NGÔN NGỮ HỌC VÀ VIỆT NGỮ HỌC

VAI TRÒ CỦA TRÍ TUỆ NHÂN TẠO (AI) TRONG VIỆC BIÊN SOẠN CẤU TRÚC VI MÔ CỦA TỪ ĐIỂN THUẬT NGỮ HÀNH CHÍNH TIẾNG VIỆT

VŨ THỊ SAO CHI*

TÓM TẮT: Dựa trên cơ sở lí luận của ngôn ngữ học từ điển và mô hình cấu trúc nghĩa, bài viết phân tích vai trò và tiềm năng của trí tuệ nhân tạo (Artificial Intelligence - AI) trong việc hỗ trợ biên soạn và hoàn thiện cấu trúc vi mô của *Từ điển thuật ngữ hành chính tiếng Việt*. Trên các bình diện như: thu thập và chuẩn hóa thuật ngữ, xây dựng và kiểm chứng định nghĩa, lựa chọn ví dụ pháp lí, thiết lập mạng lưới quan hệ thuật ngữ và cập nhật nguồn trích dẫn, AI được xem như một công cụ tăng cường trí tuệ, giúp nâng cao tính hệ thống, độ bao quát và hiệu quả của quá trình biên soạn. Bài viết đồng thời chỉ ra những thách thức về dữ liệu, phương pháp và diễn giải khái niệm pháp lí, từ đó đề xuất định hướng ứng dụng AI theo hướng kết hợp chặt chẽ giữa công nghệ hiện đại và kiểm soát học thuật của nhà biên soạn.

TỪ KHÓA: Cấu trúc vi mô; Từ điển thuật ngữ hành chính; Trí tuệ nhân tạo (AI); chuẩn hóa thuật ngữ.

NHẬN BÀI: 11/01/2026. BIÊN TẬP-CHỈNH SỬA-DUYỆT ĐĂNG: 05/03/2026

1. Mở đầu

Trong bối cảnh cải cách hành chính và chuyển đổi số ở Việt Nam, việc xây dựng *Từ điển thuật ngữ hành chính tiếng Việt* có ý nghĩa quan trọng đối với công tác chuẩn hóa và sử dụng thống nhất thuật ngữ trong văn bản pháp luật và quản lí nhà nước. Một trong những yếu tố quyết định chất lượng khoa học và giá trị ứng dụng của cuốn từ điển là *cấu trúc vi mô* - tức cấu trúc bên trong của mỗi mục từ, bao gồm hệ thống các thành phần thông tin được cung cấp và cách thức tổ chức, trình bày chúng một cách nhất quán. (Granger & Paquot, 2015; Kosem et al., 2017).

Theo định hướng biên soạn đã nêu, cấu trúc vi mô của Từ điển thuật ngữ hành chính tiếng Việt được xây dựng theo nguyên tắc *đẳng cấu*, thống nhất giữa các nhóm thuật ngữ, bao gồm: thuật ngữ dạng thức chính, phiên âm/ chuyển ngữ (nếu có), định nghĩa, ví dụ ngữ cảnh pháp lí, quan hệ thuật ngữ, ghi chú - phân biệt, và nguồn trích dẫn. Cấu trúc này nhằm bảo đảm người tra cứu không chỉ "biết nghĩa" mà còn "dùng đúng" thuật ngữ trong ngữ cảnh hành chính - pháp lí.

Hiện nay, trí tuệ nhân tạo (Artificial Intelligence - AI) đang phát triển mạnh mẽ, đặc biệt là các kĩ thuật xử lí ngôn ngữ tự nhiên và học máy. Do đó, một vấn đề đặt ra là: *AI có thể đóng vai trò gì trong việc hỗ trợ biên soạn và hoàn thiện cấu trúc vi mô của Từ điển thuật ngữ hành chính tiếng Việt?* (Russell & Norvig, 2021).

Bài viết này tiếp cận và giải quyết câu hỏi nêu trên từ góc độ ngôn ngữ học từ điển và ngôn ngữ học tính toán.

2. Khái quát về cấu trúc vi mô của từ điển thuật ngữ hành chính

Cấu trúc vi mô của Từ điển thuật ngữ hành chính được xác định là hệ thống các thành phần thông tin phản ánh đầy đủ các đặc điểm ngữ nghĩa, ngữ pháp, ngữ dụng và quan hệ từ vựng của thuật ngữ, nhằm phục vụ yêu cầu chuẩn hóa và sử dụng chính xác trong môi trường hành chính - pháp lí. (Gouws et al., 2013; Tarp, 2014). Các thành phần cơ bản trong cấu trúc vi mô của mục từ gồm:

1/Thuật ngữ (dạng thức chính); 2/Phiên âm/ chuyển ngữ (nếu có); 3/Định nghĩa; 4/Ví dụ ngữ cảnh thực tế; 5/Quan hệ thuật ngữ; 6/Ghi chú, phân biệt, lưu ý dùng; 7/Nguồn trích dẫn.

Cũng cần lưu ý rằng, trong các thành phần thông tin nêu trên, cốt lõi chính là phần định nghĩa. Để xây dựng phần định nghĩa thuật ngữ hành chính, chúng tôi dựa trên cơ sở lí thuyết cấu trúc nghĩa được Nguyễn Đức Tồn đưa ra trong công trình *Đặc trưng văn hóa - dân tộc của ngôn ngữ và tư duy*. Ông quan niệm mô hình cấu trúc ngữ nghĩa của các trường từ vựng - ngữ nghĩa nói chung, của một

**TS; Nhà xuất bản và Tạp chí KHXH - Viện Hàn lâm KHXH Việt Nam;*
Email: saochi1210@gmail.com

từ nói riêng, tương tự như mô hình cấu trúc nguyên tử của Niels Bohr, gồm ba vòng tròn đồng tâm tương ứng với ba thành tố nghĩa, trong đó tâm của cấu trúc nghĩa gồm hai vòng tròn tương ứng với thành tố I (siêu nghĩa vị chỉ phạm trù/ loại hay loài) - hạt nhân, và thành tố II (các nghĩa vị khu biệt của trường/ từ), còn thành tố III là thành tố ngoại vi (gồm các nghĩa vị tiềm tàng có thể trở thành nghĩa vị khu biệt với từ khác trong hoàn cảnh giao tiếp nhất định) (Nguyễn Đức Tồn, 2015, tr.428-430). Như vậy, để định nghĩa một từ nào đó, trong lời định nghĩa trước hết cần nêu thành tố I là siêu nghĩa vị chỉ loại của đối tượng (hiểu theo nghĩa rộng gồm sự vật, hành động, tính chất,…) được định nghĩa, tiếp đó nêu sự khác biệt về chủng - tức nêu các nghĩa vị thuộc thành tố II và III - là các đặc trưng khu biệt giữa đối tượng này với các đối tượng ngang hàng và dưới nó. Ví dụ, định nghĩa từ *bằng lăng: Cây to, tán dày, lá hình bầu dục, hoa mọc thành cụm đầu cành, màu tím hồng, gỗ nâu vàng, có thể dùng đóng thuyền và đồ đạc thông thường* (Từ điển tiếng Việt). Lời giải thích này gồm các nghĩa vị: Nghĩa vị chỉ loài: *cây*; nghĩa vị kích cỡ: *to*; nghĩa vị hình thức/ cấu tạo: *lá hình bầu dục, hoa mọc thành cụm đầu cành*; nghĩa vị màu sắc: *hoa màu tím hồng, gỗ nâu vàng*; nghĩa vị vai trò trong đời sống: *dùng đóng thuyền và đồ đạc thông thường*.

Đối với định nghĩa của thuật ngữ hành chính thì cần nêu những nghĩa vị khu biệt giữa thuật ngữ đang được định nghĩa với những thuật ngữ biểu hiện khái niệm có liên quan. Với mô hình ba thành tố: siêu nghĩa vị (thành tố I), các nghĩa vị khu biệt trung tâm (thành tố II) và các nghĩa vị ngoại vi (thành tố III), bảo đảm lời định nghĩa vừa xác lập đúng phạm trù khái niệm, vừa làm rõ những đặc trưng khu biệt của thuật ngữ trong hệ thống hành chính.

Chính hệ thống vi mô chặt chẽ này tạo tiền đề để xem xét vai trò của AI không phải như một công cụ rời rạc, mà như một phương tiện hỗ trợ toàn diện cho từng thành phần cấu trúc mục từ (Tarp, 2014).

3. Vai trò của AI trong biên soạn cấu trúc vi mô mục từ

AI có thể đóng vai trò quan trọng được trình bày sau đây trong việc hỗ trợ biên soạn và hoàn thiện cấu trúc vi mô (Russell & Norvig, 2021).

3.1. Xác lập và chuẩn hóa thuật ngữ dạng thức chính

AI có thể hỗ trợ thu thập và thống kê tự động các đơn vị thuật ngữ từ kho văn bản pháp luật, nghị quyết, thông tư, công báo và các tài liệu hành chính điện tử. Thông qua các mô hình nhận diện thuật ngữ (term extraction), hệ thống có thể: Phát hiện các tổ hợp từ có tần suất cao và tính ổn định trong văn bản hành chính; Gợi ý các biến thể hình thức (ví dụ: *xử lý/ xử lí*; *ngân sách nhà nước/ NSNN*); Hỗ trợ lựa chọn dạng thức chuẩn làm tiêu đề mục từ.

Nhờ đó, việc xác lập danh mục thuật ngữ không chỉ dựa vào rà soát thủ công, mà còn được hậu thuẫn bởi dữ liệu lớn, giúp tăng tính khách quan và bao quát (Granger & Paquot, 2015; Kosem et al., 2017).

3.2. Xử lí phiên âm và chuyển ngữ thuật ngữ vay mượn

Đối với các thuật ngữ có nguồn gốc nước ngoài như: *audit, statute, decree, charter, official gazette*, AI có thể hỗ trợ: Tự động liên kết giữa thuật ngữ tiếng Việt và dạng gốc tiếng Anh/ Pháp; Kiểm tra mức độ tương ứng giữa các cách chuyển ngữ trong thực tiễn sử dụng; Gợi ý dạng chuyển ngữ ổn định, phổ biến nhất trong văn bản chính thức. Điều này giúp bảo đảm tính nhất quán khi đưa thông tin phiên âm/ chuyển ngữ vào mục từ. Ví dụ:

Decree: trong tiếng Việt có các cách diễn đạt như *sắc lệnh*, *nghị định*, thậm chí *decree* giữ nguyên dạng. AI có thể phân tích ngữ cảnh và chỉ ra rằng trong hệ thống pháp luật Việt Nam hiện nay, *nghị định* là thuật ngữ chuẩn tương ứng với *decree*, trong khi *sắc lệnh* gắn với giai đoạn lịch sử nhất định. Nhờ đó, mục từ có thể chọn *nghị định* là dạng chính, đồng thời chú thích *decree* là dạng gốc tiếng Anh/ Pháp.

Charter: trong thực tiễn có các cách dùng như *hiến chương*, *điều lệ*, *charter*. AI có thể phân tích rằng *hiến chương* thường dùng trong ngữ cảnh quốc tế, tổ chức liên chính phủ, còn *điều lệ* dùng cho tổ chức, cơ quan, đơn vị cụ thể. Nhờ phân tích ngữ cảnh, hệ thống có thể gợi ý rằng hai thuật ngữ

tiếng Việt này tương ứng với các phạm vi nghĩa khác nhau của *charter*, giúp nhà biên soạn xử lí tách mục hoặc ghi chú phân biệt trong từ điển.

Official gazette: AI có thể nhận diện rằng trong các văn bản Việt Nam, cách chuyển ổn định là *công báo*, trong khi dạng dịch sát như *báo chính thức* hầu như không dùng. Trên cơ sở đó, hệ thống gợi ý chọn *công báo* làm dạng thức chính, liên kết với dạng gốc *official gazette*.

Qua các ví dụ trên, có thể thấy AI giúp nhà biên soạn từ điển không chỉ "biết có những cách chuyển nào", mà còn biết cách nào đang được cộng đồng chuyên môn và văn bản chính thức sử dụng ổn định nhất, từ đó đưa ra quyết định chuẩn hóa trong cấu trúc vi mô mục từ. Điều này góp phần bảo đảm tính nhất quán, khoa học và khả năng sử dụng thực tiễn của Từ điển thuật ngữ tiếng Việt trong môi trường giao tiếp chuyên môn hiện đại.

3.3. Hỗ trợ xây dựng và kiểm chứng định nghĩa thuật ngữ

Định nghĩa là "hạt nhân" của cấu trúc vi mô. AI có thể đóng vai trò hỗ trợ ở ba khía cạnh:

1) Khai thác ngữ cảnh: phân tích hàng nghìn câu chứa thuật ngữ trong văn bản pháp luật để làm nổi bật các đặc trưng nghĩa thường xuyên đi kèm.

2) Gợi ý cấu trúc định nghĩa: giúp nhận diện siêu nghĩa vị (ví dụ: *cá nhân*, *hoạt động quản lí nhà nước*, *tập hợp tài liệu*…) và các nghĩa vị khu biệt (thẩm quyền, phạm vi, đối tượng, tính cưỡng chế, nguồn ngân sách…).

3) Kiểm chứng tính nhất quán: so sánh các định nghĩa giữa những thuật ngữ liên quan (như: *cán bộ - công chức - viên chức*) để phát hiện chồng lấn hoặc mâu thuẫn.

Trên cơ sở đó, AI không thay thế việc soạn thảo định nghĩa của nhà biên soạn, nhưng giúp lời định nghĩa bám sát thực tiễn sử dụng và phù hợp hơn với mô hình cấu trúc nghĩa đã được xác lập (Gouws et al., 2013).

3.4. Tự động thu thập và đề xuất ví dụ ngữ cảnh pháp lí

Với yêu cầu mỗi mục từ có 1-2 ví dụ trích từ văn bản chính thức, AI có thể: Tìm kiếm tự động các câu chứa thuật ngữ trong luật, nghị định, thông tư; Xếp hạng các ví dụ theo mức độ điển hình, rõ nghĩa, chuẩn mực; Gắn kèm thông tin nguồn và điều khoản. Điều này giúp tiết kiệm đáng kể thời gian cho người biên soạn và bảo đảm ví dụ phản ánh đúng ngữ cảnh hành chính - pháp lí thực tế. Ví dụ:

Công chức: AI có thể quét hệ thống văn bản pháp luật và gợi ý câu: "Công chức là công dân Việt Nam, được tuyển dụng, bổ nhiệm vào ngạch, chức vụ, chức danh trong cơ quan của Đảng, Nhà nước, tổ chức chính trị - xã hội ở trung ương, cấp tỉnh, cấp huyện…" (*Luật Cán bộ, công chức*, Điều 4). Câu này được xếp hạng cao vì vừa chứa thuật ngữ ở vị trí trung tâm, vừa thể hiện rõ phạm vi nghĩa và đặc trưng pháp lí của khái niệm.

Khiếu nại: AI có thể đề xuất: "Công dân, cơ quan, tổ chức có quyền khiếu nại quyết định hành chính, hành vi hành chính khi có căn cứ cho rằng quyết định, hành vi đó là trái pháp luật." (*Luật Khiếu nại*, Điều 2). Ví dụ giúp làm rõ chủ thể, đối tượng và bản chất pháp lí của hành vi khiếu nại.

Trong mỗi trường hợp, AI không chỉ trích xuất câu chứa thuật ngữ, mà còn gắn kèm nguồn cụ thể (tên luật/ nghị định, điều khoản), đồng thời xếp hạng để nhà biên soạn từ điển lựa chọn những ví dụ: (1) ngắn gọn, (2) rõ nghĩa, (3) chuẩn mực về phong cách pháp lí. Nhờ đó, phần ví dụ minh họa trong mục từ vừa bảo đảm tính xác thực, vừa có giá trị hướng dẫn sử dụng cao đối với người tra cứu (Kosem et al., 2017).

3.5. Xây dựng mạng lưới quan hệ thuật ngữ

Dựa trên phân tích ngữ nghĩa và đồng xuất hiện, AI có thể gợi ý:

- Các thuật ngữ gần nghĩa hay có liên quan, ví dụ: *công chức - cán bộ - viên chức.*
- Các cặp dễ nhầm lẫn cần đối chiếu, ví dụ: *khiếu nại - tố cáo.*
- Các thuật ngữ trong cùng trường khái niệm, ví dụ: *thu ngân sách - chi ngân sách*.

Nhờ đó, mục "quan hệ thuật ngữ" trong cấu trúc vi mô có thể được xây dựng trên cơ sở mạng lưới dữ liệu lớn, hỗ trợ người dùng hình dung hệ thống thuật ngữ hành chính như một chỉnh thể.

3.6. Vai trò của AI trong việc phát hiện và chú giải các trường hợp thuật ngữ dễ nhầm lẫn

Trong thực tiễn sử dụng thuật ngữ hành chính - pháp lí, một trong những vấn đề nổi bật là hiện tượng dùng lẫn lộn hoặc dùng chưa chuẩn giữa các thuật ngữ có hình thức hoặc nghĩa gần nhau. Những nhầm lẫn này không chỉ làm giảm tính chính xác của văn bản hành chính, mà trong nhiều trường hợp còn có thể dẫn đến cách hiểu sai về thẩm quyền, trình tự, hoặc bản chất pháp lí của một hoạt động quản lí nhà nước. Do đó, trong cấu trúc vi mô của Từ điển thuật ngữ hành chính tiếng Việt, mục ghi chú, phân biệt hoặc lưu ý dùng giữ vai trò đặc biệt quan trọng, giúp người tra cứu nhận diện ranh giới nghĩa và phạm vi sử dụng của từng thuật ngữ.

Trong bối cảnh này, AI có thể trở thành công cụ hữu hiệu để phát hiện một cách hệ thống các trường hợp dễ nhầm lẫn và hỗ trợ biên soạn các chú giải mang tính hướng dẫn (Bender et al., 2021). Ví dụ:

- *thủ tục - quy trình*: Khi quét các văn bản hành chính, AI có thể phát hiện nhiều trường hợp dùng "*quy trình giải quyết hồ sơ*" trong những ngữ cảnh mà văn bản pháp luật chuẩn mực quy định là "*thủ tục giải quyết hồ sơ*". Từ đó, hệ thống gợi ý chú giải: "*Thủ tục* (hành chính) là trình tự, cách thức do cơ quan nhà nước có thẩm quyền quy định để giải quyết công việc đối với tổ chức, cá nhân; không dùng đồng nghĩa với *quy trình* là chuỗi các bước nội bộ trong tổ chức để thực hiện một hoạt động".

- *khiếu nại - tố cáo*: Phân tích các văn bản cho thấy hai thuật ngữ này đôi khi bị dùng lẫn trong các báo cáo, dù pháp luật quy định rõ sự khác nhau về đối tượng và mục đích. AI có thể gợi ý chú giải: "*Khiếu nại* là yêu cầu xem xét lại quyết định hành chính, hành vi hành chính liên quan trực tiếp đến quyền, lợi ích của người khiếu nại; *tố cáo* là việc báo cho cơ quan có thẩm quyền về hành vi vi phạm pháp luật của tổ chức, cá nhân".

Thông qua những ví dụ như trên, có thể thấy AI không chỉ giúp phát hiện các cặp thuật ngữ dễ nhầm lẫn dựa trên dữ liệu lớn, mà còn cung cấp cơ sở thực tiễn để nhà biên soạn từ điển xây dựng các ghi chú phân biệt, cảnh báo cách dùng trong mục từ. Nhờ đó, Từ điển thuật ngữ hành chính tiếng Việt không chỉ dừng lại ở việc giải thích nghĩa, mà còn thực sự phát huy chức năng hướng dẫn sử dụng và chuẩn hóa thuật ngữ trong hoạt động hành chính - pháp lí.

3.6.1. Phân tích lỗi dùng từ từ dữ liệu văn bản lớn

Thông qua việc khai thác các kho văn bản hành chính, pháp luật, báo chí chính thống và cả các văn bản nghiệp vụ nội bộ, AI có thể tự động:

- Thu thập các câu chứa những cặp thuật ngữ nghi dễ nhầm lẫn;
- Phân tích ngữ cảnh sử dụng để phát hiện những trường hợp một thuật ngữ xuất hiện trong môi trường ngữ nghĩa "đáng lẽ" thuộc về thuật ngữ khác;
- Thống kê tần suất và kiểu loại của các lỗi hoặc cách dùng chưa chuẩn.

Ví dụ: Khi quét hàng nghìn văn bản, AI có thể nhận ra rằng *quy trình* thường bị dùng trong những ngữ cảnh pháp lí mà về bản chất phải là *thủ tục hành chính* (ví dụ: "quy trình giải quyết hồ sơ" thay vì "thủ tục giải quyết hồ sơ"), hoặc *giấy tờ* bị dùng thay cho *hồ sơ* trong các văn bản cần nhấn mạnh tính hệ thống của tập hợp tài liệu. Việc phát hiện dựa trên dữ liệu lớn giúp chỉ ra rằng đây không phải là hiện tượng cá biệt, mà là xu hướng phổ biến cần được chú giải trong từ điển.

3.6.2. Nhận diện các cặp thuật ngữ có vùng nghĩa chồng lấn

Các mô hình ngữ nghĩa phân bố cho phép AI đo mức độ tương đồng về ngữ cảnh sử dụng giữa các thuật ngữ. Những cặp có độ tương đồng cao nhưng vẫn khác nhau về bản chất khái niệm chính là những ứng viên điển hình của nhóm "dễ nhầm lẫn".

Chẳng hạn, các cặp như: *thủ tục - quy trình*; *chức năng - nhiệm vụ - quyền hạn*; *cán bộ - công chức - viên chức*; *khiếu nại - tố cáo*; *văn bản - giấy tờ - tài liệu - hồ sơ*; *chế độ - chính sách - quy định*;… có thể được AI phát hiện thông qua sự gần gũi trong không gian ngữ nghĩa. Từ đó, hệ thống gợi ý cho nhà biên soạn những cụm thuật ngữ cần đưa vào mục "ghi chú, phân biệt" của cấu trúc vi mô.

3.6.3. So sánh ngữ cảnh "chuẩn" và "lệch chuẩn"

Một đóng góp quan trọng của AI là khả năng học từ các văn bản pháp luật chuẩn mực (luật, nghị định, thông tư) để xây dựng mô hình về ngữ cảnh sử dụng chuẩn của từng thuật ngữ. Khi đối chiếu với các văn bản khác (ví dụ: báo cáo hành chính, công văn, tài liệu hướng dẫn), AI có thể phát hiện những trường hợp thuật ngữ được dùng trong ngữ cảnh lệch khỏi chuẩn này. Ví dụ: Nếu trong các văn bản luật, *thẩm quyền* chủ yếu gắn với các động từ như: *quy định, quyết định, giao, phân định*, nhưng trong một số văn bản nghiệp vụ lại xuất hiện trong ngữ cảnh mang nghĩa năng lực chung (*có thẩm quyền làm việc này*), AI có thể gợi ý rằng cách dùng sau dễ gây mơ hồ giữa *thẩm quyền* (authority) và *năng lực/ quyền* nói chung, từ đó cần có chú giải phân biệt trong từ điển.

3.6.4. Gợi ý nội dung chú giải và cảnh báo dùng

Trên cơ sở tổng hợp các trường hợp nhầm lẫn, AI có thể hỗ trợ gợi ý khung nội dung cho các ghi chú, chẳng hạn: Chỉ ra điểm khác biệt cốt lõi về khái niệm (bản chất pháp lí, phạm vi áp dụng, chủ thể liên quan); Nêu những cách dùng phổ biến nhưng chưa chuẩn; Đề xuất cách dùng đúng theo văn bản pháp luật hiện hành.Ví dụ: Với cặp *hồ sơ - giấy tờ*, AI có thể tổng hợp rằng *hồ sơ* thường xuất hiện cùng các từ *lập, quản lí, lưu trữ, giải quyết*, trong khi *giấy tờ* đi với *xuất trình, nộp, bổ sung*, từ đó gợi ý chú giải: *"Hồ sơ" là tập hợp có hệ thống nhiều tài liệu liên quan đến một vụ việc hoặc thủ tục; không nên dùng đồng nghĩa với "giấy tờ" là từng văn bản đơn lẻ.*

Những gợi ý này giúp nhà biên soạn nhanh chóng xây dựng các chú giải có cơ sở dữ liệu rõ ràng, thay vì chỉ dựa vào trực giác.

3.6.5. Ưu tiên hóa các trường hợp cần chú giải trong từ điển

Do số lượng thuật ngữ lớn, không phải mọi cặp gần nghĩa đều cần chú giải chi tiết. AI có thể hỗ trợ xếp hạng mức độ "rủi ro nhầm lẫn" dựa trên các tiêu chí như: Tần suất xuất hiện của cách dùng lệch chuẩn; Mức độ chồng lấn ngữ nghĩa; Hệ quả pháp lí tiềm tàng nếu hiểu sai thuật ngữ.

Những cặp có điểm rủi ro cao sẽ được ưu tiên đưa vào mục "ghi chú, phân biệt" của cấu trúc vi mô, giúp từ điển tập trung vào những vấn đề có ý nghĩa thực tiễn lớn nhất đối với người dùng. Ví dụ:

- *khiếu nại - tố cáo*: AI có thể thống kê rằng hai thuật ngữ này thường xuyên bị dùng lẫn trong đơn thư và báo cáo hành chính. Mức độ chồng lấn về ngữ cảnh cao, trong khi hệ quả pháp lí của việc nhầm lẫn là rất lớn (xác định sai thẩm quyền giải quyết, trình tự, thời hạn). Do đó, cặp này được hệ thống xếp ở mức rủi ro rất cao và ưu tiên chú giải chi tiết trong từ điển.

- *thủ tục - quy trình*: Dữ liệu cho thấy *quy trình* bị dùng thay cho *thủ tục hành chính* với tần suất cao trong công văn, biểu mẫu, dù pháp luật quy định rõ về *thủ tục*. Tuy hệ quả pháp lí không nghiêm trọng như *khiếu nại - tố cáo*, nhưng ảnh hưởng trực tiếp đến chuẩn hóa ngôn ngữ hành chính. AI xếp cặp này ở mức rủi ro cao, cần có ghi chú phân biệt rõ.

- *văn bản - tài liệu*: Mặc dù hai thuật ngữ có vùng nghĩa giao nhau và đôi khi bị dùng lẫn, AI nhận thấy hệ quả pháp lí của nhầm lẫn thường không lớn, vì ít làm thay đổi thẩm quyền hay trình tự xử lí. Do đó, cặp này được xếp rủi ro thấp, có thể chỉ cần chú giải ngắn gọn hoặc không ưu tiên trong giai đoạn đầu.

- *cán bộ - công chức - viên chức*: AI thống kê rằng ba thuật ngữ này thường bị dùng thay thế cho nhau trong báo chí và một số văn bản nội bộ, trong khi pháp luật quy định rất chặt chẽ phạm vi của từng khái niệm. Hệ quả pháp lí của việc nhầm lẫn có thể ảnh hưởng đến tuyển dụng, quản lí, chế độ chính sách. Vì vậy, nhóm này được xếp rủi ro rất cao và cần chú giải, đối chiếu chi tiết trong từ điển.

Từ các ví dụ trên, có thể thấy AI giúp nhà biên soạn từ điển không chỉ phát hiện các cặp dễ nhầm lẫn, mà còn xếp hạng mức độ ưu tiên dựa trên dữ liệu thực tiễn và hệ quả sử dụng. Nhờ đó, mục "ghi chú, phân biệt" trong Từ điển thuật ngữ hành chính tiếng Việt có thể tập trung vào những trường hợp có tác động lớn nhất đến tính chuẩn xác và hiệu quả của giao tiếp hành chính - pháp lí, thay vì dàn trải nguồn lực cho mọi cặp gần nghĩa.

3.6.6. Gắn chú giải với lí thuyết cấu trúc nghĩa

Việc chú giải các trường hợp dễ nhầm lẫn không chỉ mang tính thực dụng, mà còn gắn chặt với cơ sở lí luận về cấu trúc nghĩa mà Nguyễn Đức Tồn đã đưa ra. AI, khi hỗ trợ phân tích ngữ cảnh, có thể

làm nổi bật những nghĩa vị khu biệt (thành tố II) giúp phân biệt hai thuật ngữ gần nhau hoặc những nghĩa vị ngoại vi (thành tố III) dễ bị kích hoạt trong giao tiếp, dẫn đến lẫn lộn. Ví dụ:

- *nhiệm vụ - chức năng*: AI phân tích ngữ cảnh cho thấy cả hai cùng chia sẻ thành tố I (siêu nghĩa vị) là: "hoạt động/ việc làm gắn với một cơ quan, tổ chức". Tuy nhiên, *nhiệm vụ* thường đi kèm các yếu tố như "được giao, cụ thể, trong thời hạn, phải hoàn thành" (thành tố II: "tính cụ thể, tính thời điểm"), trong khi *chức năng* lại thường gắn với "vai trò, công dụng, ổn định, lâu dài" (thành tố II: "tính ổn định, tính vai trò"). Đồng thời, trong giao tiếp, thành tố III của *nhiệm vụ* dễ kích hoạt sắc thái "áp lực, trách nhiệm trước mắt", khiến người viết đôi khi dùng thay cho *chức năng*. AI làm nổi bật sự khác biệt này để nhà biên soạn từ điển xây dựng chú giải: *"Nhiệm vụ* là công việc cụ thể được giao trong một giai đoạn nhất định; *chức năng* là vai trò, công dụng ổn định của cơ quan".

- *thẩm quyền - quyền hạn*: AI phát hiện cả hai cùng có thành tố I là "quyền được thực hiện hành vi quản lí". Tuy nhiên, *thẩm quyền* thường gắn với thành tố II "phạm vi pháp lí gắn với vị trí, chức danh, cấp quản lí", còn *quyền hạn* gắn với "các quyền cụ thể được trao để thực hiện nhiệm vụ". Trong nhiều ngữ cảnh, thành tố III của *quyền hạn* (sắc thái "cụ thể, liệt kê được") dễ khiến người dùng lẫn với *thẩm quyền* khi nói về phạm vi chung. AI giúp tách bạch các nghĩa vị này để hình thành chú giải phân biệt.

- *chức vụ - chức danh*: AI phân tích ngữ cảnh cho thấy cả hai cùng có thành tố I là "vị trí của một người trong tổ chức", nhưng *chức vụ* nhấn mạnh thành tố II "vị trí gắn với quyền lực, thẩm quyền quản lí", còn *chức danh* nhấn mạnh "tên gọi chính thức của vị trí theo hệ thống ngạch, bậc". Trong giao tiếp, thành tố III của *chức danh* (sắc thái "tên gọi") dễ khiến người viết dùng thay cho *chức vụ*. AI giúp nhà biên soạn từ điển xác lập ranh giới này trong chú giải.

Qua các ví dụ trên, có thể thấy AI không chỉ phát hiện hiện tượng nhầm lẫn trên bề mặt sử dụng, mà còn hỗ trợ nhà biên soạn từ điển đi sâu vào cấu trúc nghĩa của thuật ngữ, làm nổi bật các nghĩa vị khu biệt và ngoại vi theo khung lí luận của GS.TS. Nguyễn Đức Tồn. Nhờ đó, các chú giải phân biệt trong từ điển thuật ngữ hành chính tiếng Việt không chỉ có giá trị thực hành, mà còn có chiều sâu khoa học, góp phần kết nối công tác biên soạn từ điển với nền tảng lí thuyết ngôn ngữ học hiện đại.

3.6.7. Ý nghĩa đối với tính hướng dẫn sử dụng của từ điển

Nhờ sự hỗ trợ của AI, mục "ghi chú, phân biệt" trong cấu trúc vi mô không chỉ phản ánh kinh nghiệm cá nhân của người biên soạn, mà được xây dựng trên cơ sở phân tích hệ thống các lỗi và lệch chuẩn trong thực tiễn sử dụng. Điều này giúp: Tăng tính thực chứng của các chú giải; Nâng cao khả năng hướng dẫn người dùng "dùng đúng" thuật ngữ; Góp phần trực tiếp vào mục tiêu chuẩn hóa ngôn ngữ hành chính. Theo đó, AI đóng vai trò như một "bộ dò lệch chuẩn" ở quy mô lớn, giúp từ điển thuật ngữ hành chính tiếng Việt thực sự trở thành công cụ không chỉ để tra cứu nghĩa, mà còn để định hướng và chuẩn hóa cách dùng trong hoạt động quản lí nhà nước. Ví dụ:

- *thủ tục - quy trình*: Từ việc AI phát hiện tần suất cao các văn bản dùng *quy trình hành chính* thay cho *thủ tục hành chính*, mục từ *thủ tục (hành chính)* trong từ điển có thể kèm chú giải: "Không dùng *quy trình* để thay cho *thủ tục* trong văn bản hành chính - pháp lí; "quy trình" chỉ chuỗi bước nội bộ, còn "thủ tục" là trình tự do cơ quan có thẩm quyền quy định để giải quyết công việc cho tổ chức, cá nhân". Chú giải này giúp người soạn thảo văn bản nhận biết và tránh một lỗi rất phổ biến.

- *thẩm quyền - quyền hạn*: AI phát hiện sự lẫn lộn trong các báo cáo nội bộ. Mục từ có thể bổ sung chú giải: "*Thẩm quyền* dùng để chỉ phạm vi pháp lí gắn với vị trí, chức danh; *quyền hạn* dùng để chỉ các quyền cụ thể được trao để thực hiện nhiệm vụ. Không dùng hai thuật ngữ này thay thế cho nhau". Chú giải này giúp người đọc tránh hiểu sai về phạm vi pháp lí của chủ thể quản lí.

Thông qua những ví dụ trên, có thể thấy các chú giải do AI hỗ trợ không chỉ có tính mô tả, mà mang tính chỉ dẫn sử dụng rõ rệt: chỉ ra dùng thuật ngữ nào trong ngữ cảnh nào, tránh lỗi gì và vì sao không nên dùng lẫn. Nhờ vậy, từ điển thuật ngữ hành chính tiếng Việt thực sự trở thành một công cụ định hướng chuẩn mực, giúp người dùng không chỉ "biết nghĩa", mà còn "biết dùng đúng" trong hoạt động quản lí nhà nước, góp phần trực tiếp vào mục tiêu chuẩn hóa ngôn ngữ hành chính trong bối cảnh chuyển đổi số (Tarp, 2014; Gouws et al., 2013).

3.7. Vai trò của AI trong quản lí và liên kết nguồn trích dẫn pháp lí

Trong cấu trúc vi mô của Từ điển thuật ngữ hành chính tiếng Việt, mục nguồn trích dẫn có ý nghĩa đặc biệt quan trọng, bởi mỗi định nghĩa và ví dụ minh họa đều cần được đặt trên cơ sở các văn bản pháp luật chính thức như Hiến pháp, luật, nghị định, thông tư, công báo. Nguồn trích dẫn không chỉ bảo đảm độ tin cậy khoa học của lời giải thích, mà còn gắn từ điển với hệ thống pháp luật hiện hành, qua đó nâng cao giá trị ứng dụng của công trình trong thực tiễn quản lí nhà nước. Tuy nhiên, hệ thống văn bản pháp luật Việt Nam thường xuyên được sửa đổi, bổ sung, thay thế, khiến việc quản lí nguồn trích dẫn bằng phương thức thủ công trở nên nặng nề và khó bảo đảm tính cập nhật. Trong bối cảnh đó, AI có thể đóng vai trò then chốt trong việc tự động hóa và chuẩn hóa công việc này.

3.7.1. Tự động nhận diện và trích xuất nguồn pháp luật

Thông qua các mô hình nhận dạng thực thể (Named Entity Recognition - NER) và phân tích cấu trúc văn bản, AI có thể:

- Tự động phát hiện tên văn bản pháp luật (ví dụ: Luật Cán bộ, công chức 2025; Nghị định 30/2020/NĐ-CP; Thông tư số 01/2020/TT-BNV);
- Nhận diện điều, khoản, điểm cụ thể trong văn bản;
- Tách và chuẩn hóa thông tin nguồn theo một định dạng thống nhất.

Nhờ đó, khi biên soạn ví dụ hoặc định nghĩa, hệ thống có thể tự động gợi ý và gắn kèm nguồn trích dẫn đầy đủ, giảm thiểu sai sót về tên gọi, năm ban hành hoặc kí hiệu văn bản. Ví dụ:

- Khi quét một đoạn văn bản: "Theo quy định tại Điều 4 Luật Cán bộ, công chức năm 2025, công chức là công dân Việt Nam được tuyển dụng…", AI có thể tự động nhận diện các thực thể: *Luật Cán bộ, công chức 2025* (tên văn bản), *Điều 4* (điều khoản), và chuẩn hóa thành dạng nguồn: (*Luật Cán bộ, công chức, 2025, Điều 4*) để gắn vào ví dụ minh họa trong mục từ *công chức.*
- Trong trường hợp văn bản viết: "… theo hướng dẫn tại Khoản 2 Điều 5 Thông tư số 01/2020/TT-BNV", AI có thể nhận diện đầy đủ cấu trúc: *Khoản 2, Điều 5, Thông tư 01/2020/TT-BNV*, và chuẩn hóa thành: (*Thông tư 01/2020/TT-BNV, Bộ Nội vụ, 2020, Khoản 2, Điều 5*).
- Khi gặp các cách ghi không thống nhất như: "*NĐ 30/2020*", "*Nghị định 30/2020*", hay "*Nghị định số 30/2020/NĐ-CP*", AI có thể đối chiếu và quy về một dạng chuẩn duy nhất để tránh trùng lặp và sai khác trong toàn bộ từ điển.

Thông qua các ví dụ trên, có thể thấy AI giúp nhà biên soạn từ điển tự động hóa khâu nhận diện và chuẩn hóa nguồn pháp luật, bảo đảm rằng mỗi ví dụ và định nghĩa trong Từ điển thuật ngữ hành chính tiếng Việt đều được gắn với nguồn trích dẫn đầy đủ, thống nhất và dễ kiểm chứng. Điều này không chỉ tiết kiệm thời gian, mà còn nâng cao độ tin cậy pháp lí và tính chuyên nghiệp của công trình từ điển.

3.7.2. Liên kết hai chiều giữa mục từ và văn bản pháp luật

AI có thể giúp xây dựng mối liên kết hai chiều giữa từ điển và kho văn bản pháp luật số hóa:

Từ mục từ → văn bản: người dùng có thể truy cập trực tiếp đến điều khoản nơi thuật ngữ được sử dụng.

Từ văn bản → mục từ: khi đọc một điều luật, hệ thống có thể tự động liên kết các thuật ngữ trong đó tới mục từ tương ứng trong từ điển.

Cơ chế này biến từ điển không chỉ là công cụ tra cứu độc lập, mà thành một hệ sinh thái liên thông với cơ sở dữ liệu pháp luật, phục vụ hiệu quả cho cán bộ, công chức và người làm công tác pháp chế. Ví dụ:

- Từ mục từ → văn bản pháp luật: Khi người dùng tra cứu mục từ *công chức* trong từ điển điện tử, ngoài phần định nghĩa và ví dụ, hệ thống có thể hiển thị các liên kết như: *Luật Cán bộ, công chức (2025), Điều 4; Điều 8; Điều 10…* Người dùng chỉ cần nhấp vào *Điều 4* là có thể mở trực tiếp điều khoản trong văn bản luật, nơi thuật ngữ *công chức* được định nghĩa và sử dụng. Nhờ đó, người tra cứu có thể kiểm chứng ngay cơ sở pháp lí của định nghĩa trong từ điển.
- Từ mục từ → nhiều nguồn liên quan: Với mục từ *thủ tục hành chính*, hệ thống có thể tự động liệt kê các văn bản như: *Nghị định 30/2020/NĐ-CP; Nghị định 61/2018/NĐ-CP; các thông tư hướng*

dẫn có liên quan, kèm liên kết đến các điều khoản cụ thể. Điều này giúp người dùng tiếp cận nhanh toàn bộ hệ thống quy định pháp lí gắn với thuật ngữ.

- Từ văn bản pháp luật → mục từ: Khi người dùng đọc *Điều 2 Luật Khiếu nại*, các thuật ngữ như: *khiếu nại, người khiếu nại, quyết định hành chính, hành vi hành chính,…* có thể được hệ thống tự động gạch chân hoặc hiển thị dạng liên kết. Khi nhấp vào *khiếu nại*, người dùng sẽ được chuyển sang mục từ tương ứng trong từ điển, nơi có định nghĩa khái quát, chú giải phân biệt với *tố cáo*, và các ví dụ minh họa.

- Hỗ trợ đọc hiểu văn bản phức tạp: Trong một điều luật có nhiều thuật ngữ chuyên môn, chẳng hạn: *thẩm quyền, quyền hạn, trách nhiệm, trình tự, thủ tục*…, AI có thể tự động nhận diện và gắn liên kết đến các mục từ tương ứng. Điều này giúp cán bộ, công chức hoặc người làm pháp chế vừa đọc luật, vừa tra cứu nhanh nghĩa thuật ngữ mà không cần rời khỏi giao diện văn bản.

- Liên thông theo ngữ cảnh sử dụng: Khi người dùng đang ở mục từ *hồ sơ,* hệ thống không chỉ liên kết đến các điều luật có từ *hồ sơ*, mà còn gợi ý các mục từ liên quan như *giấy tờ, thành phần hồ sơ, tiếp nhận hồ sơ,* tạo thành mạng liên thông giữa từ điển và văn bản pháp luật.

Qua các ví dụ trên, có thể thấy cơ chế liên kết hai chiều do AI hỗ trợ đã biến Từ điển thuật ngữ hành chính tiếng Việt từ một công cụ tra cứu tĩnh thành một thành phần của hệ sinh thái pháp lí số, nơi người dùng có thể di chuyển linh hoạt giữa khái niệm và quy định pháp luật. Điều này đặc biệt hữu ích đối với cán bộ, công chức và người làm công tác pháp chế trong việc tra cứu, đối chiếu và áp dụng pháp luật trong thực tiễn.

3.7.3. Tự động cập nhật khi văn bản được sửa đổi, thay thế

Một thách thức lớn là nhiều văn bản pháp luật bị sửa đổi, bổ sung hoặc hết hiệu lực theo thời gian. AI, khi được kết nối với các cơ sở dữ liệu pháp luật chính thức, có thể: Phát hiện khi một văn bản nguồn đã bị thay thế hoặc hết hiệu lực; Tự động gợi ý cập nhật nguồn mới tương ứng; Cảnh báo cho nhóm biên soạn những mục từ chịu ảnh hưởng bởi sự thay đổi đó. Ví dụ: Nếu một nghị định được sửa đổi bằng nghị định mới, AI có thể chỉ ra rằng các mục từ đang dẫn nguồn nghị định cũ cần được rà soát lại về định nghĩa và ví dụ. Điều này giúp từ điển luôn duy trì được tính thời sự pháp lí, điều đặc biệt quan trọng đối với một từ điển thuật ngữ hành chính.

3.7.4. Kiểm chứng tính phù hợp giữa định nghĩa và nguồn trích dẫn

AI không chỉ gắn nguồn, mà còn có thể hỗ trợ kiểm tra mức độ phù hợp giữa nội dung định nghĩa và văn bản được trích dẫn. Bằng cách so sánh các đoạn định nghĩa với các điều khoản pháp luật chứa thuật ngữ, hệ thống có thể: Phát hiện những điểm không tương thích hoặc thiếu sót; Gợi ý bổ sung các yếu tố nghĩa quan trọng còn chưa được phản ánh trong định nghĩa; Nhận diện các trường hợp định nghĩa quá rộng hoặc quá hẹp so với quy định pháp lí. Ví dụ: Nếu định nghĩa *công chức* trong từ điển chưa phản ánh đầy đủ các tiêu chí về "vị trí việc làm" hay "chế độ công vụ" được nêu trong luật hiện hành, AI có thể phát hiện sự lệch này và đề xuất rà soát.

3.7.5. Chuẩn hóa cách trình bày nguồn trích dẫn trong toàn bộ từ điển

Trong một công trình lớn, việc trình bày nguồn trích dẫn rất dễ bị không thống nhất (cách ghi năm, tên văn bản, kí hiệu, viết tắt). AI có thể đảm nhiệm việc: Chuẩn hóa tất cả nguồn theo một mẫu thống nhất; Phát hiện các biến thể không nhất quán; Tự động sửa về dạng chuẩn. Điều này góp phần nâng cao tính chuyên nghiệp và khoa học của từ điển. Ví dụ:

- Trong các mục từ khác nhau, cùng một văn bản có thể được ghi là: *Luật Cán bộ, công chức năm 2025*; *Luật CBCC 2025*; *Luật Cán bộ, công chức 2025)*; hoặc *Luật Cán bộ, công chức sửa đổi 2025*. AI có thể nhận diện rằng các dạng này đều quy chiếu đến cùng một văn bản, rồi tự động chuẩn hóa về một dạng thống nhất, chẳng hạn: *Luật Cán bộ, công chức, 2025*.

- Với nghị định, có thể gặp các cách ghi: *Nghị định số 30/2020/NĐ-CP*; *NĐ 30/2020/NĐ-CP*; *Nghị định 30/2020*; *Nghị định 30/2020/NĐ-CP của CP*. AI sẽ đối chiếu và chuẩn hóa thành: *Nghị định số 30/2020/NĐ-CP, Chính phủ, 2020*.

Khi phát hiện cùng một nguồn được ghi khác nhau ở nhiều mục từ, AI không chỉ sửa về dạng chuẩn, mà còn cảnh báo để người biên soạn rà soát, tránh tình trạng một văn bản xuất hiện dưới nhiều hình thức khác nhau trong toàn bộ từ điển.

Thông qua các ví dụ trên, có thể thấy AI giúp bảo đảm rằng mọi nguồn trích dẫn trong Từ điển thuật ngữ hành chính tiếng Việt đều được trình bày nhất quán, rõ ràng và dễ kiểm chứng. Nhờ đó, công trình từ điển đạt được mức độ chuyên nghiệp cao hơn, đồng thời thuận tiện cho việc tra cứu, liên kết và cập nhật trong môi trường số.

3.7.6. Hỗ trợ tra cứu lịch sử hiệu lực của thuật ngữ

Nhiều thuật ngữ hành chính có thể thay đổi phạm vi nghĩa theo sự thay đổi của pháp luật. AI có thể giúp lưu vết: Thuật ngữ này được sử dụng trong những văn bản nào qua từng giai đoạn; Sự thay đổi trong cách quy định có dẫn đến thay đổi nội dung khái niệm hay không. Nhờ đó, từ điển không chỉ phản ánh trạng thái hiện hành, mà còn có thể cung cấp thông tin lịch sử về thuật ngữ, rất hữu ích cho nghiên cứu và cho việc giải thích các văn bản cũ. Ví dụ:

- *công chức*: AI có thể tổng hợp và hiển thị chuỗi nguồn như: *Luật Cán bộ, công chức 2008 → Luật sửa đổi, bổ sung 2019 → Luật năm 2025*, kèm theo các định nghĩa tương ứng ở từng thời điểm. Hệ thống cho thấy phạm vi *công chức* đã được điều chỉnh (ví dụ: về đối tượng, cấp hành chính, chế độ tuyển dụng), từ đó giúp người tra cứu hiểu rằng khi đọc văn bản ban hành năm 2010, nội dung khái niệm *công chức* có thể không hoàn toàn trùng với cách hiểu hiện nay.

- *viên chức*: AI có thể chỉ ra rằng trước khi Luật Viên chức được ban hành, thuật ngữ này chưa được xác lập rõ ràng trong hệ thống pháp luật; sau đó nội dung khái niệm được định nghĩa cụ thể và có những lần sửa đổi liên quan đến chế độ hợp đồng, đơn vị sự nghiệp công lập. Nhờ vậy, từ điển có thể cung cấp ghi chú lịch sử: "Khái niệm '*viên chức*' theo luật hiện hành khác với cách dùng trong một số văn bản trước năm …".

- *thủ tục hành chính*: AI có thể lưu vết cách thuật ngữ này được sử dụng trong các nghị định về cải cách thủ tục hành chính qua các giai đoạn khác nhau, cho thấy phạm vi và cách hiểu về *thủ tục hành chính* ngày càng được chuẩn hóa và chi tiết hơn. Điều này giúp giải thích vì sao cùng một thuật ngữ trong văn bản cũ và mới có thể có mức độ khái quát khác nhau.

- *xử phạt vi phạm hành chính*: AI có thể tổng hợp các luật và nghị định qua từng thời kì, chỉ ra rằng nội dung khái niệm và thẩm quyền xử phạt đã được mở rộng hoặc thu hẹp theo từng lần sửa đổi. Từ đó, từ điển có thể cung cấp thông tin lịch sử để người đọc hiểu đúng thuật ngữ trong các văn bản đã hết hiệu lực.

- *đơn vị hành chính*: AI có thể theo dõi sự thay đổi trong cách phân loại và gọi tên (tỉnh, huyện, xã; thị xã, thành phố thuộc tỉnh…) qua các luật và nghị quyết khác nhau, cho thấy nội dung khái niệm *đơn vị hành chính* đã có những điều chỉnh gắn với quá trình cải cách tổ chức bộ máy nhà nước.

Qua các ví dụ trên, có thể thấy AI giúp Từ điển thuật ngữ hành chính tiếng Việt không chỉ phản ánh trạng thái hiện hành của thuật ngữ, mà còn cung cấp chiều kích lịch sử của khái niệm. Nhờ đó, người sử dụng có thể hiểu và giải thích đúng thuật ngữ trong bối cảnh thời gian cụ thể, đặc biệt hữu ích cho nghiên cứu pháp lí, đối chiếu văn bản cũ - mới và cho việc áp dụng pháp luật trong những trường hợp liên quan đến hiệu lực hồi tố hoặc chuyển tiếp.

3.7.7. Ý nghĩa đối với độ tin cậy pháp lí và giá trị ứng dụng của từ điển

Với sự hỗ trợ của AI, mục "nguồn trích dẫn" trong cấu trúc vi mô không còn là phần phụ mang tính hình thức, mà trở thành một hệ thống liên kết động giữa từ điển và pháp luật. Điều này giúp: Bảo đảm mỗi định nghĩa và ví dụ đều có cơ sở pháp lí rõ ràng; Duy trì tính cập nhật trong môi trường pháp luật luôn biến động; Tăng cường độ tin cậy và giá trị sử dụng thực tiễn của từ điển đối với cán bộ, công chức, nhà nghiên cứu và người dân.Theo đó, AI đóng vai trò như "bộ quản gia pháp lí số", giúp từ điển thuật ngữ hành chính tiếng Việt vừa giữ được tính ổn định khoa học, vừa thích ứng linh hoạt với sự phát triển của hệ thống pháp luật trong thời đại số.

3.8. Thách thức và giải pháp khắc phục

Việc ứng dụng AI trong biên soạn cấu trúc vi mô cũng đặt ra những thách thức như: dữ liệu pháp luật chưa được chuẩn hóa hoàn toàn; nguy cơ đồng nhất máy móc các đơn vị hình thức (Bender et al., 2021); hạn chế trong diễn giải khái niệm pháp lí và văn hóa hành chính; và yêu cầu hợp tác liên ngành giữa ngôn ngữ học, luật học và khoa học máy tính.

Giải pháp để khắc phục các tình trạng trên là: 1) Xây dựng các kho ngữ liệu pháp lí chuẩn hóa; 2) Phát triển mô hình AI huấn luyện theo nguyên tắc ngôn ngữ học từ điển và cấu trúc nghĩa; 3) Đặt AI trong quy trình biên soạn có kiểm soát học thuật chặt chẽ.

4. Kết luận

Trong bối cảnh chuyển đổi số và xây dựng Chính phủ điện tử, việc biên soạn Từ điển thuật ngữ hành chính Việt Nam với cấu trúc vi mô khoa học, nhất quán là nhiệm vụ có ý nghĩa lí luận và thực tiễn sâu sắc. AI, nếu được vận dụng đúng hướng, có thể trở thành công cụ hỗ trợ đắc lực trong việc chuẩn hóa thuật ngữ, xây dựng định nghĩa, lựa chọn ví dụ, thiết lập quan hệ thuật ngữ và cập nhật nguồn trích dẫn.

Sự kết hợp giữa cơ sở lí luận ngôn ngữ học (đặc biệt là lí thuyết cấu trúc nghĩa) và các công cụ AI hiện đại hứa hẹn mở ra một hướng tiếp cận mới cho việc biên soạn Từ điển thuật ngữ hành chính ở Việt Nam, vừa bảo đảm tính khoa học, chuẩn mực, vừa nâng cao hiệu quả và khả năng cập nhật trong thời đại số.

*Bài viết là sản phẩm của Nhiệm vụ KH&CN cấp Bộ: "*Nghiên cứu biên soạn Từ điển thuật ngữ hành chính Việt Nam*". Mã số KHXH/NV/2025-70. Viện Ngôn ngữ học chủ trì.

TÀI LIỆU THAM KHẢO

Tiếng Việt

1. Đỗ Hữu Châu (1981), *Từ vựng - ngữ nghĩa tiếng Việt*. Hà Nội: Nxb Giáo dục.
2. Hoàng Phê (Chủ biên) (2010), *Từ điển tiếng Việt*. Hà Nội: Nxb Đà Nẵng.
3. Nguyễn Đức Tồn (chủ biên) (2016), *Thuật ngữ học tiếng Việt hiện đại*. Hà Nội: Nxb Khoa học xã hội.
4. Nguyễn Đức Tồn (2006), *Từ đồng nghĩa tiếng Việt*. Hà Nội: Nxb Khoa học Xã hội.
5. Nguyễn Đức Tồn (2015, tái bản có bổ sung), *Đặc trưng văn hóa - dân tộc của ngôn ngữ và tư duy*. H.: Nxb KHXH.
6. Nguyễn Đức Tồn (2026), *Từ vựng học*. Hà Nội: Nxb Thế giới.

Tiếng Anh

7. Atkins, B.T.S., & Rundell, M. (2008), *The Oxford guide to practical lexicography*. Oxford: Oxford University Press.
8. Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021), On the dangers of stochastic parrots: Can language models be too big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610-623. https://doi.org/10.1145/3442188.3445922.
9. Bergenholtz, H., & Tarp, S. (2010), *Lexicography at a crossroads: Dictionaries and encyclopedias today, lexicographical tools tomorrow*. Bern: Peter Lang.
10. Bowker, L. (2003), *Terminology tools for translators*. Amsterdam: John Benjamins.
11. Cabré, M.T. (1999), *Terminology: Theory, methods and applications*. Amsterdam: John Benjamins.
12. Eisenstein, J. (2019), *Introduction to natural language processing*. MIT Press.
13. Goldberg, Y. (2017), *Neural network methods for natural language processing*. Morgan & Claypool. https://doi.org/10.2200/S00762ED1V01Y201703HLT037.
14. Gouws, R. H., Heid, U., Schweickard, W., & Wiegand, H. E. (Eds.). (2013), *Dictionaries: An international encyclopedia of lexicography* (Vols. 1-4). Berlin: De Gruyter.

15. Granger, S., & Paquot, M. (2015), *Electronic lexicography*. Oxford: Oxford University Press.
16. Jurafsky, D., & Martin, J.H. (2023), *Speech and language processing* (3rd ed.). Draft. https://web.stanford.edu/~jurafsky/slp3/.
17. Kosem, I., et al. (Eds.). (2017), *Electronic lexicography in the 21st century: Lexicography from scratch*. Brno: Lexical Computing.
18. Manning, C.D., Raghavan, P., & Schütze, H. (2008), *Introduction to information retrieval*. Cambridge: Cambridge University Press.
19. Russell, S., & Norvig, P. (2021), *Artificial intelligence: A modern approach* (4th ed.). Pearson.
20. Sager, J. C. (1990), *A practical course in terminology processing*. Amsterdam: John Benjamins.
21. Tarp, S. (2014), *Lexicography in the borderland between knowledge and non-knowledge*. Lexikos, 24, 238-261.

The role of Artificial Intelligence (AI) in compiling the microstructure of the Vietnamese Administrative Terminology Dictionary

Abstract: Drawing on lexicographic theory and the semantic structure model proposed, this paper examines the role and potential of Artificial Intelligence (AI) in supporting the compilation and refinement of the microstructure of entries in a Vietnamese administrative terminology dictionary. Focusing on tasks such as term extraction and standardization, definition construction and validation, selection of authentic legal examples, establishment of terminological networks, and updating of source references, the paper argues that AI should be regarded as an intelligence-augmenting tool rather than a replacement for human expertise. The study also discusses major challenges related to data quality, methodological constraints, and the interpretation of legal concepts, and proposes orientations for the responsible integration of AI into the lexicographic workflow under rigorous scholarly supervision.

Key words: Microstructure; administrative terminology dictionary; Artificial Intelligence (AI); terminological standardization.