ngẫu nhiên. Thậm chí dùng các từ khóa từ tiêu đề các bài viết trên trang chủ cũng không trả về kết quả nào.	Khi gõ từ khóa tìm kiếm có dấu và đầy đủ các từ của một bài viết, một liên kết... thì mới hiện ra kết quả hoặc nếu gõ không dấu nhưng chỉ một vài từ đầu tiên (ví dụ như: "khanh thanh") thì cũng có hiện ra kết quả nhưng vì từ khóa chung chung nên sẽ hiện ra rất nhiều kết quả; còn nếu gõ không dấu, gõ thiếu chữ... với một từ khóa dài và không hoàn toàn đầy đủ như tên một bài viết, một liên kết nào đó thì sẽ không hiển thị kết quả. Nghĩa là gần như phải gõ từ khóa đúng như tiêu đề của bài viết, liên kết... thì mới cho ra kết quả.
S3 Kết quả tìm kiếm có thể hiện một cách dễ tiếp cận hay không (1.0	1 = không dễ tiếp cận chút nào	1 = không dễ tiếp cận chút nào
S3 Mô tả cụ thể lỗi đó:	Không thể tìm kiếm bất kỳ nội dung nào khi gõ có dấu, không dấu hoặc gõ ký tự ngẫu nhiên. Thậm chí dùng các từ khóa từ tiêu đề các bài viết trên trang chủ cũng không trả về kết quả nào. Một điều thú vị là trang được tải sau khi chọn tìm kiếm xuất hiện một khung tìm kiếm khác và một nút tìm kiếm với nhãn rõ ràng. Tuy nhiên, sử dụng khung tìm kiếm này cũng không mang lại kết quả khả quan hơn.	Kết quả tìm kiếm không được hiển thị dưới dạng là một tiêu đề, một bảng... nên phải dùng các phím mũi tên di chuyển lần lượt qua từng mục để đi đến, hoặc nếu như người dùng đã quen với bố cục của trang thì có thể di chuyển nhanh tới một thành phần nào đó ở gần kết quả tìm kiếm rồi dùng phím mũi tên di chuyển đến.
O1 Trang tự động chuyển nếu kích hoạt các link	5 = hoàn toàn dễ tiếp cận	1 = không dễ tiếp cận chút nào
O1 Mô tả cụ thể lỗi đó:	Trang không có các yếu tố động, không bị chuyển trang ngẫu nhiên.	Ở trang chủ có một ComboBox "Liên kết các website", khi dùng mũi tên duyệt qua danh sách rồi bấm Esc để thoát ra khỏi danh sách thì trang tự động chuyển sang một tab mới mà không có bất kì dấu hiệu nào thông báo cho người dùng biết đang chuyển sang trang mới.

O2 Trình đọc màn hình có thể tiếp cận với các thông báo lỗi khi nhập thông tin	Không áp dụng	Không áp dụng
O2 Mô tả cụ thể lỗi đó:	Trang không có mục điền thông tin nên không thể kiểm tra lỗi	Trang không có yêu cầu nhập thông tin.
O3 Có cảnh báo rõ ràng khi bất kỳ thao tác nào làm chuyển trang (1.0 -	Không áp dụng	1 = không dễ tiếp cận chút nào
O3 Mô tả cụ thể lỗi đó:	Dạng cảnh báo này chỉ phù hợp trong các trường hợp trang đang thu thập thông tin người dùng và cần đảm bảo thông tin được thu thập đầy đủ. Với dạng trang cung cấp tin tức như thể này thì việc tích hợp cơ chế cảnh báo là không cần thiết.	Ở trang chủ có một ComboBox "Liên kết các website", khi dùng mũi tên duyệt qua danh sách rồi bấm Esc để thoát ra khỏi danh sách thì trang tự động chuyển sang một tab mới mà không có bất kì dấu hiệu nào thông báo cho người dùng biết đang chuyển sang trang mới.
Một số đánh giá và nhận xét khác	Trang cần xây dựng lại từ việc cân nhắc thông tin nào cần thể hiện trên trang để phân chúng ra thành các nhóm thích hợp. Sau đó các link cần được định dạng và gắn nhãn phù hợp để tránh việc link được đặt loạn trong trang. Rất nhiều hình ảnh được dùng nhưng lại không được gắn thuộc tính là hình ảnh. Trang sẽ trở nên dễ hiểu hơn nếu heading được dùng đúng.	Các bài viết trên trang liên tục luân phiên thay đổi vì vậy sẽ rất khó để xác định được vị trí, rất khó để có thể duyệt được qua nhiều bài viết để tìm được tin tức mong muốn. Trong khi khả năng đưa ra kết quả tìm kiếm của trang lại không được tốt.

Danh sách đầy đủ các bản ghi bộ dữ liệu gốc và file dữ liệu gốc được nộp kèm đề án dưới dạng file Excel.



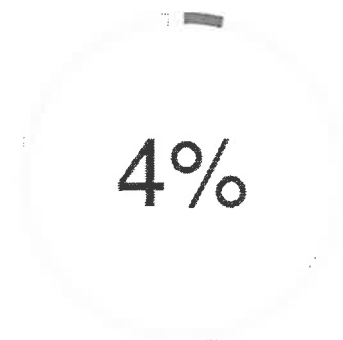
BÁO CÁO KIỂM TRA TRÙNG LẶP

Thông tin tài liệu

Tên tài liệu: PTIT_MIS_Luan Van_KhoaPD_v20250802
Tác giả: Phí Đăng Khoa
Điểm trùng lặp: 4
Thời gian tải lên: 15:10 02/08/2025
Thời gian sinh báo cáo: 15:12 02/08/2025
Các trang kiểm tra: 77/77 trang



Kết quả kiểm tra trùng lặp



Có 4% nội dung trùng
lặp



Có 96% nội
dung không
trùng lặp



Có 0% nội dung
người dùng loại
trừ



Có 0% nội dung
hệ thống bỏ qua

PGS.TS Đỗ Trung Tuấn

Phí Đăng Khoa

Nguồn trùng lặp tiêu biểu

123docz.net www.w3.org tailieu.vn

BÁO CÁO GIẢI TRÌNH
SỬA CHỮA, HOÀN THIỆN ĐỀ ÁN TỐT NGHIỆP

Họ và tên học viên: Phí Đăng Khoa

Chuyên ngành: Hệ thống thông tin

Khóa: 2023 đợt 2

Tên đề tài: Nghiên cứu công cụ tóm tắt văn bản về mô tả lỗi khi kiểm thử khả năng tiếp cận web

Người hướng dẫn khoa học: PGS.TS Đỗ Trung Tuấn

Ngày bảo vệ: 19/07/2025

Các nội dung học viên đã sửa chữa, bổ sung trong đề án tốt nghiệp theo ý kiến đóng góp của Hội đồng chấm đề án tốt nghiệp:

TT	Ý kiến hội đồng	Sửa chữa của học viên
1	Chỉnh sửa lỗi soạn thảo, lỗi ngữ pháp, chính tả	Học viên đã rà soát, chỉnh sửa các lỗi soạn thảo, các lỗi ngữ pháp
2	Chỉnh sửa phần danh mục ký hiệu, chữ viết tắt	Tiếp thu góp ý của Hội đồng, học viên đã cập nhật nội dung mô tả rõ nghĩa
3	Giải thích rõ hơn các tiêu chí trong bộ hướng dẫn WCAG	Tiếp thu góp ý của Hội đồng, học viên đã bổ sung nội dung ở mục 1.3 thuộc Chương I
4	Thiếu mô tả về xử lý tham số trong OpenAI GPT model	Tiếp thu góp ý của Hội đồng, học viên đã bổ sung nội dung về tinh chỉnh tham số trong OpenAI GPT model như temperature, instruction, ở mục hướng phát triển thuộc Kết luận và Kiến nghị

Hà Nội, ngày 4 tháng 8 năm 2025

Ký xác nhận của

CHỦ TỊCH HỘI ĐỒNG
CHẤM ĐỀ ÁN

TS. Nguyễn Duy Phương

THƯ KÝ HỘI ĐỒNG

TS. Đặng Hoàng Long

NGƯỜI HƯỚNG
DẪN KHOA HỌC

PGS.TS Đỗ Trung Tuấn

HỌC VIÊN

Phí Đăng Khoa

Phản biện 1: TS. Vũ Văn Thoa:
 - Nếu chỉ tập hệ thuật prompt & đã sử dụng trong đề án.
 - Nếu rõ hơn chỉ đánh giá tốt quá.
 Phản biện 2: PG.S. Trần Thị Danh:
 Tìm hiểu một mô hình local computer để tập luyện chi phí
 7. Trả lời của học viên:
 chạy
 trên

~~PG.TS. Trần Thu Oanh~~
~~Tìm kiếm và ứng dụng có thể chạy trên local~~
~~computer để tìm kiếm chi phí~~

Học viên phải đang khoa đã tốt nghiệp đang sẽ sáng cái lý
thuyết prompt đã sử dụng và đưa ra được phản ứng trên nhà
tốt kiếm chi phí

8. Thư ký đọc nhận xét về quá trình thực hiện đề án tốt nghiệp của học viên (có văn bản kèm theo).

9. Hội đồng hợp riêng:

- Bầu Ban kiểm phiếu:

1. Trưởng Ban kiểm phiếu: TS. Nguyễn Duy Phương
2. Ủy viên Ban kiểm phiếu: TS. Đặng Hoàng Long
3. Ủy viên Ban kiểm phiếu: TS. Vũ Văn Thảo

- Hội đồng chấm đề án tốt nghiệp bằng bỏ phiếu kín.

- Ban kiểm phiếu làm việc:

- Trưởng Ban kiểm phiếu báo cáo kết quả kiểm phiếu (có Biên bản họp Ban kiểm phiếu kèm theo)

- Điểm trung bình của đề án tốt nghiệp: 8,5

Kết luận:

1. Các nội dung cần chỉnh sửa, hoàn thiện sau bảo vệ đề án tốt nghiệp:

Học viên cần chỉnh sửa luận văn theo góp ý của
Hội đồng.
Cần nêu định nội dung luận văn chi tiết, rõ ràng hơn.

2. Đề nghị Học viện công nhận (hoặc không) và cấp bằng (hoặc không) thạc sĩ cho học viên:

Đề nghị học viện công nhận và cấp bằng thạc sĩ cho học viên

3. Đề án tốt nghiệp có thể phát triển thành đề tài nghiên cứu cho

NCS. Không

Buổi làm việc kết thúc vào 12h30... cùng ngày.

Chủ tịch

Phương

TS. Nguyễn Duy Phương

Thư ký

Long

TS. Đặng Hoàng Long

BẢN NHẬN XÉT ĐỀ ÁN TỐT NGHIỆP THẠC SĨ

Tên đề tài đề án: Nghiên cứu công cụ tóm tắt văn bản về mô tả lỗi khi kiểm thử khả năng tiếp cận Web

Chuyên ngành: Hệ thống thông tin

Mã số: 8.48.01.04

Tên học viên: Phí Đăng Khoa

Họ tên người nhận xét: Vũ Văn Thóa

Học hàm, học vị: Tiến sĩ **Chuyên ngành:** Toán học tính toán

Cơ quan công tác: Học viện Công nghệ Bưu chính Viễn thông

NỘI DUNG NHẬN XÉT

I/ Cơ sở khoa học và thực tiễn, tính cấp thiết của đề tài

Trong bối cảnh Việt Nam đang triển khai mạnh mẽ quá trình chuyển đổi số, khả năng tiếp cận các dịch vụ số của nhóm người khuyết tật trở thành vấn đề được quan tâm giải quyết của nhà nước và các doanh nghiệp cung cấp dịch vụ số. Hầu hết các dịch vụ số được cung cấp dưới dạng Website. Do đó, vấn đề kiểm thử và đánh giá khả năng tiếp cận Web của các nhóm người yếu thế trong xã hội theo các bộ hướng dẫn quốc tế có vai trò hết sức quan trọng. Do đặc thù của nhóm người khuyết tật, quá trình kiểm thử và đánh giá khả năng tiếp cận Web do các đối tượng khuyết tật thực hiện cần có các công cụ tự động hỗ trợ.

Đề án với tiêu đề “Nghiên cứu công cụ tóm tắt văn bản về mô tả lỗi khi kiểm thử khả năng tiếp cận Web” của học viên Phí Đăng Khoa tập trung nghiên cứu ứng dụng mô hình AI GPT để tạo công cụ tóm tắt văn bản về mô tả lỗi khi kiểm thử khả năng tiếp cận Web. Vì vậy, đề tài đề án có tính cấp thiết, có ý nghĩa khoa học và thực tiễn.

II/ Về nội dung chất lượng của đề án, những kết quả đạt được

Đề án được trình bày trong 51 trang bao gồm phần mở đầu, 3 chương nội dung, phần kết luận và danh mục các tài liệu tham khảo. Ngoài ra đề án còn có 3 phụ lục mô tả các dữ liệu liên quan đến nội dung đề án.

Chương 1 của đề án gồm 6 trang, trình bày tổng quan các kiến thức cơ sở về khả năng tiếp cận Web và các vấn đề liên quan. Tuy nhiên nội dung của chương này khá sơ sài và bố cục chưa thật hợp lý. Học viên cần hiệu chỉnh chương này theo hướng bổ sung các nội dung về bộ hướng dẫn WCAG 2.1AA, các vấn đề về kiểm thử và đánh giá khả năng tiếp cận Web theo WCAG. Trên cơ sở đó, trình bày bài toán cần giải quyết của đề án. Mục 1.4 Kết luận nên hiệu chỉnh thành mục Kết luận chương 1.

Nội dung chính của chương 2 là nghiên cứu các kỹ thuật tóm tắt văn bản và đề xuất phương pháp tóm tắt văn bản về mô tả lỗi khi kiểm thử khả năng tiếp cận Web được trình bày

trong 11 trang. Đề án đã trình bày khá sơ sài về một số kỹ thuật truyền thống và mô hình ngôn ngữ lớn để giải quyết bài toán tóm tắt văn bản. Để lựa chọn mô hình AI GPT để giải quyết bài toán đặt ra trong đề án cần bổ sung các nội dung về đặc điểm của các văn bản về mô tả lỗi khi kiểm thử khả năng tiếp cận Web và các yêu cầu cần đạt của các tóm tắt. Mục 2.4 Kết luận nên hiệu chỉnh thành mục Kết luận chương 2.

Chương 3 của đề án gồm 26 trang, trình bày các nội dung triển khai thử nghiệm tóm tắt văn bản về mô tả lỗi khi kiểm thử khả năng tiếp cận Web sử dụng mô hình AI GPT và các vấn đề liên quan. Để hợp lý hơn, các nội dung của mục 3.1 về giới thiệu dự án đánh giá khả năng tiếp cận 20 công thông tin điện tử nên chuyển về chương 1. Học viên cần bổ sung các nội dung mô tả các đặc điểm của bộ dữ liệu kiểm thử, các tiêu chí đánh giá kết quả tóm tắt.

Đề án được chia thành các chương mục rõ ràng và cơ bản bám sát các nội dung theo đề cương đã được phê duyệt. Tuy nhiên, đề án chưa làm rõ các đóng góp của học viên trong giải quyết vấn đề đặt ra. Vì vậy, phần kết luận của đề án cần được hiệu chỉnh và trình bày chi tiết hơn về các kết quả đạt được và hướng phát triển tiếp theo của đề án. Phần mở đầu của đề án cần hiệu chỉnh nội dung và hình thức theo quy định của Học viện và kết quả đạt được của đề án. Nên bổ sung các nội dung liên quan đến các hình trong đề án thì hợp lý hơn.

Trong đề án còn có khá nhiều lỗi in ấn phải sửa chữa tại các trang 5, 21, 31, 43, 45, 46, 50, ... Phần Danh mục tài liệu tham khảo cần được trình bày theo đúng qui định của Học viện.

III/ Những vấn đề cần giải thích thêm

- (1) Giải thích rõ hơn các tiêu chí về khả năng tiếp cận Web trong bộ hướng dẫn WCAG 2.1AA được sử dụng trong đề án?
- (2) Nêu chi tiết kỹ thuật prompt trong GPT 4.1 để tóm tắt văn bản trình bày trong hình II.1 (trang 15) của đề án.
- (3) Giải thích rõ hơn các bước trong quy trình tóm tắt văn bản mô tả lỗi kiểm thử khả năng tiếp cận Web theo WCAG trình bày tại hình III.6 (trang 34). Đồng thời, nêu rõ tiêu chí đánh giá kết quả tóm tắt văn bản mô tả lỗi kiểm thử sử dụng trong đề án?

IV/ Kết luận

Tôi đồng ý cho học viên Phí Đăng Khoa được bảo vệ đề án với đề tài “Nghiên cứu công cụ tóm tắt văn bản về mô tả lỗi khi kiểm thử khả năng tiếp cận Web” trước Hội đồng chấm đề án tốt nghiệp thạc sĩ của Học viện. Tuy nhiên, để đảm bảo chất lượng đề án tốt nghiệp thạc sĩ chuyên ngành Hệ thống thông tin, học viên cần chỉnh sửa đề án và giải trình rõ trước Hội đồng các vấn đề được góp ý trên đây.

Ngày 14 tháng 07 năm 2025

Người nhận xét



TS Vũ Văn Thỏa

CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM

Độc lập – Tự do – Hạnh phúc

BẢN NHẬN XÉT ĐỀ ÁN TỐT NGHIỆP THẠC SĨ
(Dùng cho người phản biện)

Tên đề tài đề án tốt nghiệp: Nghiên cứu công cụ tóm tắt văn bản về mô tả lỗi khi kiểm thử khả năng tiếp cận WEB.

Chuyên ngành: Hệ thống thông tin

Mã chuyên ngành: 8.48.01.04

Họ và tên học viên: Phí Đăng Khoa

Họ và tên người nhận xét: Trần Thị Oanh

Học hàm, học vị: PGS.TS

Chuyên ngành: Khoa học máy tính

Cơ quan công tác: Trường Quốc tế - ĐHQG Hà Nội

Số điện thoại: 0362220684

E-mail: tranthioanh@vnu.edu.vn

NỘI DUNG NHẬN XÉT

I/ Cơ sở khoa học và thực tiễn, tính cấp thiết của đề tài:

Đề tài bám sát thực trạng hạn chế khả năng tiếp cận số tại Việt Nam, có số liệu và minh chứng cụ thể. Điều này cho thấy nhu cầu nghiên cứu và ứng dụng giải pháp là hoàn toàn có thật và cần thiết, đặc biệt khi pháp lý còn thiếu hụt.

Đề tài dựa trên nền tảng các tiêu chuẩn quốc tế (WCAG) và các hướng nghiên cứu về tóm tắt văn bản, mô hình ngôn ngữ lớn (LLM) — có hệ thống, bám sát xu hướng công nghệ hiện nay. Giải pháp tập trung vào quy trình và công cụ có thể áp dụng ngay, phù hợp nguồn lực hạn chế, giúp phát huy vai trò người khiếm thị trong kiểm thử — đối tượng rất cần được trao quyền trong bối cảnh tiếp cận số.

II/ Nội dung của đề án tốt nghiệp, các kết quả đã đạt được:

Đề án tốt nghiệp đã thực hiện:

- + Xây dựng quy trình kiểm thử khả năng truy cập web cho người khiếm thị,
- + Thử nghiệm công cụ tóm tắt tự động mô tả lỗi dựa trên GPT

III/ Những vấn đề cần giải thích thêm:

- Cần giải thích rõ ý nghĩa trong bảng viết tắt: Vd LSA cần dịch sang tiếng Việt, sau đó mới giải thích ý nghĩa.

- Các mục của chương cần căn trái, ko căn giữa như hiện tại.
- Về quy trình thu thập dữ liệu:
 - + Tính đại diện: Các trang được chọn có đủ phản ánh sự đa dạng của công thông tin điện tử không? Ví dụ các dạng trang chuyên biệt như form đăng ký, tra cứu kết quả có được kiểm thử không?
 - + Khi 2 cộng tác viên cho kết quả khác nhau trên cùng trang, có đối chiếu không?
 - + Có bước hội ý cộng tác viên để đồng nhất cách chấm điểm không vì việc phân chia mức 1-5 với các CTV có thể hiểu khác nhau
- Về đánh giá chất lượng tóm tắt:
 - + Đề án chưa chỉ ra cách đo lường chất lượng bản tóm tắt so với tóm tắt chuyên gia: có dùng chỉ số nào không (BLEU, ROUGE, hay chỉ chấm điểm)?
 - + Thiếu mô tả về xử lý prompt và tham số: Phần mô tả prompt còn sơ sài. Chưa rõ prompt cấu trúc ra sao để đảm bảo GPT hiểu ngữ cảnh kiểm thử accessibility, tránh trả ra văn bản chung chung. Ngoài ra, các tham số như temperature, max_tokens, hệ thống instruction nên được nêu để chứng minh học viên có điều chỉnh tối ưu.
 - + Rủi ro phụ thuộc hạ tầng bên thứ ba: Việc chạy qua Azure OpenAI giúp dễ triển khai nhưng có rủi ro chi phí token biến động, quota giới hạn hoặc API hết free. Nên đề xuất hướng mở như: deploy mô hình local (nếu quy mô lớn), hay tối ưu chi phí bằng batching request.
 - + Chưa có minh chứng kiểm thử thực tế: Có khảo sát cộng tác viên hoặc chuyên gia so sánh output GPT với tóm tắt do con người làm?

Câu hỏi:

1. Tóm tắt GPT có xử lý tốt các mô tả lỗi đa dạng, không đồng nhất, có câu dài lủng củng không? Bạn đã có test case nào thể hiện hạn chế?
2. Bạn dự định mở rộng công cụ thế nào để không phụ thuộc 100% vào GPT thương mại (trong trường hợp tài chính bị cắt)?

IV/ Kết luận:

Đồng ý cho phép học viên bảo vệ đề án tốt nghiệp. Tuy nhiên, học viên cần làm rõ thêm các ý được nêu ở mục III.

Ngày tháng 7 năm 2025

NGƯỜI NHẬN XÉT



Trần Thị Oanh

