

PHÁT TRIỂN TRÍ TUỆ NHÂN TẠO LẤY CON NGƯỜI LÀM TRUNG TÂM *(Human-centered Artificial Intelligence)*

Khoa Công nghệ Thông tin

Trường Đại học Khoa học Tự nhiên

Đại học Quốc gia thành phố Hồ Chí Minh

1 Mở đầu

Chúng ta đang chứng kiến một kỷ nguyên với những thay đổi nhanh chóng chưa từng có nhờ vào những thành tựu vượt bậc trong khoa học và công nghệ. Các lĩnh vực nghiên cứu về trí tuệ nhân tạo (artificial intelligence), học máy (machine learning), dữ liệu lớn (big data) và phân tích dữ liệu (data analytics), tương tác người máy (human-computer interaction), thực tại ảo (virtual reality/augmented reality/mixed reality), và các hệ thống thông minh (intelligent systems) đã có những bước tiến nhanh chóng. Những thành tựu trong các lĩnh vực này đã và đang được áp dụng thực tế như: các giải pháp và hệ thống thông minh trong y tế (như điều trị ung thư), giáo dục (học thích ứng – adaptive learning), giao thông (xe tự hành), và môi trường (dự báo thời tiết). Những thành tựu này có triển vọng thúc đẩy mạnh mẽ sự phát triển kinh tế xã hội của các nước trên thế giới. Và ngày nay, các lĩnh vực này được xem nền tảng của cuộc Cách mạng Công nghiệp 4.0 mà chúng ta đang hướng tới.

Tuy nhiên, việc phát huy những thành tựu trong các lĩnh vực trên gặp nhiều hạn chế và thách thức ở Việt Nam. Chúng ta chưa thực sự áp dụng tốt những kết quả từ các lĩnh vực trên để giải quyết những vấn đề trong thực tế phục vụ đời sống người dân. Chúng ta cũng chưa có cơ chế và thiết bị nhằm thu thập, lưu trữ, và chia sẻ dữ liệu lớn cho nghiên cứu và xây dựng các hệ thống thông minh. Một thách thức quan trọng nữa là hạn chế trong đào tạo nguồn nhân lực phục vụ nghiên cứu, phát triển, và ứng dụng các giải pháp trí tuệ nhân tạo, học máy, thực tại ảo, v.v... vào thực tế.

Với xu thế hiện nay, những bước tiến trong các lĩnh vực nêu trên không dừng lại ở phòng thí nghiệm mà cần được áp dụng trong thực tế, phục vụ con người với tinh thần lấy con người làm trung tâm. Các bài toán nghiên cứu được hình thành từ thực tiễn và các giải pháp đáp ứng tốt nhất những nhu cầu thiết thực của con người. Các nghiên cứu trong các

lĩnh vực của ngành công nghệ thông tin đang hướng tới các hệ thống thông minh phục vụ con người.

Trong báo cáo này, chúng tôi trình bày một số vấn đề về xu hướng chung trong phát triển Trí tuệ Nhân tạo trên thế giới cũng như một số kết quả đạt được tại Trường Đại học Khoa học Tự nhiên, cụ thể là tại Khoa Công nghệ Thông tin, phòng Thí nghiệm Trí tuệ Nhân tạo (AILab) và Phòng thí nghiệm Công nghệ Phần mềm (SELab) trong những năm qua, đồng thời nêu lên những định hướng phát triển cho việc nghiên cứu và phát triển Trí tuệ Nhân tạo hướng đến tầm khu vực và thế giới cũng như tiềm năng ứng dụng phục vụ xã hội.

2 Một số kết quả đạt được về nghiên cứu và ứng dụng Trí tuệ Nhân tạo trong Trường Đại học Khoa học Tự nhiên, ĐHQG-HCM

Tại trường Đại học Khoa học Tự nhiên, Đại học Quốc gia thành phố Hồ Chí Minh, hướng nghiên cứu về Trí tuệ Nhân tạo đã được hình thành và phát triển ngay từ khi thành lập Khoa Công nghệ Thông tin từ năm 1996. Nhiều nhóm nghiên cứu được xây dựng và phát triển tại Khoa Công nghệ Thông tin trong hơn 20 năm qua, bao gồm: Tính toán mềm, Tính toán tiến hóa, Máy học, Xử lý ảnh và Thị giác Máy tính, Xử lý ngôn ngữ tự nhiên, Xử lý ngôn ngữ nói, Khai thác Dữ liệu và Khoa học Dữ liệu, Dữ liệu lớn...

Năm 2007, phòng thí nghiệm Trí tuệ Nhân tạo (AILab) do Đại học Quốc gia thành phố Hồ Chí Minh đầu tư được thành lập tại Trường Đại học Khoa học Tự nhiên. Đây là phòng thí nghiệm Trí tuệ Nhân tạo đầu tiên tại thành phố Hồ Chí Minh. Hướng nghiên cứu chính của AILab bao gồm xử lý ngôn ngữ nói (nhận dạng và tổng hợp tiếng nói), xử lý ảnh và thị giác máy tính.

Trung tâm ngôn ngữ học tính toán (Computational Language Center) được thành lập từ năm 2015 với các chủ đề nghiên cứu và ứng dụng xử lý ngôn ngữ tự nhiên.

Từ năm 2017, Phòng thí nghiệm Công nghệ Phần mềm (SELab, thành lập từ năm 2002) đã bắt đầu xây dựng và phát triển hướng nghiên cứu và phát triển các hệ thống tương tác thông minh, ứng dụng các thuật toán xử lý thông minh để phát triển các dịch vụ cho thành phố thông minh, hỗ trợ cuộc sống thường nhật của các cá nhân và tổ chức.

Trường Đại học khoa học Tự nhiên đã có nhiều công trình công bố quốc tế về nghiên cứu Trí tuệ Nhân tạo. Riêng trong 5 năm qua, đã có 96 công trình trên các tạp chí ISI (29 công trình thuộc Q1 và 31 thuộc Q2), trên 200 bài báo hội nghị khoa học quốc tế.

Ngoài ra, các nhóm nghiên cứu gồm giảng viên và sinh viên của Khoa Công nghệ Thông tin, Phòng thí nghiệm Trí tuệ Nhân tạo và Phòng thí nghiệm Công nghệ Phần mềm đã đạt được nhiều thứ hạng cao trong các cuộc thi (challenge) quốc tế về Trí tuệ Nhân tạo: 11 giải Nhất, 3 giải Nhì, 4 giải Ba, 2 hạng Sáu, 1 hạng 11, 1 hạng 24. Các challenge này được tổ chức trong các Hội nghị quốc tế có uy tín như CVPR (2017-2019), Eurographics (2016-2019), MediaEval (2018-2019), ICMR (2018-2019)...

Ví dụ như: Giải Nhất liên tiếp 3 năm về Truy vấn đối tượng 3 chiều (SHREC, EuroGraphics), Hạng Ba về phân đoạn đối tượng trong video (DAVIS, CVPR 2017, 2019), hạng Sáu về phát hiện và ước lượng vận tốc xe trong xử lý thông tin giao thông (AI City Challenge, CVPR 2018), hạng Tám thế giới về phát hiện sự kiện bất thường trong giao thông (AI City Challenge, CVPR 2019), hạng Nhất về phát hiện và chẩn đoán ảnh nội soi (MediaEval 2018), Hạng 11 về đếm nhân tế bào trong điều trị ung thư (MoNuSec, MICCAI 2018)...

Trong các cuộc thi này có các Trường/Viện/Phòng thí nghiệm nổi tiếng khác: Google, Tencent AILab, Google Deepmind, Viện Max Planck Institute (Đức), ByteDance AILab, Đại học Carnegie Mellon University (CMU, Mỹ), Đại học University of Illinois at Urbana-Champaign (UIUC-Mỹ)...

3 Định hướng đào tạo, nghiên cứu và ứng dụng Trí tuệ Nhân tạo lấy con người làm trung tâm

3.1 Sự liên kết đa ngành, chuyên ngành, hướng nghiên cứu để phục vụ định hướng phát triển Trí tuệ Nhân tạo lấy con người làm trung tâm

Trong nhiều năm qua, tại Khoa Công nghệ Thông tin, trường Đại học Khoa học Tự nhiên đã định hướng nghiên cứu tập trung vào các chủ đề sau:

- **Trí tuệ nhân tạo và khoa học dữ liệu:** đề xuất và đánh giá giải thuật trí tuệ nhân tạo và học máy, khai thác dữ liệu và phân tích dữ liệu lớn. Hai chủ đề cụ thể đã và đang được thực hiện là về các thuật toán nâng cao trong khai thác mẫu tuần tự và luật (đề

tài NAFOSTED) và phát triển các thuật toán khai thác mẫu tuần tự và luật từ cơ sở dữ liệu chuỗi (đề tài NAFOSTED).

- **Môi trường thông minh và đa phương tiện:** nhóm nghiên cứu tập trung vào sự kết hợp và giao thoa của trí tuệ nhân tạo, công nghệ phần mềm, tương tác người máy nhằm phát triển môi trường tương tác thông minh và an toàn hỗ trợ các công việc và cuộc sống thường nhật của con người.
- **Hệ thống tích hợp và thông minh:** nghiên cứu và xây dựng các hệ thống thông minh hỗ trợ con người trong việc vận hành, ra quyết định thông qua việc thu thập, xử lý, phân tích dữ liệu từ nhiều nguồn khác nhau. Các giải pháp được xây dựng dựa trên việc áp dụng các kết quả nghiên cứu trong lĩnh vực trí tuệ nhân tạo, học máy, thị giác máy tính, tối ưu hoá, v.v...
- **Xử lý ngôn ngữ tự nhiên:** chuyển ngữ, xác định phong cách văn bản tiếng Việt, thống kê văn bản tiếng Việt, so sánh độ tương đồng của văn bản tiếng Việt. Xây dựng và khai thác từ điển, ngữ liệu tiếng Việt, ngữ liệu song ngữ/song song đa ngữ, và các công cụ xử lý văn bản.
- **Xử lý tiếng nói:** tập trung các vấn đề liên quan đến xử lý tiếng Việt, như mô hình hoá âm thanh và ngôn ngữ, xây dựng kho dữ liệu âm thanh và văn bản, truy vấn âm thanh và nhạc, kiểm chứng và nhận biết người nói, và dịch tiếng nói. Một số kết quả như Hệ thống tự động phát sinh phụ đề cho các bản tin thời sự tiếng Việt trên kênh Youtube và Hệ thống hỗ trợ tạo tự động biên bản cuộc họp từ tiếng nói.
- **Tìm kiếm thông tin và khai thác văn bản:** các chủ đề bao gồm các mô hình tìm kiếm thông tin, tìm kiếm thông tin xuyên ngữ, khai thác văn bản (như văn bản luật), và phân tích dữ liệu lớn.
- **Sinh trắc học:** tập trung tìm hiểu việc áp dụng các thuật toán học máy vào các ứng dụng sinh trắc học, từ đó đề xuất những cải tiến thuật toán mới nhằm nâng cao hiệu suất các hệ thống sinh trắc học trong nhiều lĩnh vực thực tế khác nhau.
- **Khoa học dữ liệu thị giác:** các chủ đề theo hướng nghiên cứu này bao gồm khai thác dữ liệu ảnh dựa vào tái tạo thực thể ba chiều; xây dựng dịch vụ thông minh hỗ trợ

giám sát, thống kê trong hoạt động của con người dựa thông tin thị giác; và nhận dạng hành động người và dự đoán mặt người, cảm xúc theo thời gian.

- **Mật mã phi tập trung và blockchain:** mục tiêu chính là thiết kế các công cụ mật mã sáng tạo và phát triển các hệ mật mã đảm bảo tính riêng tư có tính ứng dụng thực tiễn và an ninh hậu lượng tử. Ngoài ra, nhóm cũng hướng đến phát triển các hệ nền dựa trên công nghệ blockchain.
- **Công nghệ phần mềm:** tập trung nghiên cứu các phương pháp và công cụ cho quá trình phát triển phần mềm như ước tính chi phí (đề tài NAFOSTED), kiểm thử tự động, và phát triển phần mềm trên nền tảng đám mây. Nhóm quan tâm đến việc áp dụng trí tuệ nhân tạo, học máy và xử lý ngôn ngữ tự nhiên giúp đưa ra quyết định đúng đắn và kịp thời trong phát triển phần mềm.

Các định hướng trên là cơ sở và nền tảng cho việc tích hợp các nghiên cứu liên chủ đề, liên ngành nhằm phát triển các hệ thống thông minh (với tính năng xử lý và tương tác thông minh), giúp áp dụng trí tuệ nhân tạo đến gần với cuộc sống, gắn bó với con người và lấy con người làm nhân tố trung tâm (Human-centered AI) theo xu hướng mới trên thế giới. Việc nghiên cứu và phát triển các hệ thống thông minh đòi hỏi cần có sự kết hợp chặt chẽ giữa các lĩnh vực về Trí tuệ Nhân tạo, Kỹ thuật Phần mềm, Tương tác Người – Máy và Bảo mật, An toàn thông tin. Sự kết hợp giữa các lĩnh vực trên phù hợp với xu hướng rất mới trên thế giới là nghiên cứu Trí tuệ Nhân tạo hướng đến con người làm nhân tố trung tâm (Human-centered Artificial Intelligence), tiêu biểu là tại Phòng thí nghiệm Human-centered AI của Đại học MIT (<https://hcai.mit.edu/>) hay Viện nghiên cứu Human-Centered Artificial Intelligence của Đại học Stanford (<https://hai.stanford.edu/>).

3.2 Phát triển đào tạo, hợp tác và chuyển giao công nghệ

Nhóm chiến lược đầu tiên cần tập trung là phát triển đào tạo, hợp tác và chuyển giao công nghệ, góp phần đào tạo nguồn nhân lực có kỹ năng cao về các lĩnh vực phân tích dữ liệu, trí tuệ nhân tạo, học máy, tương tác người máy, v.v... Khoa Công nghệ Thông tin, trường Đại học Khoa học Tự nhiên cũng định hướng việc đào tạo nguồn nhân lực gồm

đội ngũ chuyên gia, nhà khoa học có trình độ quốc tế, và đội ngũ kỹ sư chất lượng cao về lĩnh vực Trí tuệ Nhân tạo:

- Hiện nay, môn Trí tuệ Nhân tạo là môn bắt buộc chung cho tất cả SV của mọi ngành/chuyên ngành (mỗi khóa tuyển khoảng 900-1000 SV), trong đó khoảng 30-35% sinh viên học tiếp và nghiên cứu chuyên sâu về Trí tuệ Nhân tạo.
- Đầu tư cho việc phát triển các nhóm nghiên cứu mạnh, có uy tín trong lĩnh vực Trí tuệ Nhân tạo.
- Từng bước triển khai các nội dung đào tạo về Trí tuệ Nhân tạo cho các học sinh xuất sắc, đặc biệt là tại các Trường THPT chuyên tại thành phố
- Xây dựng và triển khai các khóa học, chương trình đào tạo ngắn hạn để bồi dưỡng, nâng cao năng lực về nghiên cứu và ứng dụng Trí tuệ Nhân tạo.
- Triển khai các chương trình theo định hướng “đại học chia sẻ” giúp lan tỏa và nâng cao năng lực nghiên cứu và phát triển ứng dụng Trí tuệ Nhân tạo tại thành phố.

3.3 Một số hướng nghiên cứu tiêu biểu đang được phát triển

Dưới đây là một số hướng nghiên cứu chiến lược đang được tập trung phát triển tại Khoa Công nghệ Thông tin, Phòng thí nghiệm Trí tuệ Nhân tạo và Phòng thí nghiệm Công nghệ Phần mềm nhằm đẩy mạnh việc nghiên cứu và ứng dụng Trí tuệ Nhân tạo phục vụ cuộc sống, lấy con người làm trung tâm:

Tương tác người máy thông minh. Hướng nghiên cứu này tập trung vào sự kết hợp và giao thoa của trí tuệ nhân tạo và công nghệ phần mềm, tương tác người máy nhằm phát triển môi trường tương tác thông minh và an toàn hỗ trợ các công việc và cuộc sống thường nhật của con người. Hai hướng nghiên cứu cụ thể bao gồm (1) áp dụng các giải pháp và công nghệ hiện đại vào thiết kế các hệ thống tương tác người-máy phục vụ các yêu cầu khác nhau ở nhiều ngành và lĩnh vực (2) phát minh các cách tiếp cận, kỹ thuật giúp con người tương tác dễ dàng và hiệu quả hơn với các thiết bị máy tính.

Các giải pháp giao thông thông minh. Chủ đề này tập trung nghiên cứu các giải pháp thu thập và rút trích thông tin giao thông từ các camera, sensor, các thiết bị IoT, từ báo

chí, mạng xã hội, v.v... nhằm xây dựng hệ thống giao thông thông minh có khả năng dự đoán lưu lượng xe, phát hiện sự kiện, tình huống bất thường trên đường phố qua camera thông minh, phát hiện và gán nhãn đối tượng, phân tích luồng giao thông và dự đoán số người trong đám đông v.v... Hệ thống này cũng có những chức năng thông minh khác như điều phối, phân luồng giao thông cho phù hợp, điều khiển đèn và tín hiệu giao thông, v.v... hay sử dụng các thuật toán chuyên dụng nhằm xây dựng các ứng dụng giao thông thông minh phục vụ cho các doanh nghiệp như điều phối xe giao hàng, xe vận chuyển rác thải v.v... cho hiệu quả và phù hợp với các yêu cầu phức tạp trên thực tế.

Lưu trữ bảo mật dữ liệu trong môi trường IoT. Chủ đề này hướng đến nghiên cứu các hệ nền cho phép xây dựng các môi trường thông minh (thành phố thông minh, nhà thông minh, công nghiệp 4.0, ...) ở đó các thực thể tương tác thời gian thực với nhau qua một mạng các cảm biến, liên mạng vạn vật kết nối - IoT. Việc quản lý môi trường như thế dựa trên sự trao đổi thông tin/dữ liệu liên tục giữa các thiết bị thông minh được thu thập từ môi trường. Không khó để hình dung một môi trường như thế cũng là mục tiêu tấn công của các tin tặc. Rủi ro bị tấn công là rất cao, và một khi việc tấn công xảy ra sẽ dẫn đến những thiệt hại nghiêm trọng đến tính riêng tư của công dân, đến an ninh trong môi trường vật lý. Công nghệ blockchain có thể xem như mảnh ghép lấp kín các lỗ hổng về độ tin cậy, tính an toàn, bảo mật trong IoT. Các đặc trưng tin cậy, tự động, và phi tập trung khiến công nghệ blockchain với công cụ hợp đồng thông minh (smart contract) thích hợp cho việc triển khai các môi trường thông minh, ở đó, các thiết bị có thể tự vận hành mà không cần sự hiện diện của người hay bộ phận kiểm soát tập trung (CA). Hơn nữa, blockchain giúp các thiết bị thông minh giao tiếp và trao đổi thông tin/dữ liệu an toàn, với các công cụ sẽ được nghiên cứu và tích hợp vào từng giải pháp cụ thể: mật mã hiện đại, kỹ thuật/giải pháp ẩn danh, giả danh, và các công cụ an ninh đã có. Hai nghiên cứu cụ thể sẽ được triển khai tại Khoa CNTT trong những năm tới (i) chứng thực phi tập trung và ứng dụng trong IoT (ii) định danh phi tập trung và ứng dụng trong FinTech.

Ngoài ra, hướng nghiên cứu này cũng tập trung phân tích và tăng cường tính cạnh tranh an toàn, bảo mật (security and privacy) và phát triển các hệ nền dựa trên công nghệ blockchain (blockchain-based platform). Các công việc chính bao gồm (1) thiết kế các

công cụ mã hóa hiện đại và phát triển các hệ mật mã bảo toàn tính riêng tư đảm bảo an toàn và bảo mật ngay cả khi tấn công bằng phương pháp/máy tính lượng tử và (2) phát triển các hệ nền dựa trên công nghệ blockchain.

Giải pháp thông minh về giáo dục. Hướng đến mục tiêu nghiên cứu và xây dựng hệ thống thông minh hỗ trợ học tập (trong đó có lập trình) và quản lý chương trình đào tạo trên nền Web và di động. Hệ thống này sẽ là nền tảng kết nối và tương tác giữa người học và người dạy. Hệ thống có những chức năng thông minh đề nghị chương trình học, môn học và bài tập phù hợp với mỗi học sinh, sinh viên (personalized learning), đề nghị giáo viên phù hợp với từng học sinh, sinh viên.

Theo dõi sông hồ, dự báo mực nước, ngập lụt. Chủ đề nghiên cứu này hướng đến thu thập dữ liệu lượng nước ở sông hồ qua vệ tinh, camera, báo chí và mạng xã hội, và các thiết bị IoT. Các giải pháp được đề xuất sẽ giải quyết các bài toán như xác định lưu lượng nước sông hồ; theo dõi lượng nước và diện tích bề mặt sông hồ; phát hiện những nơi có thể xảy ra lũ lụt và hạn hán; dự đoán mực nước hiện tại, tương lai; đo lường, theo dõi và cảnh báo.

4 Kết luận

Trong báo cáo này, chúng tôi trình bày và phân tích tình hình chung của nghiên cứu trên thế giới về Trí tuệ Nhân tạo, trong đó tập trung theo xu hướng mới là kết hợp liên ngành, chuyên ngành, chủ đề nghiên cứu nhằm phát triển và ứng dụng Trí tuệ Nhân tạo phục vụ cuộc sống, lấy con người làm trung tâm (Human-centered AI). Chúng tôi đã phân tích những kết quả đã đạt được tại Trường Đại học Khoa học Tự nhiên, ĐHQG-HCM trong việc nghiên cứu và phát triển ứng dụng Trí tuệ Nhân tạo tại Khoa Công nghệ Thông tin, Phòng thí nghiệm Trí tuệ Nhân tạo, Phòng thí nghiệm Công nghệ Phần mềm, từ đó nêu lên định hướng chiến lược trong việc phát triển đào tạo, nghiên cứu, ứng dụng Trí tuệ Nhân tạo để phục vụ xã hội.