BỘ THÔNG TIN VÀ TRUYỀN THÔNG

**VIỆN CÔNG NGHIỆP PHẦN MỀM VÀ NỘI DUNG SỐ VIỆT NAM**

--------------------------------

**THUYẾT MINH**

**Dự thảo Tiêu chuẩn quốc gia**

**Công nghệ thông tin – Trí tuệ nhân tạo – Các khái niệm và thuật ngữ trí tuệ nhân tạo**

**Hà Nội – 10/2022**

**MỤC LỤC**

[1 Tên gọi và ký hiệu của dự thảo TCVN 3](#_Toc114757174)

[1.1 Tên tiếng Việt của dự thảo TCVN: 3](#_Toc114757175)

[1.2 Tên tiếng Anh của dự thảo TCVN: 3](#_Toc114757176)

[2 Đặt vấn đề 3](#_Toc114757177)

[2.1 Sự cần thiết xây dựng tiêu chuẩn thuật ngữ, định nghĩa về trí tuệ nhân tạo 3](#_Toc114757178)

[2.2 Tình hình hoạt động chuẩn hóa liên quan đến định nghĩa, thuật ngữ về Trí tuệ nhân tạo 7](#_Toc114757179)

[2.3 Nhận xét về hiện trạng hoạt động chuẩn hóa thuật ngữ, khái niệm trong Ai của các tổ chức tiêu chuẩn lớn trên thế giới 14](#_Toc114757180)

[3 Sở cứ xây dựng các yêu cầu kỹ thuật 15](#_Toc114757181)

[3.1 Sở cứ chung để đưa ra các yêu cầu 15](#_Toc114757182)

[3.2 Lựa chọn tài liệu, tiêu chuẩn tham chiếu để xây dựng các yêu cầu kỹ thuật về chuẩn hóa thuật ngữ, khái niệm trong AI 18](#_Toc114757183)

[4 Giải thích nội dung dự thảo tiêu chuẩn 22](#_Toc114757184)

[4.1 Xây dựng dự thảo tiêu chuẩn 22](#_Toc114757185)

[4.2 Các nội dung chính của dự thảo tiêu chuẩn 22](#_Toc114757186)

[5 Khuyến nghị áp dụng tiêu chuẩn 30](#_Toc114757187)

[6 Kiến nghị về hướng phát triển tiếp theo 31](#_Toc114757188)

**DANH MỤC BẢNG BIỂU**

[Bảng 1 – Bảng đối chiếu nội dung dự thảo tiêu chuẩn với tài liệu tiêu chuẩn tham chiếu ISO/IEC 22989:2022 26](#_Toc115101889)

# Tên gọi và ký hiệu của dự thảo TCVN

## Tên tiếng Việt của dự thảo TCVN

Công nghệ thông tin – Trí tuệ nhân tạo – Các khái niệm và thuật ngữ trí tuệ nhân tạo

## Tên tiếng Anh của dự thảo TCVN

Information technology — Artificial intelligence — Artificial intelligence concepts and terminology

# Đặt vấn đề

## Sự cần thiết xây dựng tiêu chuẩn thuật ngữ, định nghĩa về trí tuệ nhân tạo

Trí tuệ nhân tạo (AI) không phải là một trào lưu công nghệ mới gần đây, nó được hình thành, xây dựng và phát triển trong suốt 60 năm qua. Thực tiễn phát triển đó cho thấy rằng AI một ý tưởng công nghệ xuất phát từ việc đề ra mục tiêu thực hiện của nó là dựa vào các loại hình công nghệ tiên tiến (hay chính xác hơn là tập hợp các công nghệ đã có hoặc các công nghệ mới nổi được phát triển theo thời gian) để tạo dựng các hệ thống máy móc nhằm hỗ trợ, thay thế khả năng nhận thức, hành vi của con người hoặc máy móc thế hệ cũ. Nó được hình thành từ khì các yếu tố công nghệ còn chưa có thể đáp ứng được những mục đích đề ra. Do đó trong những năm 1980 khi sự phát triển công nghệ chưa cho phép, các ý tưởng về thực hiện AI gặp khó khăn về tính khả thi, các khoản đầu từ nghiên cứu trong lĩnh vực này bị ngừng trệ (giai đoạn mùa đông của AI). Trong giai đoạn hiện nay, các thành tựu công nghệ trong lĩnh vực công nghệ điện tử, công nghệ thông tin, truyền thông, tự động hóa, công nghệ vật liệu mới, công nghệ sinh học v.v.. đã tạo tiền đề cho AI thực hiện các mục tiêu công nghệ của nó. Các chip vi xử lý cho phép tạo ra các hệ thống điện toán, máy tính di động, cầm tay, các thiết bị nhỏ gọn có khả năng đáp ứng xử lý các khối lượng thông tin, thuật toán, phần mềm cho hầu hết các yêu cầu đề ra. Sự phát triển bùng nổ của Internet, của cách mạng công nghệ 4.0, chuyển đổi số trong các lĩnh vực kinh tế, xã hội đã sản sinh ra lượng dữ liệu số khổng lồ cần phải xử lý trong các lĩnh vực, theo các mục đích khác nhau để tạo ra giá trị mới. Như vậy trong giai đoạn hiện nay và sắp tới, các yếu tiền đề tạo môi trường cho AI đã sẵn sàng để nó phát triển một cách bùng nổ. Chính vì vậy AI được coi là một trong các trụ cột công nghệ chính để phát triển cuộc cách mạng 4.0 và chuyển đổi số.

Trí tuệ nhân tạo ngày nay đang ngày càng tham gia nhiều hơn vào cuộc sống của chúng ta nhưng các quy định pháp lý về nó đang phát triển khá chậm. Một trong những lý do cản trở quá trình phát triển của nó cũng như việc xây dựng các quy định quản lý là sự không rõ ràng, không nhất quán về các định nghĩa, khái niệm trong lĩnh vực này. Chẳng hạn chỉ đối với cái gọi là “trí tuệ nhân tạo” hiện tồn tại rất nhiều các định nghĩa, khái niệm khác nhau. Đi cùng với nó là các định nghĩa, khái niệm liên quan tới các công nghệ, ứng dụng, các quy định luật pháp quản lý trong lĩnh vực này. Điều này tùy thuộc cách thức tiếp cận, nhìn nhận ở các khía cạnh khác nhau về nó. Một số nhà nghiên cứu cho rằng “trí tuệ nhân tạo” là khái niệm không phù hợp và không phù hợp, thay vào đó thuật ngữ này được đề xuất đổi thành “phần mềm thông minh” cho nhóm công nghệ được gọi là AI hiện nay. Thuật ngữ như vậy bao gồm các loại thuật toán, chương trình và phương pháp khác nhau được sử dụng để phân tích dữ liệu và được công nhận là công nghệ có khả năng hoạt động “có ý thức”: có thể học hỏi, tự cải thiện, đưa ra quyết định v.v.. Theo các chuyên gia này, bản thân trí tuệ nhân tạo chỉ là một phương thức thể hiện sự thông minh của các chức năng phần mềm tiên tiến. Cùng với sự phát triển của trí tuệ nhân tạo; khái niệm “trí tuệ nhân tạo” và các định nghĩa, khái niệm khác liên quan đang dần được đưa ra, được định nghĩa, giải thích. Các thuật ngữ, khái niệm đó có thể đầu tiên xuất hiện trong các bài báo khoa học, các báo cáo, thông báo kỹ thuật, các bài viết v.v.., và chúng dần được sáng tỏ trong quá trình phát triển công nghệ, thiết kế, xây dựng, triển khai và vận hành AI trong thực tiễn.

Nhu cầu thiết lập danh mục thống nhất, mang tính chuẩn hóa về từ vựng, định nghĩa, khái niệm liên quan đến AI bắt nguồn từ việc trong giai đoạn khởi đầu nó không có sự đồng thuận, hoặc tồn tại các cách thức tiếp cận định nghĩa, phân loại thuật ngữ AI khác nhau. Điều này sẽ cản trở việc đạt được mục tiêu về hợp tác, phát triển, ứng dụng, quản lý giám sát các hệ thống sản phẩm và dịch vụ AI. Việc thiết lập các định nghĩa, thuật ngữ và khái niệm trong AI sẽ cung cấp một nhận thức, phân loại theo quan điểm chung, bao quát toàn bộ bối cảnh phát triển và ứng dụng của AI, cả ở trong lĩnh vực, môi trường học thuật, nghiên cứu phát triển, sản xuất kinh doanh, công nghiệp, cộng đồng xã hội.

AI thường được mô tả liên quan đến trí thông minh của con người, hoặc trí thông minh do con người tạo ra bằng hệ thống phần mềm, máy móc. Nó đề cập đến các khái niệm, định nghĩa mô tả máy móc hoạt động giống như con người hoặc có khả năng thực hiện các hành động đòi hỏi trí thông minh. Vì trí thông minh của con người cũng rất khó để xác định và đo lường, do vậy mặc dù đã có những nỗ lực định lượng khác nhau, nhưng để định nghĩa một cách khách quan về một vấn đề mang tính chủ quan và trừu tượng như trí thông minh, tạo ra ấn tượng về một độ chính xác là khó có thể đạt được. Do đó hầu hết các định nghĩa, khái niệm được tìm thấy trong báo cáo nghiên cứu, chính sách hoặc thị trường đều mang tính mơ hồ và đề xuất một mục tiêu đạt được về mặt lý thuyết hơn là một khái niệm nghiên cứu có thể đo lường được. Nghiên cứu các tài liệu giúp chúng ta xác định bốn đặc điểm thường được đề cập đến trong việc mô tả các định nghĩa và khái niệm về AI:

1. Nhận thức về môi trường và sự phức tạp của thế giới thực.
2. Xử lý thông tin: thu thập và giải thích đầu vào
3. Ra quyết định, bao gồm suy diễn, học tập và thực hiện các hành động.
4. Đạt được các mục tiêu đã xác định trước.

Nhóm chuyên gia cao cấp về AI của Ủy ban châu Âu cho rằng việc xác định các định nghĩa, thuật ngữ thống nhất trong AI cần phải tiếp cận từ việc coi hệ thống trí tuệ nhân tạo là hệ thống phần mềm và có thể cả phần cứng được thiết kế bởi con người để thực hiện một tập hợp các mục tiêu tương đối phức tạp, thực hiện bằng các giải pháp vật lý hoặc kỹ thuật số để nhận thức môi trường hoạt động của chúng thông qua thu thập dữ liệu, diễn giải dữ liệu có cấu trúc hoặc phi cấu trúc, suy luận các kiến ​​thức hoặc thông tin đã được xử lý để đưa ra các quyết định và hành động tốt nhất đối với sự thay đổi của môi trường hoạt động. Các thuật ngữ, định nghĩa có thể sử dụng các quy tắc biểu trưng hoặc học tập thông qua các mô hình số để điều chỉnh hành vi của chúng cho phù hợp với cá mục tiêu thực hiện.

Một khía cạnh quan trọng của việc làm rõ các định nghĩa, thuật ngữ trong lĩnh vực AI đó là tác động của công nghệ này đối với hệ thống luật pháp hiện hành. Đặc thù của các ứng dụng AI có nhiều điểm khác với các ứng dụng, dịch vụ công nghệ thông tin truyền thống. Nó cung cấp các chức năng xử lý thông tin, đưa ra các dự đoán hoặc quyết định hoặc hành động trợ giúp hoặc thay thế hẳn con người ở các lĩnh vực ứng dụng của nó. Ảnh hưởng của AI đối với luật pháp và ngược lại là rất quan trọng: luật pháp thiết lập các quy tắc có thể thúc đẩy hoặc ngừng phát triển hơn nữa của công nghệ; AI tham gia vào các trường hợp khác nhau liên quan đến thị trường tài chính, thao túng xã hội và chính trị, thực thi pháp luật, vận tải, v.v. Lĩnh vực ứng dụng của AI được kỳ vọng là rất rộng do đó trong tương lai gần nó có thể ảnh hưởng tới mọi mặt của đời sống kinh tế xã hội của một quốc gia.

Do vậy AI bây giờ không còn chỉ là một khái niệm từ toán học hay tin học thuận túy về lý thuyết hoặc kỹ thuậ nữa - các công nghệ được phát triển, ứng dụng và trở thành một phần của cuộc sống hàng ngày của chúng ta. Các nhà lập pháp từ khắp nơi trên thế giới và các thành viên của các tổ chức quốc tế hiện đang thảo luận về chủ đề này một cách thực tế và nghiêm túc để cố gắng trả lời những câu hỏi: những công nghệ này là gì, những nguyên tắc nào cần được đặt lên trên sự phát triển và ứng dụng của chúng, luật pháp phải trả lời như thế nào các mối quan hệ đang thay đổi do tác động của AI.

Khi nói đến quy định về AI, các bên liên quan cần hiểu liệu quy định đó có áp dụng cho một tổ chức, sản phẩm hoặc dịch vụ hay không. Các tổ chức khác nhau có thể có bộ thuật ngữ và định nghĩa khác nhau, phạm vi đề cập có thể hữu hạn tùy thuộc vào quan điểm, kỳ vọng và mục đích của họ. Bất kỳ cách tiếp cận như thế nào thì các định nghĩa, khái niệm cần được xây dựng xung quanh các thuộc tính cụ thể của hệ thống (ví dụ: học máy; suy luận thống kê) và ngữ cảnh ứng dụng. Do đó theo ý kiến của các chuyên gia trong lĩnh vực AI thì cách tiếp cận khác nhau về xây dựng các định nghĩa, thuật ngữ đều có thể phát sinh từ sự hiểu lầm hoặc thậm chí là không hiểu được các định nghĩa, khái niệm về trí tuệ nhân tạo cũng như đối tượng, phạm vi áp dụng của nó. Để có thể đưa ra được các định nghĩa, khái niệm về AI thì có thể cần phải xây dựng danh mục các thuật ngữ đại diện cho các lĩnh vực, khía cạnh thực hiện của AI. Danh mục các thuật ngữ này có thể được thu thập, tổng hợp từ các tài liệu, kết quả nghiên cứu, báo cáo kỹ thuật, các tài liệu kỹ thuật về sản phẩm dịch vụ AI, các tiêu chuẩn liên quan đã được ban hành. Sau đó, danh sách các thuật ngữ, khái niệm được các chuyên gia AI ở mọi lĩnh vực xem xét chấp thuận về cách hiểu mang tính đồng thuận đối với một bản rút gọn trên cơ sở các thuật ngữ, định nghĩa, khái niệm đã được đưa ra bàn thảo. Các công việc về phân loại thuật ngữ, lĩnh vực áp dụng cũng được thực hiện để tạo thuận lợi cho các đối tượng khác nhau sử dụng theo từng khía cạnh, lĩnh vực phù hợp.

Một vấn đề quan trọng đối với khuôn khổ quy định cụ thể trong tương lai về trí tuệ nhân tạo là xác định phạm vi áp dụng của nó. Nếu như đưa ra được các định nghĩa rõ ràng về các sản phẩm và dịch vụ để từ đó áp dụng các khuôn khổ pháp lý có liên quan hoặc cần thiết phải xây dựng mới thực sự là cần thiết để có được sự chắc chắn về mặt pháp lý cho tất cả các bên liên quan trong thực thi và triển khai AI.

Trong tương lai, những rủi ro và thách thức mới đối với việc phân bổ trách nhiệm pháp lý có thể xuất hiện với sự phức tạp ngày càng tăng của các ứng dụng AI. Ví dụ như độ chính xác của AI trong việc điều khiển các hệ thống thời gian thực, chẳng hạn như xe tự lái, có thể cần sớm được định nghĩa một cách rõ ràng để mang lại sự chắc chắn về mặt pháp lý cho việc phát triển và sử dụng AI trong lĩnh vực cụ thể này. Một thách thức pháp lý cụ thể đối với việc phát triển và sử dụng phần mềm AI có thể nằm ở việc phụ thuộc vào các mô-đun phần mềm của bên thứ ba với số lượng mã không đồng nhất ngày càng tăng, nơi có thể cần đến các công cụ xác định nguồn gốc. Các tiêu chuẩn về định nghĩa, thuật ngữ có tính rõ ràng, minh bạch có thể được sử dụng làm công cụ để hỗ trợ việc áp dụng các quy tắc trách nhiệm trong điều kiện thực hành tốt nhất và trong khuôn khổ xác định nghĩa vụ về mặt pháp lý đối với các bên liên quan trong hệ sinh thái AI.

Như vậy, cũng như đối với mỗi một loại hình, lĩnh vực công nghệ nào đó đều cần phải có một tập hợp các khái niệm, định nghĩa mang tính chuyên môn phù hợp với nó. Điều này cũng đúng đối với sự phát triển của AI. Nó cần thiết để có các nhận thức chung, mang tính nhất quán đối với mọi đối tượng tham gia vào hệ sinh thái phát triển công nghệ đó. Chẳng hạn như các nhà nghiên cứu, phát triển, thiết kế AI cần phải thống nhất về các định nghĩa, khái niệm, chức năng mang tính nguyên tắc, nguyên lý, phương pháp tiếp cận để nghiên cứu phát triển các kỹ thuật, thiết kế các thành phần cho AI. Các bên liên quan đến ứng dụng và vận hành quản lý cần phải nhận thức, xác định, phân biệt rõ các định nghĩa, khái niệm, đặc điểm của thành phần, chức năng của AI trongc quy trình hoạt động, ứng dụng trong từng lĩnh vực của nó. Các nhà lập pháp, quản lý nhà nước cần phải được định nghĩa rõ ràng các đối tượng cần quản lý trong AI cũng như mối liên quan đến các đối tượng, bên liên quan khác khi đưa các hệ thống, sản phẩm cung cập dịch vụ AI cho xã hội.

Do vậy mục tiêu xây dựng tiêu chuẩn nhằm đáp ứng yêu cầu về cung cấp thông tin, thông hiểu; phổ biến và tăng cường hiểu biết về các quá trình, kỹ thuật, công nghệ; các vấn đề liên quan đến hệ thống, sản phẩm, dịch vụ trong lĩnh vực AI cho các đối tượng áp dụng tiêu chuẩn.

## Tình hình hoạt động chuẩn hóa liên quan đến định nghĩa, thuật ngữ về Trí tuệ nhân tạo

### Hoạt động chuẩn hóa của Viện Kỹ thuật Điện và Điện tử (IEEE) của Mỹ

Các hoạt động chuẩn hóa trong lĩnh vực AI nói chung và trong lĩnh vực thuật ngữ, định nghĩa của AI nói riêng được thực hiện bởi nhóm Sáng kiến toàn cầu về các hệ thống thông minh tự động (The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems) và nhóm Sáng kiến về các hệ thống trí tuệ nhân tạo (Artificial Intelligence System (AIS)) thực hiện. Năm 1996 IEEE đã thực hiện một dự án (IEEE P610) chuẩn hóa về trí tuệ nhân tạo. Một trong những kết quả của dự án này là ban hành tài liệu phục vụ tiêu chuẩn hóa (IEEE.610.8) xuất bản lần cuối vào năm2016 về từ vựng về trí tuệ nhân tạo và ứng dụng người máy. Tuy nhiên kể từ đó đến nay tài liệu phục vụ tiêu chuẩn hóa này không có bất cứ hoạt động sửa đổi cập nhật nào trong khi số lượng thuật ngữ, định nghĩa về lĩnh vực này đã phát triển rất nhiều cùng với sự phát triển của các công nghệ, ứng dụng và quản lý trong lĩnh vực này.

Các dự án xây dựng tiêu chuẩn về AI hiện nay của IEEE không hướng tới việc xây dựng tiêu chuẩn mang tính tổng thể và toàn diện về định nghĩa, thuật ngữ về trong lĩnh vực trí tuệ nhân tạo. Thay vào đó các định nghĩa, thuật ngữ về AI được xây dựng theo từng khía cạnh, vấn đề cần chuẩn hóa của thể. Hoặc là chúng dược bao gồm trong các nội dung “thuật ngữ và định nghĩa của các tiêu chuẩn cụ thể được ban hành.

Dưới đây là một số tiêu chuẩn, tài liệu chứa đựng các nội dung liên quan đến thuật ngữ, định nghĩa về AI của IEEE:

#### IEEE 610.8

**Tên tài liệu:** Từ vựng về trí tuệ nhân tạo và các ứng dụng người máy, tái bản lần thứ 6, 1996 (Glossary of Artificial Intelligence and Robotics Applications)

**Tình trạng:** Tài liệu của nhóm nghiên cứu về từ vựng kỹ thuật phần mềm của IEEE (IEEE Standard Glossary of Software Engineering Terminology) công bố năm 1996.

Tài liệu đưa ra các khái niệm, thuật ngữ liên quan tới lĩnh vực trí tuệ nhân tạo và ứng dụng người máy. Tuy nhiên kể từ lần tái bản làn cuối năm 1996, tài liệu này không được sửa đổi cập nhật thêm bất cứ nội dung nào

#### Báo cáo kỹ thuật của nhóm Sáng kiến thiết kế phù hợp với đạo đức

**Tên tài liệu:** Thiết kế phù hợp với đạo đức – Từ vựng xuất bản lần thứ nhất (Ethically aligned design – First edition Glossary)

**Tình trạng:** Tài liệu báo cáo kỹ thuật của nhóm Sáng kiến về thiết kế phù hợp với đạo đức (IEEE Global Initiative for Ethically Aligned Design) xuất bản năm 2018.

Tài liệu cung cấp các khái niệm, định nghĩa phổ biến được sử dụng bởi các chuyên gia trong trong lĩnh vực hệ thống thông minh và tự động, hướng tới việc thiết kế các hệ thống có tính đạo đức. Mục tiêu của về đưa ra đề xuất về thuật ngữ, định nghĩa là nhằm đạt được sự đồng thuận của các thành viên trong nhóm sáng kiến này, để các nhóm sáng kiến khác có thể sử dụng dựa trên bối cảnh của từng lĩnh vực cụ thể trong AI.

#### Các tài liệu tiêu chuẩn có các nội dung thuật ngữ, định nghĩa liên quan tới AI

Ngoài ra các thuật ngữ, định nghĩa chung hoặc liên quan đến các lĩnh vực tiêu chuẩn hóa cụ thể trong AI được bao hàm trong nội dung “định nghĩa và thuật ngữ” của các tiêu chuẩn được ban hành dưới đây:

* IEEE 2941-2021: Tiêu chuẩn IEEE về trình bày, nén, phân phối và quản lý mô hình trí tuệ nhân tạo (IEEE Standard for Artificial Intelligence (AI) Model Representation, Compression, Distribution, and Management)
* IEEE 2801-2022: Thực hành được đề xuất của IEEE để quản lý chất lượng bộ dữ liệu cho trí tuệ nhân tạo trong y tế (IEEE Recommended Practice for the Quality Management of Datasets for Medical Artificial Intelligence)
* IEEE 1232.3-2014: Hướng dẫn IEEE về việc sử dụng trao đổi trí tuệ nhân tạo và ràng buộc dịch vụ với tất cả các môi trường thử nghiệm (IEEE Guide for the Use of Artificial Intelligence Exchange and Service Tie to All Test Environments (AI-ESTATE)).
* IEEE 7010-2020: Thực hành được đề xuất của IEEE để đánh giá tác động của các hệ thống tự trị và thông minh đối với sức khỏe con người (IEEE Recommended Practice for Assessing the Impact of Autonomous and Intelligent Systems on Human Well-Being)
* IEEE 3652.1-2020: Hướng dẫn IEEE chok khung kiến trúc và ứng dụng học máy liên kết (EEE Guide for Architectural Framework and Application of Federated Machine Learning)

### Liên minh viễn thông quốc tế ITU-T

Liên minh viễn thông quốc tế ITU-T đã thành lập các nhóm nghiên cứu trọng điểm (Focus Group – FG) để xúc tiến các hoạt động chuẩn hóa về AI. Tuy nhiên hiện tại ITU-T không có các hoạt động chuyên sâu về chuẩn hóa các thuật ngữ và định nghĩa trong lĩnh vực trí tuệ nhân tạo. Các nhóm nghiên cứu trọng điểm của AI, chẳng hạn như Nhóm nghiên cứu về AI trong y tế (AI4H), Nhóm nghiên cứu về AI trong trợ giúp và tự động lái xe (FG- FGAI4AD), Nhóm nghiên cứu về bảo mật trong AI (SAI) đã công bố các báo cáo kỹ thuật, tài liệu phục vụ chuẩn hóa trong từng lĩnh vực mà nó phụ trách. Một phần các nội dung tài liệu nói trên được dụng để ban hành hoặc đưa vào bổ sung cho các khuyến nghị đã có liên quan đến AI.

Phần lớn thuật ngữ, định nghĩa liên quan đến AI được ban hành trong các khuyến nghị của ITU-T được bao hàm trong các nội dung “định nghĩa – definitions), “khái niệm – terms” theo bố cục của các khuyến nghị này. Qua khảo sát nhóm thực hiện nhiệm vụ thấy rằng: phần lớn các định nghĩa, khái niệm trong các khuyến nghị được ITU-T ban hành là được tham chiếu từ các tài liệu, tiêu chuẩn của ISO/IEC. Các khái niệm được nêu có thể là các định nghĩa, thuật ngữ, khái niệm chung hoặc chuyên sâu tuy thuộc vào mục tiêu chuẩn hóa của từng khuyến nghị đưa ra.

Các khuyến nghị điển hình của ITU-T ban hành liên quan tới đưa ra các định nghĩa, thuật ngữ, khái niệm về AI là:

#### F.749.4 (06/2021)

**Tên khuyến nghị:** Các trường hợp sử dụng và yêu cầu đối với hệ thống xe hỗ trợ truyền thông đa phương tiện sử dụng trí tuệ nhân tạo (Use cases and requirements for multimedia communication enabled vehicle systems using artificial intelligence).

**Tình trạng:** Khuyến nghị được ban hành chính thức

Khuyến nghị có nội dung đưa ra các định nghĩa, thuật ngữ, khái niệm liên quan tới AI

#### F.749.13 (06/2021)

**Tên khuyến nghị:** Khuôn khổ và yêu cầu đối với điều khiển máy bay không người lái dân dụng sử dụng trí tuệ nhân tạo (Framework and requirements for civilian unmanned aerial vehicle flight control using artificial intelligence).

**Tình trạng:** Khuyến nghị được ban hành chính thức

Khuyến nghị có nội dung đưa ra các định nghĩa, thuật ngữ, khái niệm liên quan tới AI

#### L.1305 (11/2019)

**Tên khuyến nghị:** Hệ thống quản lý cơ sở hạ tầng trung tâm dữ liệu dựa trên dữ liệu lớn và công nghệ trí tuệ nhân tạo (Data centre infrastructure management system based on big data and artificial intelligence technology).

**Tình trạng:** Khuyến nghị được ban hành chính thức

Khuyến nghị có nội dung đưa ra các định nghĩa, thuật ngữ, khái niệm liên quan tới AI

#### M.3080 (02/2021)

**Tên khuyến nghị:** Khung vận hành và quản lý viễn thông được tăng cường bởi trí tuệ nhân tạo (AITOM) (Framework of artificial intelligence enhanced telecom operation and management (AITOM)).

**Tình trạng:** Khuyến nghị được ban hành chính thức

Khuyến nghị có nội dung đưa ra các định nghĩa, thuật ngữ, khái niệm liên quan tới AI

#### M.3381 (01/2022)

**Tên khuyến nghị:** Yêu cầu đối với quản lý tiết kiệm năng lượng của hệ thống mạng truy cập vô tuyến (RAN) 5G với trí tuệ nhân tạo (AI) (Requirements for energy saving management of 5G radio access network (RAN) systems with artificial intelligence (AI)).

**Tình trạng:** Khuyến nghị được ban hành chính thức

Khuyến nghị có nội dung đưa ra các định nghĩa, thuật ngữ, khái niệm liên quan tới AI

#### M.3382 (06/2022)

**Tên khuyến nghị:** Yêu cầu đối với xử lý trình tự công việc trong quản lý viễn thông với trí tuệ nhân tạo (Requirements for work order processing in telecom management with artificial intelligence).

**Tình trạng:** Khuyến nghị được ban hành chính thức.

Khuyến nghị có nội dung đưa ra các định nghĩa, thuật ngữ, khái niệm liên quan tới AI

#### Y.3116 (02/2022)

**Tên khuyến nghị:** Quản lý IMT-2020 định dạng hóa lưu lượng dựa trên phương pháp tiếp cận bằng trí tuệ nhân tạo (Traffic typization IMT-2020 management based on an artificial intelligence approach)

Khuyến nghị có nội dung đưa ra các định nghĩa, thuật ngữ, khái niệm liên quan tới AI

#### Y.3177 (02/2021)

**Tên khuyến nghị:** Khung kiến ​​trúc cho tự động hóa mạng dựa trên trí tuệ nhân tạo để quản lý tài nguyên và lỗi trong các mạng tương lai bao gồm cả IMT-2020 (Architectural framework for artificial intelligence-based network automation for resource and fault management in future networks including IMT-2020).

**Tình trạng:** Khuyến nghị được ban hành chính thức

Khuyến nghị có nội dung đưa ra các định nghĩa, thuật ngữ, khái niệm liên quan tới AI

#### Y.3178 (07/2021)

**Tên khuyến nghị:** Khung chức năng của cung cấp dịch vụ mạng dựa trên trí tuệ nhân tạo trong các mạng tương lai bao gồm cả IMT-2020 (Functional framework of artificial intelligence-based network service provisioning in future networks including IMT-2020).

**Tình trạng:** Khuyến nghị được ban hành chính thức

Khuyến nghị có nội dung đưa ra các định nghĩa, thuật ngữ, khái niệm liên quan tới AI

#### Y.4470 (08/2020)

**Tên khuyến nghị:** Kiến trúc tham chiếu về tiếp xúc với dịch vụ trí tuệ nhân tạo cho các thành phố bền vững thông minh (Reference architecture of artificial intelligence service exposure for smart sustainable cities).

**Tình trạng:** Khuyến nghị được ban hành chính thức

Khuyến nghị có nội dung đưa ra các định nghĩa, thuật ngữ, khái niệm liên quan tới AI

Ngoài ra ITU duy trì một cơ sở dữ liệu trực tuyến về từ vựng cho tất cả các lĩnh vực mà các khuyến nghị ban hành đề cập tới, có thể truy cập theo đường liên kết: [https://www.itu.int/br\_tsb\_terms/#/](https://www.itu.int/br_tsb_terms/%23/)

### Viện Tiêu chuẩn Viễn thông châu Âu ETSI

Có thể nói hoạt động chuẩn hóa chuyên sâu trong lĩnh vực định nghĩa, thuật ngữ của tổ chức này còn chưa được chú trọng triển khai. Điều này xuất phát từ việc ETSI hiện mới chỉ tập trung vào hoạt động chuẩn hóa trong vấn đề bảo mật AI với hoạt động của Nhóm nghiên cứu đặc biệt về bảo mật trí tuệ nhân tạo (ISG SAI). Cũng tương tự như đối với tổ chức ITU-T; các thuật ngữ, định nghĩa, khái niệm liên quan đến AI được bao hàm trong các tiêu chuẩn, thông báo kỹ thuật hoặc tài liệu đặc tính kỹ thuật được ETSI ban hành. Thông thường phần này thường được đưa vào bố cục nội dung của các tài liệu nói trên ở phần “Định nghĩa các thuật ngữ, ký hiệu và chữ viết tắt (Definition of terms, symbols and abbreviations). Các thuật ngữ, định nghĩa này có thể là các định nghĩa chung hoặc định nghĩa mang tính đặc thù của lĩnh vực AI mà các tài liệu đề cập đến.

Một số tiêu chuẩn, tài liệu kỹ thuật mà ETSI ban hành dưới đây chứa đựng các nội dung về thuật ngữ, định nghĩa về AI.

#### ETSI GR ENI 004 V2.2.1 (2021-12)

**Tên tài liệu:** Mạng thông minh dựa trên kinh nghiệm (ENI); Thuật ngữ cho các khái niệm chủ yếu trong ENI (Experiential Networked Intelligence (ENI); Terminology for Main Concepts in ENI)

**Tình trạng:** Thông báo kỹ thuật của nhóm nghiên cứu Mang thông minh dựa trên kinh nghiệm (ENI) công bố

Tài liệu này chứa các thuật ngữ, định nghĩa liên quan đến AI trong mục “Định nghĩa các thuật ngữ, ký hiệu và chữ viết tắt (Definition of terms, symbols and abbreviations). Đây cũng là nội dung chính của tài liệu này (định nghĩa cho hơn 50 thuật ngữ thuộc lĩnh vực này)

#### ETSI TR 103 748 V1.1.1 (2022-06)

**Tên tài liệu:** Kiểm tra mạng lõi và khả năng tương tác (INT); Trí tuệ nhân tạo (AI) trong hệ thống kiểm tra và kiểm tra mô hình AI; Sử dụng và lợi ích của công nghệ AI trong hoạt động kiểm tra (Core Network and Interoperability Testing (INT); Artificial Intelligence (AI) in Test Systems and Testing of AI Models; Use and Benefits of AI Technologies in Testing)

**Tình trạng:** Báo cáo kỹ thuật ban hành chính thức bởi ETSI

Tài liệu này chứa các thuật ngữ, định nghĩa liên quan đến AI trong mục “Định nghĩa các thuật ngữ, ký hiệu và chữ viết tắt (Definition of terms, symbols and abbreviations).

#### ETSI GR SAI 006 V1.1.1 (2022-03)

**Tên tài liệu:** Bảo mật Trí tuệ nhân tạo (SAI); Vai trò của phần cứng trong bảo mật AI (Securing ArtificiaI Intelligence (SAI); The role of hardware in security of AI)

**Tình trạng:** Báo cáo kỹ thuật của nhóm SAI

Tài liệu này chứa các thuật ngữ, định nghĩa liên quan đến AI trong mục “Định nghĩa các thuật ngữ, ký hiệu và chữ viết tắt (Definition of terms, symbols and abbreviations).

#### ETSI GR SAI 001 V1.1.1 (2022-01)

**Tên tài liệu:** Bảo mật Trí tuệ nhân tạo (SAI); Bản thể luận tấn công AI (Securing Artificial Intelligence (SAI); AI Threat Ontology)

**Tình trạng:** Báo cáo kỹ thuật của nhóm SAI.

Tài liệu này chứa các thuật ngữ, định nghĩa liên quan đến AI trong mục “Định nghĩa các thuật ngữ, ký hiệu và chữ viết tắt (Definition of terms, symbols and abbreviations).

#### ETSI GR SAI 002 V1.1.1 (2021-08)

**Tên tài liệu:** Bảo mật Trí tuệ nhân tạo (SAI); Bảo mật chuỗi cung ứng dữ liệu (Securing Artificial Intelligence (SAI); Data Supply Chain Security).

**Tình trạng:** Báo cáo kỹ thuật của nhóm SAI

Tài liệu này chứa các thuật ngữ, định nghĩa liên quan đến AI trong mục “Định nghĩa các thuật ngữ, ký hiệu và chữ viết tắt (Definition of terms, symbols and abbreviations).

#### ETSI GR SAI 005 V1.1.1 (2021-03)

**Tên tài liệu:** Bảo mật Trí tuệ nhân tạo (SAI); Báo cáo chiến lược giảm thiểu (Securing Artificial Intelligence (SAI); Mitigation Strategy Report)

Tình trạng: Báo cáo kỹ thuật của nhóm SAI

Tài liệu này chứa các thuật ngữ, định nghĩa liên quan đến AI trong mục “Định nghĩa các thuật ngữ, ký hiệu và chữ viết tắt (Definition of terms, symbols and abbreviations).

Ngoài ra ETSI duy trì một cơ sở dữ liệu trực tuyến về từ vựng cho tất cả các lĩnh vực mà các tiêu chuẩn, tài liệu đặc tính kỹ thuật, báo cáo kỹ thuật được tổ chức này ban hành, truy cập theo đường liên kết: <https://webapp.etsi.org/Teddi/>

### ISO/IEC

Các hoạt động xúc tiến chuẩn hóa về AI được thực hiện chủ yếu bởi nhóm nghiên cứu SC42 (chuyên về chuẩn hóa trong lĩnh vực AI và dữ liệu lớn) trực thuộc Ban kỹ thuật phối hợp 1 giữa ISO và IEC. Nhóm nghiên cứu này của ISO/IEC đã ban hành chính thức 13 tiêu chuẩn liên quan đến trí tuệ nhân tạo và dữ liệu lớn, 26 tiêu chuẩn đang được xây dựng và trong quá trình tiến hành các bước thẩm định để ban hành.

Các tiêu chuẩn ban hành và đang trong quá trình xây dựng đều đưa các thuật ngữ và định nghĩa (mới hoặc tham chiếu) liên quan tới AI. Các thuật ngữ và định nghĩa được bao hàm trong mục “Các thuật ngữ và định nghĩa” (Terms and definitions) là nội dung bắt buộc theo bố cục tiêu chuẩn của ISO/IEC.

Đặc biệt, trong các tiêu chuẩn đã được ban hành thì tiêu chuẩn dưới đây được lưu ý nhất về tính chuyên sâu và toàn diện của nó về lĩnh vực chuẩn hóa thuật ngữ, định nghĩa về AI.

#### ISO/IEC 22989:2022

**Tên tiêu chuẩn:** Công nghệ thông tin – Trí tuệ nhân tạo – Các khái niệm và thuật ngữ trí tuệ nhân tạo (Information technology — Artificial intelligence — Artificial intelligence concepts and terminology).

**Tình trạng:** Tiêu chuẩn đã được ban hành, công bố chính thức là tiêu chuẩn quốc tế.

Tiêu chuẩn này đề cập một cách toàn diện về chuẩn hóa thuật ngữ, định nghĩa và khái niệm về AI; bao gồm cả việc phân loại theo các lĩnh vực, khía cạnh thực hiện khác nhau như kỹ thuật học máy, dữ liệu, vòng đời đặc thù của AI; các khái niệm về chức năng, hệ sinh thái AI v.v..

Tổng cộng tiêu chuẩn này đã đưa ra gần 130 thuật ngữ và rất nhiều các khái niệm liên quan đến các vấn để như đã đề cập ở trên. Phần lớn các định nghĩa, thuật ngữ, khái niệm có trong tiêu chuẩn này được tổng hợp, phân loại từ các tiêu chuẩn, tài liệu khác của ISO/IEC.

## Nhận xét về hiện trạng hoạt động chuẩn hóa thuật ngữ, khái niệm trong Ai của các tổ chức tiêu chuẩn lớn trên thế giới

Những thông tin được khảo sát ở trên về hiện trạng tiêu chuẩn hóa về thuật ngữ, khái niệm trong AI của các tổ chức tiêu chuẩn hóa trên thế giới (IEEE, ITU-T, ETSI, ISO/IEC) cho thấy:

* Hoạt động tiêu chuẩn hóa thuật ngữ, định nghĩa, khái niệm trong lĩnh vực AI đã được các tổ chức triển khai chính thức trong những năm gần đây. Tuy nhiên, tuy từng mục tiêu hoạt động chuẩn hóa trong lĩnh vực này mà các tổ chức có cách tiếp cận khác nhau
* Đối với các tổ chức tiêu chuẩn IEEE, ITU-T, ETSI không xúc tiến dự án xây dựng và ban hành tiêu chuẩn riêng về thuật ngữ, định nghĩa, khái niệm trong lĩnh vực AI. Các thuật ngữ về AI liên quan đến lĩnh vực cụ thể được xác định và bao hàm trong các tiêu chuẩn, đặc tính kỹ thuật, thông báo kỹ thuật được ban hành. Chẳng hạn như các định nghĩa, thuật ngữ về AI được hiểu trong các lĩnh vực cụ thể như an toàn, bảo mật thông tin; y tế, lái xe tự động. Do vậy việc xác định các thuật ngữ, khái niệm được hiểu trong các ngữ cảnh tổng quát, được phân loại có tính hệ thống về công nghệ, kỹ thuật, ứng dụng là chưa được đề cập.
* ISO/IEC là tổ chức tiêu chuẩn quốc tế có hoạt động xây dựng tiêu chuẩn riêng về thuật ngữ, định nghĩa và khái niệm trong AI thông qua nhóm nghiên cứu SC42. Gần như có sự ngầm định mang tính “phân công” giữa các tổ chức tiêu chuẩn quốc tế này trong việc xây dựng và ban hành tiêu chuẩn về thuật ngữ, định nghĩa và các khái niệm trong AI

# Sở cứ xây dựng các yêu cầu kỹ thuật

## Sở cứ chung để đưa ra các yêu cầu

Do tính chất quan trọng như đã đề cập ở trên về việc xác định, xây dựng khái niệm, định nghĩa; nên các tổ chức tiêu chuẩn quốc tế, các hiệp hội, tổ chức quốc tế, diễn đàn trên hoạt động liên quan đến AI đã xúc tiến các hoạt động nhằm chuẩn hóa các định nghĩa khái niệm trong lĩnh vực AI. Đối với các tổ chức tiêu chuẩn quốc tế, hoạt động chuẩn hóa các khái niệm, thuật ngữ, định nghĩa có thể được thực hiện ở các hình thức khác nhau, chẳng hạn như các định nghĩa, khái niệm, thuật ngữ có thể được nêu trong các tài liệu, báo cáo, tiêu chuẩn kỹ thuật liên quan đến từng khía cạnh, lĩnh vực AI cụ thể. Cách tiếp cận toàn diện hơn, chẳng hạn như ISO/IEC hướng tới việc xây dựng riêng về các định nghĩa, khái niệm, thuật ngữ trong lĩnh vực AI, thực chất là sự tổng hợp, phân loại các định nghĩa, khái niệm và thuật ngữ từ các tiêu chuẩn đã được ban hành trong lĩnh vực này và từ các nguồn tài liệu của các tổ chức tiêu chuẩn hoặc từ các nguồn tin cậy khác. Đối với một số tổ chức quốc tế, diễn đàn học thuật hoặc các hiệp hội công nghiệp họ có cách tiếp cận khác; đó là việc đề xuất, xây dựng danh mục từ vựng (vocabulary, glossary) về AI và phổ biến danh mục này thông qua các trang web của tổ chức, diễn đàn đó.

Hiện tại có những hình thức khác nhau để xây dựng những khái niệm, định nghĩa và thuật ngữ trong lĩnh vực AI như đề cập ở ở trên. Tuy nhiên nội hàm của những hoạt động này đều có tính hội tụ, tập trung vào những vấn đề, khía cạnh cần làm sáng tỏ trong quá trình phát triển và ứng dụng của AI. Những vấn đề đó có thể là tập hợp các khái niệm, định nghĩa mô tả khả năng mô phỏng các chức năng nhận thức của con người (phân loại mô hình học máy, phương pháp tiếp cận của các loại hình thuật toán v.v..); tổ hợp các giải pháp công nghệ bao gồm cơ sở hạ tầng thông tin và truyền thông, phần mềm (kể cả những phần mềm sử dụng phương pháp học máy), các quy trình và dịch vụ để xử lý dữ liệu. Mỗi một loại hình công nghệ đều không phải là bất biến. Trong quá trình phát triển nó luôn có các phát minh, cải tiến, phát triển, thậm chí là biến đổi vể hình thái, phương pháp, tính ứng dụng. Do vậy các định nghĩa, khái niệm, thuật ngữ trong lĩnh vực này không phải là bất biến. Chúng cần được hiệu chỉnh, sửa đổi và bổ sung theo tiến trình phát triển của công nghệ. Do vậy hoạt động chuẩn hóa trong lĩnh vực này là một quá trình theo dõi sự phát triển, tiếp thu, cải tiến và bổ sung về tri thức trong lĩnh vực AI.

Trên cơ sở đó, các hoạt động chuẩn hóa về định nghĩa, khái niệm, thuật ngữ AI hiện nay đang được tập trung vào giải quyết các yêu cầu trọng tâm dưới đây.

### Chuẩn hóa các thuật ngữ và định nghĩa về AI

Hoạt động chuẩn hóa về AI trước tiên cần phải đưa ra được các thuật ngữ liên quan đến các lĩnh vực, các đối tác liên quan tham gia vào tất cả các quá trình nghiên cứu phát triển, thiết kế, thực thi và ứng dụng và vận hành của nó trong thực tế. Công việc khởi đầu là việc thống nhất các thuật ngữ. Trong thực tế có thể một đối tượng AI có các đặc điểm, thuộc tính giống nhau nhưng tồn tại các tên gọi khác nhau (có thể thuộc về từ vựng, sự tương đồng về ngữ nghĩa). Do đó cần phải có sự đồng thuận, thống nhất về tên của các thuật ngữ đó. Điều này chỉ có thể thực hiện bằng cách chuẩn hóa chúng trong các tiêu chuẩn được xay dựng và ban hành bởi các tổ chức tiêu chuẩn có uy tín.

Các thuật ngữ, nếu như được chuẩn hóa về tên gọi thì công việc tiếp theo là phải xác định được nó là gì, ý nghĩa, mục đích sử dụng trong các trường hợp, tình huống cụ thể như thế nào. Để giải quyết vấn đề này, mỗi một thuật ngữ cần được định nghĩa một cách rõ ràng, đầy đủ thông qua các phát biểu mang tính mô tả đối tượng mà thuật ngữ đề cập tới. Định nghĩa có thể kèm theo các giải thích hoặc chú thích cần thiết, các tình huống thể hiện hoặc sử dụng để người đọc có thể hiểu và nhận thức rõ hơn về thuật ngữ đó và đối tượng mô tả của nó.

Phần lớn các thuật ngữ được xây dựng cùng với các định nghĩa đi kèm thường có phạm vi áp dụng nhất định. Nghĩa là việc hiểu và áp dụng nó cần phải đặt trong những ngữ cảnh, tình huống hoặc lĩnh vực xác định của AI. Do đó việc phân loại các thuật ngữ theo các khía cạnh thực thi của AI là cần thiết, tạo thuận lợi cho người dùng tra cứu, áp dụng trong các trường hợp cụ thể.

Theo đó, yêu cầu về chuẩn hóa thuật ngữ, định AI được xác định và phân loại một cách tương đối trong các lĩnh vực dưới đây:

* Các định nghĩa, thuật ngữ chung về AI
* Các định nghĩa, thuật ngữ liên quan đến dữ liệu trong AI
* Các định nghĩa, thuật ngữ liên quan đến học máy
* Các định nghĩa, thuật ngữ liên quan đến mạng nơ-ron
* Các định nghĩa, thuật ngữ liên quan đến tính đáng tin cậy trong AI
* Các định nghĩa thuật ngữ liên quan đến xử lý ngôn ngữ tự nhiên
* Các định nghĩa thuật ngữ liên quan đến thị giác máy tính

### Chuẩn hóa các khái niệm về AI

Thuật ngữ và định nghĩa thông thường chỉ nêu tên gọi. Định nghĩa về chúng được phát biểu ngắn gọn dưới dạng mô tả biểu hiện, tính chất bên ngoài của các đối tượng mà không đề cập chi tiết về các tính chất, thuộc tính cũng như mối liên hệ của nó đối với các đối tượng khác. Trong AI, nếu chỉ đưa ra các thuật ngữ và định nghĩa có thể là chưa đủ. Một số đối tượng trong AI tương đối phức tạp, do vậy nó không thể bao hàm được trong các thuật ngữ và định nghĩa một cách ngắn gọn về nó. Do đó các đối tượng, quá trình này trong AI cần được hiểu ở dạng các khái niệm. Cũng giống như thuật ngữ, các khái niệm này cũng được đặt tên, nghĩa là chúng có tên gọi cụ thể. Nhưng khác với định nghĩa cho thuật ngữ, các khái niệm mô tả chi tiết các thuộc tính, đặc điểm, các dạng thể hiện, có thể là sự phân loại về hình thái, phương pháp, cách thức sử dụng, các mối liên hệ, tương tác với các đối tượng hoặc quá trình khác. Do đó một khái niệm được nêu ra với cách thức như vậy sẽ làm cho người dùng hiểu biết sâu sắc và ở phạm vi rộng hơn về một đối tượng, quá trình, kỹ thuật hoặc công nghệ cụ thể nào đó trong AI.

Cũng giống như việc xây dựng các định nghĩa, thuật ngữ; việc xây dựng các khái niệm trong AI cũng cần được phân loại. Hiện nay xu hướng phân loại các khái niệm AI thường dựa trên cơ sở các kỹ thuật, công nghệ chủ chốt liên quan đến AI. Ngoài ra, một số đối tượng, chẳng hạn như mục tiêu, các đối tượng liên quan các vấn đề thực thi, triển khai trong AI cũng được đề cập đến và phân loại một cách phù hợp.

Do đó yêu cầu xây dựng các khái niệm về AI đang được triển khai và phân loại theo các vấn đề dưới đây:

* Các loại hình AI (AI mạnh, AI yếu, AI tổng quát, AI hẹp)
* Các tác nhân
* Tri thức
* Nhận thức và điện toán nhận thức
* Điện toán ngữ nghĩa
* Điện toán mềm
* Thuật toán di truyền
* Học máy
* Dữ liệu trong AI
* Học máy
* Mức độ tự động hóa trong AI
* IoT và các hệ thống thực - ảo
* Tính đáng tin cậy trong AI
* Xác minh và thẩm định trong AI
* Pháp lý trong AI
* Tác động xã hội
* Vai trò của các bên liên quan

### Chuẩn hóa các khái niệm mang tính chỉ dẫn trong AI

Một số khái niệm mang tính chất cung cấp các hiểu biết, tri thức hoặc có tính hướng dẫn, đặc thù hóa trong lĩnh vực AI cũng cần phải được xác định, chẳng hạn như đối với các vấn đề dưới đây:

* Vòng đời của các hệ thống AI: Mô tả các đặc thù về các quá trình như khởi đầu, thiết kế phát triển, triển khai, vận hành, đánh giá v.v..
* Các tính năng chủ yếu của hệ thống AI, chẳng hạn như chức năng dự đoán, ra quyết định, hành động v.v...
* Hệ sinh thái AI: mô tả các đối tượng liên quan trong hệ sinh thái, như các phương pháp kỹ thuật, dữ liệu và dữ liệu lớn, hệ thống điện toán đám mây, quản lý tài nguyên v.v..
* Lĩnh vực ứng dụng AI: xử lý ngôn ngữ tự nhiên, khai phá dữ liệu, ứng dụng điển hình trong chống gian lận, xe tự lái, bảo trì theo dự đoán v.v..

## Lựa chọn tài liệu, tiêu chuẩn tham chiếu để xây dựng các yêu cầu kỹ thuật về chuẩn hóa thuật ngữ, khái niệm trong AI

Hoạt động xây dựng tiêu chuẩn hóa trong lĩnh vực trí tuệ nhân tạo tại Việt Nam hiện đang ở trong giai đoạn bắt đầu. Hiện tại chưa có đơn vị, tổ chức nào tại Việt Nam triển khai hoạt động chuẩn hóa về thuật ngữ, định nghĩa về AI. Do vậy việc xây dựng và ban hành tiêu chuẩn thuật ngữ, định nghĩa về AI là cần thiết.

Để có thể xây dựng được tiêu chuẩn thuật ngữ, định nghĩa về AI thì cần phải lựa chọn được tài liệu, tiêu chuẩn tham chiếu để làm sở cứ xây dựng các nội dung của tiêu chuẩn này. Một số tiêu chí dưới đây cần được xem xét trong việc lựa chọn tài liệu, tiêu chuẩn tham chiếu:

* Tiêu chuẩn, tài liệu tham chiếu cần đề cập một cách toàn diện đến các khía cạnh liên quan đến thuật ngữ, định nghĩa về AI như đã giới thiệu ở nội dung trước. Cụ thể là tiêu chuẩn phải đề cập đến các thuật ngữ, định nghĩa các thuật ngữ, các khái niệm; có sự phân loại theo các lĩnh vực, khía cạnh thực hiện của AI
* Tiêu chuẩn, tài liệu làm tài liệu tham chiếu đưa ra các định nghĩa, thuật ngữ dựa trên trên phương pháp tiếp cận tổng quát và toàn diện nhất; không căn cứ vào cách tiếp cận cụ thể về định nghĩa, thuật ngữ từ việc phát triển và xây dựng một loại hình sản phẩm, ứng dụng AI nhất định.
* Tiêu chuẩn, tài liệu tham chiếu đã được ban hành chính thức bởi các tổ chức tiêu chuẩn hóa quốc tế có uy tín; có tài liệu, tiêu chuẩn đã được chọn là tài liệu tham chiếu xây dựng ban hành thành TCVN trong lĩnh vực AI.

Xét theo hiệu trạng xây dựng tiêu chuẩn về tính tin cậy trong AI như đã trình bày ở nội dung trước, căn cứ vào các tiêu chí đề ra ở trên về lựa chọn tài liệu tham chiếu để xây dựng; nhóm xây dựng tiêu chuẩn nhận thấy rằng:

Đối hoạt động tiêu chuẩn hóa của tổ chức tiêu chuẩn IEEE về thuật ngữ, định nghĩa về AI; tổ chức này đã ban hành một số tài liệu, tiêu chuẩn liên quan đến một số khía cạnh thực thi AI. Tuy nhiên tổ chức này chưa ban hành một tài liệu, tiêu chuẩn nào mang tính chuẩn hóa các định nghĩa, thuật ngữ về AI. Có hai tài liệu phục vụ chuẩn hóa của tổ chức này là IEEE 610.8 (Từ vựng về trí tuệ nhân tạo và các ứng dụng người máy) xuất bản năm 1996 và “Thiết kế phù hợp với đạo đức – Từ vựng xuất bản lần thứ nhất” năm 2018 liên quan đến đề xuất chuẩn hóa thuật ngữ, định nghĩa trong AI. Tuy nhiên tài liệu IEEE 610.8 từ năm 1996 tới nay chưa được cập nhật các thay đổi. Tài liệu “Thiết kế phù hợp với đạo đức – Từ vựng xuất bản lần thứ nhất” tập trung vào các định nghĩa thuật ngữ liên quan tới thiết kế, thực thi đạo đức trong AI. Do vậy thuật ngữ đưa ra trong các tài liệu này là không đầy đủ và không toàn diện. Hơn nữa đây là các tài liệu chưa được ban hành thành tiêu chuẩn chính thức.

Liên minh viễn thông quốc tế ITU-T cũng không xúc tiến hoạt động chuẩn hóa riêng về thuật ngữ, định nghĩa trong AI. Một số thuật ngữ, định nghĩa được đưa ra hoặc tham chiếu và bao hàm trong các khuyến nghị, tài liệu được tổ chức này ban hành chỉ liên quan đến các lĩnh vực cụ thể mà chúng đề cập tới. Do vậy việc lựa chọn khuyến nghị hoặc tài liệu tham chiếu để xây dựng tiêu chuẩn định nghĩa, thuật ngữ AI là không đầy đủ.

Cũng giống như vậy, Viện tiêu chuẩn Viễn thông châu Âu (ETSI) thực hiện hoạt động chuẩn hóa về AI tập trung vào vấn đề bảo mật thông qua nhóm nghiên cứu về bảo mật trong AI (SAI) hoặc vẫn đề áp dụng AI trong hạ tầng mạng (ENI). Do vậy các định nghĩa, thuật ngữ cũng chỉ được bao hàm là nội dung theo bố cục tiêu chuẩn hoặc tài liệu và cho lĩnh vực hẹp mà chúng đề cập tới. Do vậy cũng không có tiêu chuẩn nào của ETSI phù hợp để lựa chọn làm tiêu chuẩn tham chiếu cho việc xây dựng tiêu chuẩn quốc gia về định nghĩa, thuật ngữ về AI.

Đối với ISO/IEC, hoạt động chuẩn hóa về AI được thực hiện bởi nhóm nghiên cứu SC42 trực thuộc tiểu ban kỹ thuật phối hợp 1 (ISO/IEC JTC1). SC42 là nhóm nghiên cứu đã tập hợp được rất nhiều các chuyên gia từ các tổ chức tiêu chuẩn quốc tế khác và từ các lĩnh vực công nghiệp, các nhà nghiên cứu phát triển, cung cấp các sản phẩm về AI. Do vậy có thể nói đây là nhóm nghiên cứu hiệu quả nhất trong hoạt động chuẩn hóa về AI. Phần lớn các tiêu chuẩn, tài liệu của các tổ chức tiêu chuẩn khác (đề cập ở trên, chẳng hạn như IEEE. ITU-T, ETSI) đều tham chiếu đến các định nghĩa, thuật ngữ có trong các tiêu chuẩn, tài liệu kỹ thuật do ISO/IEC ban hành.

Trên cơ sở đề ra các tiêu chí về lựa chọn tài liệu tham chiếu, cũng như xem xét tình trạng ban hành, nội dung của các tiêu chuẩn, tài liệu của các tổ chức tiêu chuẩn hóa ở trên; nhóm xây dựng tiêu chuẩn quan tâm đặc biệt tới tiêu chuẩn ISO/IEC 22989:2022: Công nghệ thông tin – Trí tuệ nhân tạo – Các khái niệm và thuật ngữ trí tuệ nhân tạo (Information technology – Artificial intelligence –Artificial intelligence concepts and terminology). Trên cơ sở xem xét các nội dung tiêu chuẩn, đối chiếu với tiêu chí lựa chọn tài liệu, tiêu chuẩn tham chiếu ở trên; nhóm nghiên cứu xây dựng tiêu chuẩn nhận thấy chuẩn này đề cập mang tính toàn diện đến các vấn đề chuẩn hóa thuật ngữ, định nghĩa, khái niệm về AI, cụ thể như sau:

Tiêu chuẩn đã đưa ra tập hợp các thuật ngữ, định nghĩa, khái niệm đầy đủ và được phân loại theo tình lĩnh vực, khía cạnh thực thi của AI. Do đó người dùng có thể hiểu và nhạn biết được các thuật ngữ, định nghĩa, khái niệm thuộc về vấn đề gì và áp dụng nó như thề nào cho phù hợp. Chẳng hạn như các thuật ngữ được phân loại vào hầu hết các lĩnh vực vấn đề tướng ứng liên quan tới AI như là dữ liệu, học máy, mạng nơ-ron, xử lý ngôn ngữ tự nhiên, thị giác máy tính v.v.. Tương tự như vậy, các khái niệm đưa ra trong tiêu chuẩn cũng được phân loại theo lĩnh vực liên quan như AI tổng quát, AI hẹp, học máy, thuật toán, mức độ tự động hóa trong AI, vòng đời AI v.v..

Tiêu chuẩn cũng xác định các khái niệm để làm rõ các loại hình, đối tượng, bên liên quan trong hệ sinh thái AI, các chức năng cơ bản, ứng dụng điện hình của AI trong thực tiễn. Điều này là rất cần thiết để các bên liên quan nhận thức rõ vai trò, phạm vi của họ trong quá trình phát triển, thiết kế, vận hành, quản lý nhà nước các hệ thống, sản phẩm, dịch vụ AI

Tiêu chuẩn đã đưa ra các định nghĩa, xác định các yêu cầu đối với các khía cạnh liên quan đến thực thi tính đáng tin cậy của Ai, chẳng hạn như trách nhiệm, cách nhiẹm giải trình, tính an toàn, các vấn đề rủi ro về bảo mật, vấn đề về sự thiên vị, quyền riêng tư, tính minh bạch, khả năng giải thích v.v.. Ngoài ra, các nội dung trong tiêu chuẩn cũng đã cung cấp các thông tin mang tính hướng dẫn thực thi, tuân thủ các yêu cầu; chiến lược giảm thiểu rủi ro ở mọi phạm vi thực hiện có tác động đến tính đang tin cậy của hệ thống, sản phẩm AI.

Các tiếp cận xây dựng các định nghĩa, thuật ngữ, khái niệm trong tiêu chuẩn mang tính tổng quát và toàn diện. Nó đề cập đến hầu hết các vấn đề chủ yếu trong thực thi AI mà cần phải làm rõ và thống nhất về cách hiểu và sử dụng nó phù hợp. Nó không phải được xây dựng từ một hệ thống, giải pháp, mô hình cụ thể nào mà được lựa chọn bởi các chuyên gia trong lĩnh vực này từ các phương pháp tiếp cận, thực hành tốt nhất đã có.

Tiêu chuẩn ISO/IEC 22989:2022 được soạn thảo từ năm 2019. Tiêu chuẩn này trải qua quá trình sửa đổi, biểu quyết, phê duyệt theo quy trình phê duyệt của ISO/IEC và được thông qua ban hành chính thức thành tiêu chuẩn quốc tế vào tháng 6/2022.

Hiện tại Bộ Khoa học và Công nghệ cũng đã xây dựng một số tiêu chuẩn trong lĩnh vực AI dựa trên các tài liệu tham chiếu chính là các tiêu chuẩn do nhóm SC42 (là nhóm soạn thảo tiêu chuẩn ISO/IEC TR 24028:2020 xây dựng, cụ thể như sau:

* TCVN 13238:2020: ISO/IEC 20546:2019 – Công nghệ thông tin – Dữ liệu lớn – Tổng quan và từ vựng (Information technology – Big data – Overview and vocabulary)
* TCVN 13239-2:2020: ISO/IEC TR 20547-2:2018 – Công nghệ thông tin – Kiến trúc tham chiếu dữ liệu lớn – Phần 2: Các trường hợp sử dụng và yêu cầu dẫn xuất (Information technology – Big data reference architecture – Part 2: Use cases and derived requirements)
* TCVN 13239-5:2020: ISO/IEC TR 20547-5:2018 – Công nghệ thông tin – Kiến trúc tham chiếu dữ liệu lớn – Phần 5: Lộ trình tiêu chuẩn (Information technology – Big data reference architecture – Part 5: Standards roadmap)

Các tiêu chuẩn tham chiếu nói trên có cùng số hiệu phân loại của ISO/IEC với tiêu chuẩn ISO/IEC TR 24028:2020 (35.020). Điều này tạo điều kiện thuận lợi hơn trong việc hướng tới xây dựng và ban hành tiêu chuẩn một cách nhất quán, hài hòa với hệ thống tiêu chuẩn quốc tế cụ thể cũng như công tác phân loại tiêu chuẩn sau này.

Hiện tại Việt Nam chưa ban hành một tiêu chuẩn nào liên quan đến lĩnh vực AI. Trong khi đó các sản phẩm, dịch vụ AI hiện nay đã được các công ty trong và ngoài nước nghiên cứu phát triển, cung cấp, triển khai ở nhiều lĩnh vực. AI là một lĩnh vực đặc thù nó liên quan đến rất nhiều các bên tham gia và được hiểu trong hệ sinh thái các sản phẩm, dịch vụ AI. Do vậy các bên liên quan này cần hiểu rõ vị trí, vai trò, chức năng, phạm vi công việc mà họ thực hiện hoặc liên quan cũng như các ảnh hưởng, tác động tới họ hoặc tới các đối tượng khác. Các vấn đề này được giải quyết bằng việc tiêu chuẩn được xây dựng giúp các đối tượng áp dụng để hiểu rõ, có những nhận thức chung thông qua các thuật ngữ, khái niệm trong các bối cảnh, trường hợp, vai trò cụ thể mà các bên tham gia.

Như vậy việc xây dựng tiêu chuẩn thuật ngữ và khái niệm trong AI là một yêu cầu cần thiết hiện nay để đáp ứng nhu cầu triển khai, sử dụng và quản lý các sản phẩm, dịch vụ và hệ thống AI trong nước. Tiêu chuẩn ISO/IEC 22989:2022 đáp ứng đầy đủ các tiêu chí cho việc lựa chọn tài liệu, tiêu chuẩn tham chiếu để xây dựng tiêu chuẩn quốc gia về thuật ngữ, khái niệm trong AI.

Với các lý do như trên, nhóm nghiên cứu đã chọn tài liệu **tiêu chuẩn ISO/IEC 22989:2022** làm tiêu chuẩn tham chiếu để xây dựng dự thảo tiêu chuẩn quốc gia về định nghĩa, thuật ngữ về AI.

# Giải thích nội dung dự thảo tiêu chuẩn

## Xây dựng dự thảo tiêu chuẩn

Dự thảo được xây dựng theo hình thức chấp thuận nguyên vẹn các nội dung theo tài liệu tham chiếu được chọn và phù hợp với các tiêu chí để thực hiện mục tiêu xây dựng TCVN. Cụ thể như sau:

* Biên dịch các nội dung của tiêu chuẩn tham chiếu
* Thực hiện so sánh đối chiếu các định nghĩa, thuật ngữ bằng tiếng Anh trong tài liệu tham chiếu và các định nghĩa thuật ngữ đã dùng tại các tiêu chuẩn phù hợp được ban hành trong nước để có sự đồng nhất về các định nghĩa, thuật ngữ dịch ra tiếng Việt
* Trình bày thể thức, định dạng dự thảo tiêu chuẩn quy định trong TCVN 1-2: 2008.

## Các nội dung chính của dự thảo tiêu chuẩn

***1 Phạm vi áp dụng***

Phạm vi của tiêu chuẩn là chuẩn hóa các thuật ngữ, khái niệm trong AI, áp dụng cho các bên liên quan trao đổi thông tin để có sự hiểu, nhận thức rõ ràng, nhất quán trong các hoạt động liên quan đến thiết kế, phát triển, vận hành, sử dụng và quản lý các hệ thống, sản phẩm, dịch vụ AI.

Đối tượng sử dụng tiêu chuẩn là các cơ quan quản lý nhà nước, doanh nghiệp, tổ chức và cá nhân.

***2. Tài liệu viện dẫn***

Các tài liệu tham chiếu cho các nội dung của dự thảo tiêu chuẩn

***3. Các thuật ngữ và định nghĩa***

Nội dung này đưa ra thuật ngữ và phát biểu các định nghĩa để hiểu rõ về bản chất các thuật ngữ được sử dụng trong lĩnh vực AI

Các thuật ngữ và cách hiểu thông qua các định nghĩa giúp cho các bên liên quan như bên thiết kế phát triển, vận hành hiểu rõ các công nghệ, kỹ thuật, mô hình, thuật toán chính trong AI để có sự hợp tác trao đổi thông tin rõ ràng và thông suốt trong quá trình hợp tác. Các bên liên quan đến cung cấp và sử dụng dữ liệu trong AI có thể áp dụng để hiểu, phân loại các loại hình dữ liệu được sử dụng trong vòng đời của hệ thống, sản phẩm, dịch vụ AI.

Các thuật ngữ và định nghĩa trong nội dung này giúp các cơ quan quản lý có những hiểu biết và nhận thức rõ và phù hợp với các nhà cung cấp đặc tính tổng quát về AI, các công đoạn của quá trình sản xuất, vận hành đối với một hệ thống, sản phẩm, dịch vụ AI để từ đó có những hoạt động quản lý phù hợp cho mỗi một giai đoạn của vòng đời sản phẩm AI.

Các thuật ngữ, định nghĩa được chuẩn hóa trong tiêu chuẩn sẽ được sử dụng trong các tiêu chuẩn được ban hành sau này.

Theo đó các nội dung thuật ngữ, định nghĩa trong phần này bao gồm:

3.1 Các thuật ngữ liên quan đến AI

3.2 Các thuật ngữ liên quan đến dữ liệu

3.3 Các thuật ngữ liên quan đến học máy

3.4 Các thuật ngữ liên quan đến mạng nơ-ron

3.5 Các thuật ngữ liên quan đến tính đáng tin cậy

3.6 Các thuật ngữ liên quan đến xử lý ngôn ngữ tự nhiên

3.7 Các thuật ngữ liên quan đến thị giác máy tính

***4 Thuật ngữ viết tắt***

Nội dung này nêu các thuật ngữ viết tắt bằng tiếng Anh, tên đầy đủ của thuật ngữ bằng tiếng Anh và tiếng Việt

Các thuật ngữ viết tắt này được sử dụng trong tiêu chuẩn này và các tiêu chuẩn được ban hành sau này.

***5 Các khái niệm AI***

Một số đối tượng trong AI tương đối phức tạp, do vậy nó không thể bao hàm được trong các thuật ngữ và định nghĩa một cách ngắn gọn về nó như ở nội dung 3 của dự thảo tiêu chuẩn. Do đó các đối tượng về thuật ngữ, định nghĩa như nêu ra ở mục 3 của dự thảo tiêu chuẩn cần được diễn giải để hiểu rõ ở dạng các khái niệm. Nó mô tả chi tiết các thuật ng về thuộc tính, đặc điểm, các dạng thể hiện, phân loại về hình thái, phương pháp, cách thức sử dụng, các mối liên hệ, tương tác với các đối tượng hoặc quá trình khác.

Các khái niệm nêu trong nội dung này cho phép các đối tượng sử dụng tiêu chuẩn (như đã chỉ ra trong phạm vi áp dụng của tiêu chuẩn) hiểu biết sâu sắc và ở phạm vi rộng hơn về các đối tượng, quá trình, kỹ thuật công nghệ, các tác động, ảnh hưởng để từ đó các bên liên quan xác định được vị trí vai trò của mình trong hoạt động thiết kế, phát triển, vận hành và quản lý nhà nước đối với các hệ thống, sản phẩm, dịch vụ AI.

Theo đó các nội dung phần này đưa ra các khái niệm bao gồm:

5.1 Tổng quan

5.2 Từ AI mạnh và yếu đến AI tổng quát và hẹp

5.3 Tác nhân

5.4 Tri thức

5.5 Nhận thức và điện toán nhận thức

5.6 Điện toán ngữ nghĩa

5.7 Điện toán mềm

5.8 Thuật toán di truyền

5.9 Các phương pháp tiếp cận biểu trưng và biểu trưng phụ cho AI

5.10 Dữ liệu

5.11 Các khái niệm về học máy

5.12 Ví dụ về thuật toán học máy

5.13 Tự chủ, can thiệp và tự động hóa

5.14 Internet vạn vật và các hệ thống thực - ảo

5.15 Tính đáng tin cậy

5.16 Xác minh và thẩm định trong AI

5.17 Các vấn đề pháp lý

5.18 Tác động xã hội

5.19 Vai trò của các bên liên quan đến AI

***6 Vòng đời hệ thống AI***

Nội dung trong mục có thể được coi là sự mở rộng các mô tả yêu cầu vòng đời hệ thống phần mềm đưa ra trong các tiêu chuẩn đã được ban hành như TCVN 10539:2014 và TCVN 10607- 4:2014, trong đó có tính đến các đặc thù đối với hệ thống, sản phẩm, dịch vụ AI.

Các yêu cầu, mô tả, các hoạt động cần được triển khai có trong nội dung này là hướng dẫn cho các đối tác liên quan đến quá trình phát triển, triển khai, vận hành và kết thúc các sản phẩm thuộc lĩnh vực AI. Các nội dung, yêu cầu cho hoạt động đánh giá, xác minh, thẩm định là cần thiết để hướng dẫn triển khai quản lý nhà nước về hoạt động kiểm định, chứng nhận sản phẩm, dịch vụ AI.

Theo đó các nội dung phần này mô tả các yêu cầu về vòng đời hệ thống, sản phẩm, dịch vụ AI bao gồm:

6.1 Mô hình vòng đời hệ thống AI

6.2 Các giai đoạn và quá trình trong vòng đời của hệ thống AI

***7 Tổng quan chức năng của hệ thống AI***

Nội dung mô tả các chức năng chính mà bất cứ một hệ thống Ai nào cũng cần phải có. Thông qua các mô tả chức năng, các đối tượng sử dụng tiêu chuẩn nắm rõ về mô hình chức năng và các thành phần cơ bản của một hệ thống AI. Các thể hiện hoạt động của AI, chẳng hạn như hoạt động dự đoán, ra quyết định và hành động được mô tả để các đối tượng sử dụng tiêu chuẩn, đặc biệt là cá nhân, tổ chức sử dụng và quản lý sản phẩm dịch vụ AI hiểu rõ về bản chất, cách thức hoạt động của chúng để từ đó sử dụng đúng và có cách tiếp cận đề ra các quy định quản lý phù hợp.

Theo đó phần này bao gồm các nội dung:

7.1 Yêu cầu chung

7.2 Dữ liệu và thông tin

7.3 Tri thức và học tập

7.4 Từ dự đoán đến hành động

***8 Hệ sinh thái AI***

Nội dung này của tiêu chuẩn mô tả mô hình hệ sinh thái AI được thể hiện bởi các thành phâm thực thi của chúng. Các thành phần này có sự gắn kết với nhau và trải rộng ở rất nhiều các lĩnh vực, chẳng hạn như cơ sở hạ tầng, kỹ thuật, công nghệ cả ở trong và ngoài lĩnh vực AI (dữ liệu lớn, điện toán đám mây, hệ tri thức chuyên gia ở nhiều lĩnh vực kinh tế, xã hội khác nhau.)

Nội dung này giúp các đối tượng áp dụng tiêu chuẩn hiểu rõ mô hình hệ sinh thái AI, các mối liên quan giữa các thực thể bên trong cũng như các chỉ dẫn về cách thức thực hiện các thành phần của hệ sinh thái trong thực tế.

Theo đó phần này có các nội dung dưới đây:

8.1 Yêu cầu chung

8.2 Hệ thống AI

8.3 Chức năng AI

8.4 Học máy

8.5 Kỹ thuật

8.6 Dữ liệu lớn và nguồn dữ liệu - điện toán đám mây và điện toán biên

8.7 Vùng tài nguyên

***9 Lĩnh vực của AI***

Các hệ thống, sản phẩm, dịch vụ AI được xây dựng dựa trên các kỹ thuật, công nghệ chính của nó là học máy, lô-gic mờ, mạng nơ-tron. Trong thực tế triển khai đã có rất nhiều sự nhầm lẫn giữa công nghệ và lĩnh vực áp dụng của AI. Do đó nội dung này đề ra sự phân biệt, xác định một số lĩnh vực ứng dụng triển khai chủ yếu của AI như mô tả ở các nội dung dưới đây:

9.1 Thị giác máy tính và nhận dạng hình ảnh

9.2 Xử lý ngôn ngữ tự nhiên

9.3 Khai phá dữ liệu

9.4 Lập kế hoạch

Từ các nội dung chính này có thể phát triển các ứng dụng cụ thể phù hợp với từng ngành nghề cụ thể.

***10 Ứng dụng của hệ thống AI***

Nội dung đưa ra các tham khảo về ứng dụng của các hệ thống AI đã được triển khai (đã thành công, có tiềm năng hoặc các thực hành tốt nhất đã được khảo sát) tại thời điểm hiện tại cũng như trong tương lai.

Các chỉ dẫn này giúp cho các nhà cung cấp sản phẩm AI, các cơ quan quản lý tham khảo để xem xét, đánh giá, định hướng phát triển các sản phẩm AI, tính khả thi và phù hợp với thực tế trong nước.

Do đó nội dung này bao gồm các phần dưới đây:

10.1 Yêu cầu chung

10.2 Phát hiện gian lận

10.3 Xe tự động

10.4 Bảo trì theo dự đoán

***Phụ lục***

Mô tả sự phù hợp về mô hình vòng đời hệ thống AI đưa ra trong tiêu chuẩn này với định nghĩa của OECD

Bảng 1 – Bảng đối chiếu nội dung dự thảo tiêu chuẩn với tài liệu tiêu chuẩn tham chiếu ISO/IEC 22989:2022

|  |  |  |
| --- | --- | --- |
| **Dự thảo tiêu chuẩn xxxxx:2022: Công nghệ thông tin – Trí tuệ nhân tạo – Các khái niệm và thuật ngữ trí tuệ nhân tạo** | **Tài liệu tham chiếu**  **(ISO/IEC 22989:2022)** | **Sửa đổi, bổ sung** |
| **1 Phạm vi áp dụng** | **1 Scope** | Chấp thuận nguyên vẹn |
| **2 Tài liệu viện dẫn** | **2 Normative references** | Chấp thuận nguyên vẹn |
| **3 Các thuật ngữ và định nghĩa** | **3 Terms and definitions** | Chấp thuận nguyên vẹn |
| 3.1 Các thuật ngữ liên quan đến AI | 3.1 Terms related to AI | Chấp thuận nguyên vẹn |
| 3.2 Các thuật ngữ liên quan đến dữ liệu | 3.2 Terms related to data | Chấp thuận nguyên vẹn |
| 3.3 Các thuật ngữ liên quan đến học máy | 3.3 Terms related to machine learning | Chấp thuận nguyên vẹn |
| 3.4 Các thuật ngữ liên quan đến mạng nơ-ron | 3.4 Terms related to neuron networks | Chấp thuận nguyên vẹn |
| 3.5 Các thuật ngữ liên quan đến tính đáng tin cậy | 3.5 Terms related to trustworthiness | Chấp thuận nguyên vẹn |
| 3.6 Các thuật ngữ liên quan đến xử lý ngôn ngữ tự nhiên | 3.6 Terms related to natural language processing | Chấp thuận nguyên vẹn |
| 3.7 Các thuật ngữ liên quan đến thị giác máy tính | 3.7 Terms related to computer vision | Chấp thuận nguyên vẹn |
| **4 Thuật ngữ viết tắt** | **4 Abbreviated** | **Chấp thuận nguyên vẹn** |
| **5 Các khái niệm AI** | **5 AI concepts** |  |
| 5.1 Tổng quan | 5.1 General | Chấp thuận nguyên vẹn |
| 5.2 Từ AI mạnh và yếu đến AI tổng quát và hẹp | 5.2 From strong and weak AI to general and narrow AI | Chấp thuận nguyên vẹn |
| 5.3 Tác nhân | 5.3 Agent | Chấp thuận nguyên vẹn |
| 5.4 Tri thức | 5.4 Knowledge | Chấp thuận nguyên vẹn |
| 5.5 Nhận thức và điện toán nhận thức | 5.5 Cognition and cognitive computing | Chấp thuận nguyên vẹn |
| 5.6 Điện toán ngữ nghĩa | 5.6 Semantic computing | Chấp thuận nguyên vẹn |
| 5.7 Điện toán mềm | 5.7 Soft computing | Chấp thuận nguyên vẹn |
| 5.8 Thuật toán di truyền | 5.8 Genetic algorithms | Chấp thuận nguyên vẹn |
| 5.9 Các phương pháp tiếp cận biểu tượng và biểu tượng phụ cho AI | 5.9 Symbolic and subsymbolic approaches for AI | Chấp thuận nguyên vẹn |
| 5.10 Dữ liệu | 5.10 Data | Chấp thuận nguyên vẹn |
| 5.11 Các khái niệm về học máy | 5.11 Machine learning concepts | Chấp thuận nguyên vẹn |
| 5.12 Ví dụ về thuật toán học máy | 5.12 Examples of machine learning algorithms | Chấp thuận nguyên vẹn |
| 5. 13 Tự chủ, can thiệp và tự động hóa | 5.13 Autonomy, heteronomy and automation | Chấp thuận nguyên vẹn |
| 5.14 Internet vạn vật và các hệ thống thực - ảo | 5.14 Internet of things and cyber-physical systems | Chấp thuận nguyên vẹn |
| 5.15 Tính đáng tin cậy | 5.15 Trustworthiness | Chấp thuận nguyên vẹn |
| 5.16 Xác minh và thẩm định trong AI | 5.16 AI verification and validation | Chấp thuận nguyên vẹn |
| 5.17 Các vấn đề pháp lý | 5.17 Jurisdictional issues | Chấp thuận nguyên vẹn |
| 5.18 Tác động xã hội | 5.18 Societal impact | Chấp thuận nguyên vẹn |
| 5.19 Vai trò của các bên liên quan đến AI | 5.19 AI stakeholder roles | Chấp thuận nguyên vẹn |
| **6 Vòng đời hệ thống AI** | **6 AI system life cycle** |  |
| 6.1 Mô hình vòng đời hệ thống AI | 6.1 AI system life cycle model | Chấp thuận nguyên vẹn |
| 6.2 Các giai đoạn và quá trình trong vòng đời của hệ thống AI | 6.2 AI system life cycle stages and processes | Chấp thuận nguyên vẹn |
| **7 Tổng quan về chức năng của hệ thống AI** | **7 AI system functional overview** |  |
| 7.1 Yêu cầu chung | 7.1 General | Chấp thuận nguyên vẹn |
| 7.2 Dữ liệu và thông tin | 7.2 Data and information | Chấp thuận nguyên vẹn |
| 7.3 Tri thức và học tập | 7.3 Knowledge and learning | Chấp thuận nguyên vẹn |
| 7.4 Từ dự đoán đến hành động | 7.4 From predictions to actions | Chấp thuận nguyên vẹn |
| **8 Hệ sinh thái AI** | **8 AI ecosystem** |  |
| 8.1 Yêu cầu chung | 8.1 General | Chấp thuận nguyên vẹn |
| 8.2 Hệ thống AI | 8.2 AI systems | Chấp thuận nguyên vẹn |
| 8.3 Chức năng AI | 8.3 AI function | Chấp thuận nguyên vẹn |
| 8.4 Học máy | 8.4 Machine learning | Chấp thuận nguyên vẹn |
| 8.5 Kỹ thuật | 8.5 Engineering | Chấp thuận nguyên vẹn |
| 8.6 Dữ liệu lớn và nguồn dữ liệu – điện toán đám mây và điện toán biên | 8.6 Big data and data sources – cloud and edge computing | Chấp thuận nguyên vẹn |
| 8.7 Vùng tài nguyên | 8.7 Resource pools | Chấp thuận nguyên vẹn |
| **9 Các lĩnh vực của AI** | **9 Fields of AI** |  |
| 9.1 Thị giác máy tính và nhận dạng hình ảnh | 9.1 Computer vision and image recognition | Chấp thuận nguyên vẹn |
| 9.2 Xử lý ngôn ngữ tự nhiên | 9.2 Natural language processing | Chấp thuận nguyên vẹn |
| 9.3 Khai phá dữ liệu | 9.3 Data mining | Chấp thuận nguyên vẹn |
| 9.4 Lập kế hoạch | 9.4 Planning | Chấp thuận nguyên vẹn |
| **10 Các ứng dụng của hệ thống AI** | **10 Applications of AI systems** |  |
| 10.1 Tổng quan | 10.1 General | Chấp thuận nguyên vẹn |
| 10.2 Phát hiện gian lận | 10.2 Fraud detection | Chấp thuận nguyên vẹn |
| 10.3 Xe tự động | 10.3 Automated vehicles | Chấp thuận nguyên vẹn |
| 10.4 Bảo trì theo dự đoán | 10.4 Predictive maintenance | Chấp thuận nguyên vẹn |
| **Phụ lục A (tham khảo)** | **Annex A (informative)** | Chấp thuận nguyên vẹn |

# Khuyến nghị áp dụng tiêu chuẩn

Một số khuyến nghị dưới đây về áp dụng tiêu chuẩn trong thực tế:

***Mục tiêu chuẩn hóa:***

Tiêu chuẩn đề cập đến phạm vi chuẩn hóa các thuật ngữ và mô tả các khái niệm trong lĩnh vực AI. Do vậy các nội dung có trong tiêu chuẩn này đáp ứng mục tiêu thông tin, thông hiểu; phổ biến và tăng cường hiểu biết về các quá trình, kỹ thuật, công nghệ; các vấn đề liên quan đến hệ thống, sản phẩm, dịch vụ trong lĩnh vực AI.

***Đối với các đối tượng áp dụng tiêu chuẩn:***

Các đối tượng áp dụng tiêu chuẩn này là các cá nhân, tổ chức, doanh nghiệp, cơ quan quản lý nhà nước hoạt động trong mọi lĩnh vực liên quan tới AI như nghiên cứu kỹ thuật công nghệ, thiết kế, phát triển, sản xuất, triển khai và quản lý nhà nước đối với các hệ thống, sản phẩm, dịch vụ dựa trên công nghệ AI (hoặc một phần dựa trên công nghệ AI). Mỗi một đối tượng áp dụng tiêu chuẩn có thể tham gia vào một hoặc nhiều khâu trong toàn bộ quá trình vòng đời của một hệ thống, sản phẩm AI. Do đó các đối tượng áp dụng cần phải có những nhận thức, hiểu biết rõ ràng và nhất quán về các thuật ngữ, định nghĩa, khái niệm liên quan; nhất là các khía cạnh liên quan đến vị trí, vai trò của mình để từ đó phân loại, xác định các loại hình hoạt động cụ thể cũng như môi tương tác, ảnh hưởng qua lại với các đối tượng khác trong toàn bộ vòng đời của một hệ thống dịch vụ, sản phẩm AI. Tác động của tiêu chuẩn đối với các lớp đối tượng cụ thể là:

* *Đối tượng tham gia thiết kế, phát triển triển khai các sản phẩm, dịch vụ AI:* Để tạo ra các sản phẩm, dịch vụ, hệ thống AI các đối tác tham gia cần phải có những hiểu biết, nhận thức chung, nhất quán và rõ ràng về các thuật ngữ, khái niệm, chức năng, kỹ thuật, mô hình từ đó thực hiện các nội dung, công đoạn, quy trình công việc theo đúng yêu cầu của các đối tác khác đưa ra, cũng như yêu cầu cho các đối tác khác thực hiện.
* *Các đối tượng vận hành, sử dụng hệ thống, sản phẩm, dịch vụ AI:* Tiêu chuẩn thuật ngữ, khái niệm AI giúp các bên vận hành, sử dụng hiểu rõ tính năng, đặc tính kỹ thuật, các chức năng, dịch vụ được cung cấp bởi hệ thống AI; hiểu rõ các khâu, công đoạn cần tiến hành trong vận hành và sử dụng; giới hạn phạm vi và cách thức vận hành, sử dụng cũng như vai trò, nghĩa vụ của người vận hành, sử dụng đối với các sản phẩm, dịch vụ về AI
* *Các đối tượng tham gia quản lý:* Tiêu chuẩn thuật ngữ, khái niệm này cung cấp cho các cơ quan quản lý hiểu rõ những vấn đề liên quan đến chức năng quản lý nhà nước về sản phẩm trong lĩnh vực AI. Nó cũng giúp các nhà quản lý nhận thức các tác động tích cực cũng như tiêu cực đến đến các lĩnh vực kinh tế, xã hội để từ đó áp dụng các quy định hiện hành phù hợp cũng như nhận biết các vấn đề mới cần đưa ra trong thực thi tuân thủ liên quan đến hoạt động quản lý nhà nước ở từng lĩnh vực triển khai cụ thể, chẳng hạn hoạt động như xác minh, thẩm định, chứng nhận đối với các hệ thống, sản phẩm, dịch vụ AI.

***Đối với hoạt động xây dựng các tiêu chuẩn trong lĩnh vực AI:***

* Các thuật ngữ, khái niệm này làm tiền đề để chuẩn hóa các thuật ngữ, khái niệm, được hiểu một cách nhất quán và được sử dụng trong các tiêu chuẩn, quy chuẩn về AI được xây dựng và ban hành sau này.
* Tiêu chuẩn áp dụng tài liệu tham chiếu là tiêu chuẩn quốc tế của tổ chức ISO/IEC theo hình thức chấp thuận nguyên vẹn. Do vậy tiêu chuẩn đáp ứng sự hài hòa đối với hệ thống tiêu chuẩn quốc tế, góp phần thúc đẩy chủ trương phát triển hội nhập trong giai đoạn hiện nay.

# Kiến nghị về hướng phát triển tiếp theo

Dự thảo tiêu chuẩn TCVN xxxxx:2022 “Công nghệ thông tin – Trí tuệ nhân tạo – Các khái niệm và thuật ngữ trí tuệ nhân tạo” dự kiến được ban hành thành tiêu chuẩn quốc gia sẽ là tiêu chuẩn đầu tiên trong lĩnh vực AI được ban hành trong nước. Các thuật ngữ, định nghĩa, khái niệm và các nội dung dung đưa ra trong tiêu chuẩn đề cập hoặc chỉ dẫn đến nhiều khía cạnh chuẩn hóa khác về AI. Điều này hoàn toàn phù hợp với xu hướng thúc đẩy các hoạt động chuẩn hóa trong lĩnh vực AI của các tổ chức tiêu chuẩn quốc tế nói chung và các quốc gia, khu vực nói riêng. Dự kiến trong thời gian tới cần tiếp tục triển khai hoạt động chuẩn hóa AI trong nước, trọng tâm hướng tới các vấn đề dưới đây:

* Tính thiên vị và ra quyết định được trợ giúp bởi AI
* Khung cho hệ thống AI sử dụng học máy
* Tổng quan về tính đáng tin cậy trong AI
* Tổng quan về đạo đức và mối quan tâm về xã hội đối với AI
* Tổng quan các phương pháp tính toán đối với hệ thống AI
* Quản trị trí tuệ nhân tạo – Các chỉ dẫn về sử dụng trí tuệ nhân tạo trong các tổ chức.

Ngoài ra cần tổ chức thực hiện các nhiệm vụ nghiên cứu tìm hiểu, cập nhật xu hướng phát triển công nghệ, thị trường, sản phẩm; xu hướng chuẩn hóa trong lĩnh vực AI để đưa ra các kết quả phục vụ mục tiêu thúc đẩy phát triển ứng dụng các sản phẩm dịch vụ AI trong nước.

------------------------------------------------------