

TRỢ LÝ AI TRONG HÀNH CHÍNH CÔNG VIỆT NAM: KHẢO LƯỢC CÓ HỆ THỐNG VÀ KHUYẾN NGHỊ TRIỂN KHAI

● **VƯƠNG THÚY NGỌC¹ - NGUYỄN VĂN TIẾN SƠN^{2*}**
- NGÔ MINH VŨ³ - NGUYỄN HỮU HUÂN⁴

¹Công ty Cổ phần Công nghệ Promete, TP. Hồ Chí Minh, Việt Nam

²Đại học Kinh tế TP. Hồ Chí Minh - UEH Mekong, Vĩnh Long, Việt Nam

³Đại học Kinh tế TP. Hồ Chí Minh, TP. Hồ Chí Minh, Việt Nam

⁴Đại học Kinh tế TP. Hồ Chí Minh, TP. Hồ Chí Minh, Việt Nam

* Email tác giả liên hệ: sonnvu@ueh.edu.vn

TÓM TẮT:

Chuyển đổi số và chiến lược AI của Việt Nam tạo áp lực triển khai trợ lý AI hỗ trợ dịch vụ công, kèm rủi ro về quyền riêng tư và an ninh mạng. Bài viết khảo lược hệ thống 25 nghiên cứu quốc tế theo chuẩn PRISMA 2020 nhằm tổng hợp bằng chứng về (i) yếu tố ảnh hưởng chấp nhận/niềm tin, (ii) nguyên tắc thiết kế, (iii) mô hình quản trị khi triển khai trợ lý AI trong khu vực công; từ đó đề xuất khuyến nghị cho Việt Nam. Kết quả nhấn mạnh: truyền thông mục tiêu tăng niềm tin; thiết kế theo “vai công chức” và giá trị công; quản trị AI nội bộ là điều kiện tiên quyết; và khoảng trống đánh giá hệ thống đã vận hành bằng thước đo giá trị công. Triển khai cần lộ trình thí điểm có kiểm soát, quản trị dữ liệu cá nhân theo pháp luật, cơ chế human-in-the-loop và bộ chỉ số đa mục tiêu.

Từ khóa: trợ lý AI, chatbot chính phủ, hành chính công, bảo vệ dữ liệu cá nhân, quản trị AI, PRISMA.

1. Đặt vấn đề

Việt Nam đang thúc đẩy mạnh mẽ chuyển đổi số trong khu vực công thông qua Chương trình Chuyển đổi số quốc gia và Chiến lược phát triển Chính phủ điện tử hướng tới Chính phủ số giai đoạn 2021-2025 (Thủ tướng Chính phủ Việt Nam, 2021b), song song với Chiến lược quốc gia về nghiên cứu, phát triển và ứng dụng Trí tuệ nhân tạo đến năm 2030 (Thủ tướng Chính phủ Việt Nam, 2021a). Trong bối cảnh đó, các trợ lý AI - bao gồm chatbot, virtual assistant và gần

đây là các mô hình ngôn ngữ lớn (LLM) - nổi lên như một giải pháp giao diện triển vọng, giúp cơ quan nhà nước giảm tải tiếp nhận thông tin, phân luồng yêu cầu và hỗ trợ hướng dẫn thủ tục hành chính (Cortés - Cediél et al., 2023).

Tuy nhiên, ứng dụng trợ lý AI trong tương tác công dân - nhà nước không chỉ là bài toán kỹ thuật. Đây còn là thách thức về giá trị công, niềm tin, tính chính danh và trách nhiệm giải trình (König & Wenzelburger, 2021). Nghiên cứu thực nghiệm cho

thấy, niềm tin của công dân vào chatbot khu vực công phụ thuộc nhiều vào lĩnh vực yêu cầu và cách cơ quan truyền thông mục tiêu triển khai (Aoki, 2020). Ở chiều ngược lại, các nghiên cứu học thuật cũng chỉ ra, nghiên cứu hiện nay còn thiên về mô tả, thiếu bằng chứng thực nghiệm gắn chặt bối cảnh khu vực công và chưa có thước đo đánh giá giá trị công đủ toàn diện (Zuiderwijk et al., 2021; Cortés - Cediell et al., 2023).

Từ khoảng trống đó, bài viết đặt ra 4 câu hỏi nghiên cứu phục vụ triển khai trợ lý AI trong hành chính công Việt Nam: (RQ1) Trợ lý AI trong khu vực công hiện tập trung vào những chức năng nào và tác động đến giá trị công ra sao? (RQ2) Các yếu tố nào ảnh hưởng đến niềm tin của công dân và sự chấp nhận của công chức? (RQ3) Những nguyên tắc thiết kế nào - identity, public value - by - design, human fallback - được chứng minh là quan trọng? (RQ4) Quản trị AI nên được cấu trúc thế nào để giảm rủi ro pháp lý - đạo đức và tăng tính chính danh tại Việt Nam? Để trả lời, bài viết thực hiện khảo lược có hệ thống 25 nghiên cứu quốc tế theo chuẩn PRISMA 2020, đồng thời đối chiếu với khung pháp lý hiện hành của Việt Nam (Larsen & Folstad, 2024).

2. Đối tượng và phương pháp nghiên cứu

Nghiên cứu áp dụng phương pháp khảo lược có hệ

thống (SLR) theo chuẩn PRISMA 2020, với quy trình năm bước: nhận diện → loại trùng → sàng lọc tiêu đề/tóm tắt → đánh giá toàn văn → đưa vào tổng hợp (Page et al., 2021). Nguồn dữ liệu chính gồm Scopus và Web of Science, bổ trợ bằng ScienceDirect và Google Scholar; khung thời gian 2015-2025, mở rộng trọng tâm 2022-2025 cho nhóm LLM và generative AI (Pesch, 2025). Chuỗi truy vấn kết hợp 3 nhóm: (A) trợ lý hội thoại (chatbot, virtual assistant...); (B) LLM/generative AI (large language model, ChatGPT...); (C) khu vực công (government, public administration, e-government...), theo cú pháp (A OR B) AND C. Nhóm văn bản pháp lý Việt Nam được bổ sung để đối chiếu khuyến nghị với yêu cầu tuân thủ thực tế (Chính phủ Việt Nam, 2023).

Tiêu chí đưa vào gồm: bài báo peer-reviewed tập trung vào chatbot/AI trong khu vực công, có dữ liệu thực nghiệm hoặc khung phân tích quản trị, liên quan trực tiếp đến dịch vụ công hoặc tương tác công dân-nhà nước (Aoki, 2020). Loại trừ các bài thuần mô tả kỹ thuật thương mại, không kiểm chứng được nguồn, hoặc case study không có phản biện học thuật (Cortés-Cediell et al., 2023). Sau sàng lọc, 25 bài báo được đưa vào tổng hợp; 10 nghiên cứu trọng tâm được trình bày tại Bảng 1.

Bảng 1. Tổng hợp 10 nghiên cứu trọng tâm về trợ lý AI trong hành chính công Việt Nam

Tác giả (năm)	Tiêu đề	Phát hiện chính
Aoki (2020)	An Experimental Study of Public Trust in AI Chatbots in the Public Sector	Niềm tin công dân vào chatbot nhà nước không đồng nhất theo lĩnh vực; truyền thông mục tiêu có thể cải thiện niềm tin ban đầu.
Larsen & Følstad (2024)	The Impact of Chatbots on Public Service Provision: A Qualitative Interview Study with Citizens and Public Service Providers	Chatbot cải thiện tiếp cận dịch vụ và tạo hiểu biết sâu hơn về nhu cầu công dân thông qua dữ liệu tương tác.
Li & Wang (2024)	Should Government Chatbots Behave Like Civil Servants? The Effect of Chatbot Identity Characteristics on Citizen Experience	Đặc tính “công chức số” (lịch sự, công bằng, chuyên nghiệp) ảnh hưởng trực tiếp đến trải nghiệm và niềm tin của công dân.
Hemesath & Tepe (2024)	Public Value Positions and Design Preferences Toward AI - Based Chatbots in E - Government: Evidence from a Conjoint Experiment with Citizens and Municipal Front Desk Officers	Công dân và công chức ưu tiên nhà phát triển nội địa, giới hạn tự quyết chatbot trong kịch bản thường lệ, và can thiệp người khi hệ thống thất bại.

Tác giả (năm)	Tiêu đề	Phát hiện chính
Ahn & Chen (2022)	Digital Transformation Toward AI - Augmented Public Administration: The Perception of Government Employees and the Willingness to Use AI in Government	"Buy - in" của công chức phụ thuộc nhận thức lợi/hại dài hạn, tác động việc làm, đạo đức AI và kinh nghiệm thực tế.
Cortés - Cediel et al. (2023)	Trends and Challenges of E - Government Chatbots: Advances in Exploring Open Government Data and Citizen Participation Content	Chatbot e - government tập trung tra cứu thông tin; thiếu đánh giá hệ thống đã triển khai và thước đo giá trị công.
Busuioc (2021)	Accountable Artificial Intelligence: Holding Algorithms to Account	Tăng trách nhiệm giải trình AI; ưu tiên mô hình dễ giải thích; gắn kết account - giving trong quy trình hành chính.
König & Wenzelburger (2021)	The Legitimacy Gap of Algorithmic Decision - Making in the Public Sector: Why It Arises and How to Address It	Ra quyết định thuật toán tạo "khoảng trống chính danh"; cần giám sát đa bên và kênh tham gia của công dân.
Pesch (2025)	Potentials and Challenges of Large Language Models (LLMs) in the Context of Administrative Decision - Making	LLM hữu ích trong dự thảo văn bản nhưng có thách thức pháp lý; rủi ro "memorisation" và thiên kiến đầu ra.
de Almeida & dos Santos Júnior (2025)	Artificial Intelligence Governance: Understanding How Public Organizations Implement It	Đào tạo stakeholder và chuẩn hóa quy trình nâng mức trưởng thành quản trị AI; khung AIGov4Gov.

Nguồn: Tổng hợp từ nhóm tác giả

3. Kết quả nghiên cứu

Tổng hợp từ 25 nghiên cứu học thuật cho thấy, bằng chứng tập trung thành 5 cụm chủ đề liên kết, phản ánh những điều kiện cốt lõi để triển khai trợ lý AI hiệu quả và chính danh trong hành chính công.

Chức năng trợ lý AI và giá trị công: Chatbot khu vực công hiện chủ yếu đóng vai trò "tuyến 0" - cung cấp thông tin, phân luồng yêu cầu và giảm tải tiếp nhận. Tiềm năng tạo giá trị công thực sự nằm ở khả năng chuyển hóa dữ liệu tương tác thành tri thức cải tiến dịch vụ, với AI được định vị là công cụ hỗ trợ quyết định chứ không thay thế phán đoán của con người (Larsen & Følstad, 2024).

Niềm tin và chấp nhận của công dân: Niềm tin của công dân vào chatbot nhà nước mang tính phi đồng nhất, phụ thuộc vào lĩnh vực dịch vụ, và có thể

được củng cố thông qua truyền thông rõ ràng về mục tiêu triển khai hướng đến lợi ích công dân (Aoki, 2020). Về thiết kế, chatbot cần mang đặc tính "công chức số" - lịch sự, công bằng, chuyên nghiệp - thay vì lối giao tiếp máy móc thuần túy (Li & Wang, 2024).

"Buy - in" của công chức và năng lực tổ chức: Sự chấp nhận của đội ngũ công chức là điều kiện tiên quyết để triển khai AI thành công, được quyết định bởi nhận thức lợi ích lâu dài, lo ngại rủi ro việc làm và quan điểm đạo đức AI (Ahn & Chen, 2022). Năng lực quản trị AI của tổ chức phụ thuộc vào chất lượng đào tạo nhân sự và mức độ chuẩn hóa quy trình từ lãnh đạo cơ quan, trong khi quyền phê chuẩn cuối cùng phải luôn thuộc về con người (de Almeida & dos Santos Júnior, 2025).

Thiết kế theo giá trị công và cơ chế human - in - the - loop: Bằng chứng thực nghiệm cho thấy, công dân và công chức tuyến đầu chia sẻ 3 ưu tiên thiết kế nhất quán: ưu tiên nhà phát triển trong nước, giới hạn quyền tự quyết của chatbot trong phạm vi kịch bản thường lệ, và bảo đảm can thiệp tức thời của con người khi hệ thống xử lý sai (Hemesath & Tepe, 2024).

Quản trị, minh bạch và tính chính danh: Tích hợp thuật toán vào quy trình hành chính đòi hỏi khung quản trị rủi ro với cấu trúc trách nhiệm pháp lý rõ ràng và mô hình AI có khả năng giải thích được lý do khuyến nghị trước công dân (Busuioc, 2021). Khi AI tham gia ra quyết định hành chính, nguy cơ "khoảng trống chính danh" gia tăng nếu thiếu cơ chế giám sát đa bên, đặc biệt với LLM khi rủi ro rò rỉ dữ liệu và thiên kiến đầu ra khiến kiểm duyệt bởi con người trở thành yêu cầu bắt buộc (König & Wenzelburger, 2021; Pesch, 2025).

Đối chiếu khung pháp lý - chính sách Việt Nam: Chương trình Chuyển đổi số quốc gia và Chiến lược AI đến năm 2030 tạo nền tảng thể chế để trợ lý AI vận hành trên nguồn dữ liệu ổn định (Thủ tướng Chính phủ Việt Nam, 2020, 2021a). Về tuân thủ, Nghị định 13/2023/NĐ - CP ràng buộc bảo vệ dữ liệu cá nhân trong giao tiếp AI, trong khi Luật An ninh mạng và Nghị định 53/2022/NĐ - CP tạo lớp bảo vệ hạ tầng số trước rủi ro lộ lọt từ bên ngoài (Chính phủ Việt Nam, 2022, 2023).

4. Thảo luận

4.1. Cơ hội triển khai

Trợ lý AI mở ra 3 cơ hội có tính chiến lược cho hành chính công Việt Nam: (i) cải thiện khả năng tiếp cận dịch vụ công liên tục, đa kênh và tích lũy dữ liệu nhu cầu công dân để cải tiến dịch vụ (Larsen & Følstad, 2024); (ii) nâng cao năng suất nội bộ theo hướng AI tư vấn - tóm tắt, tra cứu, soạn thảo - phù hợp với mức chấp nhận thực tế của đội ngũ công chức (Haesevoets et al., 2025); (iii) được thúc đẩy bởi lộ trình chuyển đổi số quốc gia và Chiến lược AI đến năm 2030, tạo nền tảng thể chế và dữ liệu để hệ thống vận hành ổn định (Thủ tướng Chính phủ Việt Nam, 2021a).

4.2. Thách thức và rủi ro

3 rủi ro trọng yếu cần được kiểm soát. Thứ nhất, niềm tin công dân mang tính phi đồng nhất theo lĩnh vực - đặc biệt thấp ở các dịch vụ nhạy cảm - đòi hỏi chiến lược truyền thông và lựa chọn phạm vi thí điểm thận trọng (Aoki, 2020). Thứ hai, khi AI thay thế một phần tương tác công chức, "khoảng trống chính danh" xuất hiện làm mờ nhạt trách nhiệm giải trình và quyền khiếu nại của công dân (König & Wenzelburger, 2021). Thứ ba, với LLM, rủi ro "memorisation" và rò rỉ dữ liệu qua prompt/output đòi hỏi kiểm soát kỹ thuật và quy trình nghiêm ngặt, bên cạnh nghĩa vụ tuân thủ pháp lý về bảo vệ dữ liệu cá nhân (Pesch, 2025).

4.3. Khoảng trống nghiên cứu

Khoảng trống lớn nhất là thiếu đánh giá nghiêm ngặt đối với các hệ thống chatbot đã vận hành thực tế, với bộ thước đo vượt ra ngoài hiệu suất kỹ thuật để bao phủ công bằng, minh bạch và tác động đến công chức; đồng thời, các câu hỏi về giới hạn pháp lý của LLM trong nghĩa vụ nêu lý do quyết định hành chính vẫn chưa được nghiên cứu đủ sâu (Cortés - Cediel et al., 2023).

Khuyến nghị thiết kế và chính sách triển khai trợ lý AI trong hành chính công Việt Nam

Khuyến nghị được cấu trúc theo 6 lớp tích hợp, phản ánh bằng chứng tổng hợp về quản trị AI, rủi ro thuật toán, niềm tin công dân và yêu cầu đánh giá hệ thống (de Almeida & dos Santos Júnior, 2025).

Lớp Quản trị và giải trình: Phân định 3 tuyến quản trị đồng bộ - chính sách, vận hành và giải trình - trong đó tuyến giải trình có nhiệm vụ chuẩn hóa năng lực giải thích quyết định AI một cách minh bạch trước công dân (Busuioc, 2021).

Lớp Quản trị dữ liệu: Kết nối trợ lý AI trực tiếp vào kho dữ liệu hành chính chính thống để thiết lập "nguồn sự thật" định danh, tránh nhiễu thông tin và bảo đảm cập nhật tri thức hệ thống liên tục (Chính phủ Việt Nam, 2020).

Lớp Bảo mật và quyền riêng tư: Tuân thủ nguyên lý tối giản lưu trữ dữ liệu cá nhân theo Nghị định 13/2023/NĐ - CP; với generative AI, bắt buộc kiểm duyệt định kỳ cơ chế huấn luyện để phòng ngừa rò rỉ dữ liệu nhận dạng (Chính phủ Việt Nam, 2023).

Lớp Thiết kế theo giá trị công: Xây dựng giao diện chatbot theo chuẩn mực văn hóa hành chính - lịch sự, tận tâm, chuyên nghiệp - nhằm đáp ứng kỳ vọng người dùng và tạo sự ủng hộ bền vững từ xã hội (Li & Wang, 2024).

Lớp Human - in - the - loop: Giới hạn tính tự chủ của hệ thống trong tác vụ quy trình hóa; mọi quyết định có tác động pháp lý phải có phê duyệt từ công chức và thiết lập cơ chế chuyển tuyến tức thời khi phát sinh tình huống ngoài kịch bản (Hemesath & Tepe, 2024).

Lớp Đánh giá đa mục tiêu: Mở rộng khung đo lường vượt ra ngoài chỉ số vận hành kỹ thuật, hướng đến định lượng đóng góp vào giá trị công; triển khai theo lộ trình thí điểm có kiểm soát trước khi mở rộng toàn quốc (Cortés - Cediell et al., 2023).

5. Kết luận

Tổng hợp bằng chứng từ 25 nghiên cứu học

thuật khẳng định trợ lý AI có tiềm năng nâng cao khả năng tiếp cận và hiệu quả cung cấp dịch vụ công, song các rủi ro về niềm tin, chính danh, trách nhiệm giải trình và bảo vệ dữ liệu cá nhân là có thực và cần được quản trị chủ động (Larsen & Følstad, 2024). Tại Việt Nam, Chiến lược AI đến năm 2030 và lộ trình chuyển đổi số quốc gia tạo động lực thể chế, trong khi Nghị định 13/2023/NĐ - CP cùng các quy định về an ninh mạng và an toàn thông tin xác lập ranh giới tuân thủ bắt buộc (Chính phủ Việt Nam, 2023).

Kết luận mang tính hành động là: Việt Nam nên triển khai trợ lý AI theo nguyên tắc "quản trị trước - tự động hóa sau", ưu tiên phạm vi thủ tục ít rủi ro, thiết lập cơ chế human - in - the - loop bắt buộc và đánh giá đa mục tiêu bao gồm thước đo giá trị công, nhằm bảo đảm hiệu quả vận hành đi đôi với tính chính danh và tuân thủ pháp luật (Hemesath & Tepe, 2024) ■

Lời cảm ơn

Nghiên cứu này được thực hiện trong khuôn khổ nhiệm vụ khoa học và công nghệ "Nghiên cứu, xây dựng ứng dụng trí tuệ nhân tạo (AI) mô hình ngôn ngữ lớn phục vụ công tác quản lý nhà nước", được phê duyệt theo Quyết định số 980/QĐ-SKH-CN ngày 06/11/2024 của Sở Khoa học và Công nghệ Thành phố Hồ Chí Minh. Nhóm tác giả xin trân trọng cảm ơn Đại học Kinh tế Thành phố Hồ Chí Minh (UEH) với tư cách đơn vị chủ trì, và Sở Khoa học và Công nghệ TP.HCM đã cấp kinh phí và tạo điều kiện thuận lợi trong suốt quá trình triển khai nghiên cứu.

TÀI LIỆU THAM KHẢO:

Tài liệu tiếng Việt:

Chính phủ Việt Nam. (2020). Nghị định số 47/2020/NĐ-CP về quản lý, kết nối và chia sẻ dữ liệu số của cơ quan nhà nước. Truy cập tại <https://vanban.chinhphu.vn/?docid=199754&pageid=27160>

Chính phủ Việt Nam. (2022). Nghị định số 53/2022/NĐ-CP quy định chi tiết một số điều của Luật An ninh mạng. Truy cập tại <https://vanban.chinhphu.vn/?docid=206381&pageid=27160>

Chính phủ Việt Nam. (2023). Nghị định số 13/2023/NĐ-CP về bảo vệ dữ liệu cá nhân. Truy cập tại <https://vanban.chinhphu.vn/?docid=207759&pageid=27160>

Thủ tướng Chính phủ Việt Nam. (2020). Quyết định số 749/QĐ-TTg: Phê duyệt Chương trình Chuyển đổi số quốc gia đến năm 2025, định hướng đến năm 2030. <https://chinhphu.vn/default.aspx?docid=200163&pageid=27160>

Thủ tướng Chính phủ Việt Nam. (2021a). Quyết định số 127/QĐ-TTg: Ban hành Chiến lược quốc gia về nghiên cứu, phát triển và ứng dụng Trí tuệ nhân tạo đến năm 2030. <https://vanban.chinhphu.vn/default.aspx?docid=202565&pageid=27160>

Thủ tướng Chính phủ Việt Nam. (2021b). Quyết định số 942/QĐ-TTg: Phê duyệt Chiến lược phát triển Chính phủ điện tử hướng tới Chính phủ số giai đoạn 2021-2025, định hướng đến năm 2030. Truy cập tại <https://vanban.chinhphu.vn/?docid=203403&pageid=27160>

Tài liệu tiếng nước ngoài:

Ahn, M. J., & Chen, Y.-C. (2022). Digital transformation toward AI-augmented public administration: The perception of government employees and the willingness to use AI in government. *Government Information Quarterly*, 39(2), 101664. Available at <https://doi.org/10.1016/j.giq.2021.101664>

Aoki, N. (2020). An experimental study of public trust in AI chatbots in the public sector. *Government Information Quarterly*, 37(4), 101490. Available at <https://doi.org/10.1016/j.giq.2020.101490>

Busuioc, M. (2021). Accountable artificial intelligence: Holding algorithms to account. *Public Administration Review*, 81(5), 825-836. Available at <https://doi.org/10.1111/puar.13293>

Cortés-Cediel, M. E., Segura-Tinoco, A., Cantador, I., & Rodríguez Bolívar, M. P. (2023). Trends and challenges of e-government chatbots: Advances in exploring open government data and citizen participation content. *Government Information Quarterly*, 40(4), 101877. Available at <https://doi.org/10.1016/j.giq.2023.101877>

de Almeida, P. G. R., & dos Santos Júnior, C. D. (2025). Artificial intelligence governance: Understanding how public organizations implement it. *Government Information Quarterly*, 42(1), 102003. Available at <https://doi.org/10.1016/j.giq.2024.102003>

Haesevoets, T., Verschuere, B., & Roets, A. (2025). AI adoption in public administration: Perspectives of public sector managers and public sector non-managerial employees. *Government Information Quarterly*, 42(2), 102029. Available at <https://doi.org/10.1016/j.giq.2025.102029>

Hemesath, S., & Tepe, M. (2024). Public value positions and design preferences toward AI-based chatbots in e-government: Evidence from a conjoint experiment with citizens and municipal front desk officers. *Government Information Quarterly*, 41(4), 101985. Available at <https://doi.org/10.1016/j.giq.2024.101985>

König, P. D., & Wenzelburger, G. (2021). The legitimacy gap of algorithmic decision-making in the public sector: Why it arises and how to address it. *Technology in Society*, 67, 101688. <https://doi.org/10.1016/j.techsoc.2021.101688>

Larsen, A. G., & Følstad, A. (2024). The impact of chatbots on public service provision: A qualitative interview study with citizens and public service providers. *Government Information Quarterly*, 41(2), 101927. Available at <https://doi.org/10.1016/j.giq.2024.101927>

Li, X., & Wang, J. (2024). Should government chatbots behave like civil servants? The effect of chatbot identity characteristics on citizen experience. *Government Information Quarterly*, 41(3), 101957. Available at <https://doi.org/10.1016/j.giq.2024.101957>

Page, M. J., McKenzie, J. E., Bossuyt, P. M., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *International Journal of Surgery*, 88, 105906. <https://doi.org/10.1016/j.ijssu.2021.105906>

Pesch, P. J. (2025). Potentials and challenges of large language models (LLMs) in the context of administrative decision-making. *European Journal of Risk Regulation*, 16(Special Issue 1), 76-95. Available at <https://doi.org/10.1017/err.2024.99>

Zuiderwijk, A., Chen, Y.-C., & Salem, F. (2021). Implications of the use of artificial intelligence in public governance: A systematic literature review and a research agenda. *Government Information Quarterly*, 38(3), 101577. Available at <https://doi.org/10.1016/j.giq.2021.101577>

Ngày nhận bài: 18/01/2026

Ngày phản biện đánh giá và sửa chữa: 6/02/2026

Ngày chấp nhận đăng bài: 27/02/2026

**AI ASSISTANTS FOR
VIETNAMESE PUBLIC ADMINISTRATION:
A SYSTEMATIC LITERATURE REVIEW
AND DEPLOYMENT RECOMMENDATIONS**

● **VUONG THUY NGOC¹**

● **NGUYEN VAN TIEN SON²**

● **NGO MINH VU³**

● **NGUYEN HUU HUAN³**

¹Promete Technology Joint Stock Company

²University of Economics Ho Chi Minh City - Vinh Long Campus (UEH Mekong)

³University of Economics Ho Chi Minh City

ABSTRACT:

Vietnam's digital transformation and AI strategy have created growing pressure to deploy AI assistants in public service delivery, accompanied by risks related to privacy and cybersecurity. This paper systematically reviews 25 international studies following the PRISMA 2020 guidelines to synthesize evidence on (i) factors influencing acceptance and trust, (ii) design principles, and (iii) governance models for deploying AI assistants in the public sector, thereby proposing recommendations for Vietnam. Key findings emphasize that purpose - oriented communication enhances public trust; design based on a "civil servant identity" and public value principles is critical; internal AI governance is a prerequisite for successful deployment; and a significant gap exists in evaluating operational systems using public value metrics. Deployment requires a controlled piloting roadmap, lawful personal data governance, human - in - the - loop mechanisms, and a multi- objective evaluation framework.

Keywords: AI assistant, government chatbot, digital government, personal data protection, AI governance, PRISMA.