**Hoàn thiện Luật Công nghiệp công nghệ số: Bài học từ "Sóng thần công nghệ"**

Dự thảo Luật Công nghiệp công nghệ số (CNCNS) đang trong quá trình lấy ý kiến đóng góp rộng rãi, cho thấy nỗ lực kịp thời của Việt Nam trong việc đón đầu và định hình khung pháp lý cho một lĩnh vực then chốt của nền kinh tế trong kỷ nguyên số.

**Bối cảnh**

Sự ra đời của dự luật này không chỉ là xu hướng tất yếu của thời đại mà còn là yêu cầu cấp bách, được thúc đẩy bởi chính những đặc điểm mang tính bước ngoặt của làn sóng công nghệ mới, mà trọng tâm là sự kết hợp giữa trí tuệ nhân tạo (AI) và sinh học bên cạnh các công nghệ tiên tiến khác như dữ liệu lớn, điện toán đám mây (ĐTĐM), Internet vạn vật (IoT), chuỗi khối, thực tại ảo/thực tại tăng cường.

Trong cuốn sách "Sóng thần công nghệ" [1], tác giả Mustafa Suleyman đã vẽ nên bức tranh toàn cảnh về làn sóng công nghệ này, nhấn mạnh tính “siêu tiến hóa”, “bất đối xứng”, “đa dụng” và “tự chủ” của nó. Ông chỉ ra rằng, khác với những làn sóng công nghệ trước đây, làn sóng này mang đến tiềm năng to lớn chưa từng có trong việc thay đổi mọi mặt đời sống con người, từ y tế, giáo dục, sản xuất đến quốc phòng, an ninh.

Tuy nhiên, đi kèm với những hứa hẹn về một tương lai tươi sáng là những nguy cơ tiềm tàng khổng lồ, đặt ra "thế lưỡng nan của kiềm tỏa": một mặt, nhân loại cần khai thác triệt để tiềm năng của công nghệ để giải quyết những thách thức toàn cầu như biến đổi khí hậu, dịch bệnh, bất bình đẳng; mặt khác, việc kiểm soát và điều tiết công nghệ là yêu cầu sống còn để ngăn chặn những thảm họa tiềm ẩn như tấn công mạng tự động hóa, deepfake, vũ khí tự động, sự phân tán quyền lực và nguy cơ mất kiểm soát đối với AI.

Dựa trên những luận điểm chính trong cuốn sách "Sóng thần công nghệ", kết hợp với quan điểm của các nhà hoạch định chính sách công nghệ khác, bài viết phân tích lý do vì sao Dự luật CNCNS là cần thiết, đồng thời góp ý cho Dự thảo, nhằm hoàn thiện khung pháp lý, tạo điều kiện cho việc phát triển và ứng dụng công nghệ số tại Việt Nam diễn ra một cách có trách nhiệm, an toàn, bền vững, mang lại lợi ích tối đa cho người dân, đồng thời hạn chế tối thiểu những tác động tiêu cực tiềm ẩn của làn sóng công nghệ mới.

**Bổ sung định nghĩa về AI trong dự thảo Luật**

Một trong những điểm thiếu cần được khắc phục trong Dự thảo Luật CNCNS là việc thiếu định nghĩa rõ ràng về trí tuệ nhân tạo (AI). Điều 3 của Dự thảo hiện tại chỉ đề cập đến “Hệ thống trí tuệ nhân tạo” mà chưa định nghĩa chính xác AI là gì. Đây là một điểm thiếu quan trọng, bởi AI là một khái niệm rộng lớn và đang phát triển nhanh chóng, bao gồm nhiều kỹ thuật, phương pháp và ứng dụng khác nhau.

Thiếu định nghĩa rõ ràng về AI sẽ gây khó khăn cho việc xây dựng các quy định và chính sách liên quan đến AI sau này. Cụ thể, nó sẽ tạo ra sự mơ hồ về phạm vi điều chỉnh của Luật, dẫn đến khó khăn trong việc xác định các hành vi vi phạm, áp dụng chế tài, và đảm bảo tính hiệu lực của Luật.

Trong cuốn sách "Sóng thần công nghệ", tác giả Mustafa Suleyman đã phân tích sâu sắc sự bùng nổ và tầm ảnh hưởng ngày càng tăng của AI trong mọi mặt đời sống. Ông dự đoán rằng AI sẽ là "công cụ khuếch đại sức mạnh lớn nhất trong lịch sử", thúc đẩy tái phân bổ quyền lực ở quy mô chưa từng có và mang đến những thay đổi sâu rộng cho xã hội.

Chính vì vậy, việc định nghĩa AI một cách chính xác và đầy đủ là điều kiện tiên quyết để "kiềm tỏa" làn sóng công nghệ, đảm bảo cho sự phát triển và ứng dụng AI diễn ra một cách an toàn, có kiểm soát và có lợi cho con người.

Để khắc phục thiếu sót này, Dự thảo Luật CNCNS cần bổ sung định nghĩa về AI. Định nghĩa này cần đảm bảo tính bao quát, phản ánh đúng bản chất của AI là công nghệ dựa trên học máy, có khả năng tự học hỏi, thích nghi và đưa ra quyết định độc lập. Đồng thời, định nghĩa cũng cần phù hợp với bối cảnh Việt Nam, phản ánh đúng trình độ phát triển công nghệ và đặc thù văn hóa, xã hội của đất nước.

Ví dụ, Dự thảo Luật có thể tham khảo định nghĩa AI của Liên minh châu Âu: "Trí tuệ nhân tạo (AI) là một tập hợp các khoa học, lý thuyết và kỹ thuật với mục đích là tái tạo bằng máy các khả năng nhận thức của con người”. [2]

Việc bổ sung định nghĩa về AI trong Điều 3 của Dự thảo Luật CNCNS sẽ là bước đi quan trọng, tạo nền tảng pháp lý vững chắc cho việc xây dựng các quy định và chính sách liên quan đến AI trong tương lai, góp phần "kiềm tỏa" làn sóng công nghệ, hướng đến một tương lai AI an toàn, bền vững và có lợi cho con người.

**Hoàn thiện các quy định về đạo đức AI**

Dự thảo Luật CNCNS đã đề cập đến đạo đức AI trong Điều 82, thể hiện sự quan tâm của Nhà nước đối với vấn đề này. Tuy nhiên, những quy định hiện tại còn khá sơ sài, chưa đủ mạnh mẽ để đảm bảo cho việc phát triển và ứng dụng AI diễn ra một cách có trách nhiệm và có đạo đức.

Cuốn sách "Sóng thần công nghệ" đã chỉ ra nhiều nguy cơ tiềm ẩn do AI gây ra nếu không có những quy định đạo đức rõ ràng. Ví dụ, AI có thể bị lợi dụng để xâm phạm quyền riêng tư, phân biệt đối xử, thao túng thông tin, thậm chí gây ra những thảm họa khó lường.

Yoshua Bengio, một nhà khoa học máy tính nổi tiếng thế giới trong một bài gần đây [3] đã cảnh báo về những mối nguy hiểm tiềm ẩn của AI không được kiểm soát. Ông lập luận rằng chúng ta không nên xem nhẹ khả năng AI phát triển các mục tiêu không phù hợp với các giá trị của con người, dẫn đến những hậu quả không mong muốn, thậm chí là thảm khốc.

Để khắc phục hạn chế này, Dự thảo Luật CNCNS cần hoàn thiện các quy định về đạo đức AI theo những hướng cụ thể. Trước hết, Điều 82 cần được bổ sung, chi tiết hóa các nguyên tắc đạo đức AI. Cụ thể, Luật cần làm rõ những nguyên tắc như công bằng, minh bạch, trách nhiệm, an toàn và bảo mật trong phát triển và ứng dụng AI.

AI cần được phát triển và ứng dụng một cách công bằng, không phân biệt đối xử dựa trên chủng tộc, giới tính, tôn giáo, nguồn gốc quốc gia, hoặc bất kỳ yếu tố nào khác.

Quá trình ra quyết định của AI cần được minh bạch và dễ hiểu, cho phép con người có thể kiểm tra, giám sát và hiểu được cách thức AI hoạt động. Các chủ thể liên quan đến AI (nhà phát triển, người sử dụng, tổ chức quản lý) cần chịu trách nhiệm về các tác động của AI, bao gồm cả những tác động tiêu cực ngoài ý muốn. AI cần được phát triển và ứng dụng một cách an toàn, không gây hại cho con người, môi trường và xã hội. Dữ liệu sử dụng cho AI cần được bảo mật và bảo vệ quyền riêng tư của người dùng.

Bên cạnh việc chi tiết hóa các nguyên tắc đạo đức, Dự thảo Luật cũng cần xây dựng cơ chế giám sát việc tuân thủ các nguyên tắc này. Một trong những đề xuất khả thi là **thành lập Hội đồng đạo đức AI độc lập**, bao gồm các chuyên gia từ nhiều lĩnh vực (công nghệ, luật, đạo đức, xã hội học,...).

Hội đồng này có nhiệm vụ đánh giá các hệ thống AI, đưa ra khuyến nghị và giám sát việc tuân thủ các nguyên tắc đạo đức. Đồng thời, Luật cũng có thể yêu cầu các nhà phát triển AI phải báo cáo về các khía cạnh đạo đức của hệ thống AI, bao gồm cả những rủi ro tiềm ẩn và biện pháp giảm thiểu rủi ro.

Cuối cùng, việc xây dựng văn hóa AI có trách nhiệm cũng là yếu tố quan trọng. Luật cần khuyến khích các hoạt động giáo dục, nâng cao nhận thức về đạo đức AI cho cộng đồng, bao gồm nhà phát triển, người sử dụng và xã hội. Bên cạnh đó, việc xây dựng các bộ quy tắc ứng xử, hướng dẫn thực hành tốt nhất về đạo đức AI cho các tổ chức, DN và cá nhân hoạt động trong lĩnh vực AI cũng cần được chú trọng.

**Rủi ro AI**

Dự thảo Luật CNCNS đã đề cập đến rủi ro AI trong Điều 83 và 84, tuy nhiên, những quy định hiện tại chưa đầy đủ để bao quát hết những nguy cơ tiềm ẩn do AI gây ra. Cuốn sách "Sóng thần công nghệ" đã phân tích chi tiết những yếu tố khuếch đại bất ổn định do AI mang lại, bao gồm tấn công mạng tự động hóa, deepfake, robot vũ trang, và tự động hóa việc làm.

AI có thể được sử dụng để tạo ra những cuộc tấn công mạng tinh vi, tự động hóa và khó lường, gây thiệt hại nghiêm trọng cho hệ thống thông tin, cơ sở hạ tầng quan trọng và an ninh quốc gia. Deepfake, với khả năng tạo ra những video, hình ảnh, âm thanh giả mạo với độ chân thực cao, có thể gây khó khăn cho việc phân biệt thật giả, làm xói mòn niềm tin vào thông tin và gây bất ổn xã hội. Robot vũ trang tự động, có khả năng giết người mà không cần sự can thiệp của con người, đặt ra những vấn đề nghiêm trọng về đạo đức và an ninh. Tự động hóa việc làm, với khả năng thay thế con người trong nhiều công việc, có thể dẫn đến tình trạng thất nghiệp hàng loạt, bất bình đẳng xã hội và bất ổn chính trị.

Để ứng phó với những rủi ro này, Dự thảo Luật CNCNS cần hoàn thiện các quy định về rủi ro AI theo những hướng cụ thể. Trước hết, Dự thảo Luật cần phân loại các hệ thống AI theo cấp độ rủi ro, dựa trên mức độ ảnh hưởng tiềm ẩn của chúng đối với con người, xã hội và môi trường.

Ví dụ, các hệ thống AI đơn giản, không ảnh hưởng lớn đến con người và xã hội, ví dụ như AI dùng trong trò chơi điện tử, có thể được xếp vào nhóm rủi ro thấp. Các hệ thống AI phức tạp hơn, có thể ảnh hưởng đến con người và xã hội ở một mức độ nhất định, ví dụ như AI dùng trong hệ thống khuyến nghị sản phẩm, có thể được xếp vào nhóm rủi ro trung bình. Các hệ thống AI có khả năng gây ra những thiệt hại nghiêm trọng cho con người và xã hội, ví dụ như AI dùng trong xe tự lái, hệ thống vũ khí tự động, cần được xếp vào nhóm rủi ro cao.

Dựa trên việc phân loại rủi ro, Dự thảo Luật CNCNS cần quy định những biện pháp kiểm soát rủi ro phù hợp với từng cấp độ. Đối với các hệ thống AI có rủi ro cao, việc kiểm tra, đánh giá định kỳ là vô cùng cần thiết để kịp thời phát hiện và ngăn chặn những nguy cơ tiềm ẩn. Việc thử nghiệm các hệ thống AI trước khi đưa vào ứng dụng thực tế cũng là biện pháp quan trọng, đặc biệt là đối với các hệ thống AI có rủi ro cao.

Bên cạnh đó, việc thiết lập cơ chế cảnh báo sớm để xử lý kịp thời các sự cố liên quan đến AI cũng cần được chú trọng. Quan trọng hơn, Luật cần đảm bảo con người luôn có khả năng giám sát, kiểm soát và can thiệp vào hoạt động của AI, đặc biệt là đối với các hệ thống AI có rủi ro cao.

Cuối cùng, việc xây dựng hệ thống phòng thủ vững chắc trước những nguy cơ tấn công mạng và thông tin sai lệch do AI cũng là yếu tố then chốt. Luật cần khuyến khích đầu tư vào nghiên cứu và phát triển các công nghệ an ninh mạng tiên tiến, có khả năng phát hiện và ngăn chặn các cuộc tấn công mạng tự động hóa. Nâng cao nhận thức cho người dân về nguy cơ deepfake và thông tin sai lệch do AI tạo ra cũng là biện pháp quan trọng. Bên cạnh đó, việc xây dựng cơ chế xác minh thông tin, chống lại thông tin sai lệch trên mạng cũng cần được chú trọng.

**Trách nhiệm pháp lý**

Dự thảo Luật CNCNS đã đề cập đến trách nhiệm pháp lý liên quan đến AI trong Điều 83, tuy nhiên, những quy định hiện tại còn chung chung, chưa đủ rõ ràng để xác định trách nhiệm của các chủ thể liên quan khi xảy ra sự cố do AI gây ra. Cuốn sách "Sóng thần công nghệ" đã phân tích sự phân tán quyền lực và khó khăn trong việc quy trách nhiệm khi AI ngày càng tự chủ và phổ biến. Khi AI có khả năng tự học hỏi, thích nghi và đưa ra quyết định độc lập, việc xác định ai chịu trách nhiệm khi AI gây ra thiệt hại sẽ trở nên phức tạp hơn.

Để khắc phục hạn chế này, Dự thảo Luật CNCNS cần hoàn thiện các quy định về trách nhiệm pháp lý liên quan đến AI theo những hướng cụ thể. Trước hết, Dự thảo Luật cần quy định rõ ràng trách nhiệm pháp lý của các chủ thể liên quan đến AI, bao gồm nhà phát triển, người sử dụng, và tổ chức quản lý.

Nhà phát triển chịu trách nhiệm về thiết kế, phát triển, thử nghiệm và bảo mật của hệ thống AI, đảm bảo AI hoạt động an toàn, không gây hại và tuân thủ các quy định pháp luật. Người sử dụng chịu trách nhiệm về việc sử dụng hệ thống AI một cách phù hợp, an toàn và tuân thủ các quy định pháp luật. Tổ chức quản lý chịu trách nhiệm về việc giám sát, kiểm soát và quản lý các hoạt động liên quan đến AI, đảm bảo AI được phát triển và ứng dụng một cách có trách nhiệm và có đạo đức.

Bên cạnh việc xác định rõ ràng trách nhiệm của các chủ thể, Dự thảo Luật cũng cần xây dựng cơ chế xử lý vi phạm, bồi thường thiệt hại do AI gây ra. Luật cần quy định rõ ràng các hình phạt áp dụng cho các hành vi vi phạm pháp luật liên quan đến AI, chẳng hạn như phạt tiền, đình chỉ hoạt động, thu hồi giấy phép,... Đồng thời, Luật cũng cần quy định rõ ràng cơ chế bồi thường thiệt hại cho các nạn nhân của các hành vi vi phạm pháp luật liên quan đến AI, bao gồm cả thiệt hại về vật chất, tinh thần và danh dự. Cuối cùng, việc xây dựng cơ chế giải quyết tranh chấp liên quan đến AI cũng là yếu tố quan trọng. Dự thảo Luật cần quy định cơ chế giải quyết tranh chấp liên quan đến AI, bao gồm cả tranh chấp giữa các chủ thể liên quan đến AI và tranh chấp giữa con người và AI.

**Thúc đẩy hợp tác quốc tế trong việc quản lý AI**

Dự thảo Luật CNCNS đã đề cập đến hợp tác quốc tế trong Điều 40 và 41, tuy nhiên, những quy định hiện tại chưa đủ mạnh mẽ để thúc đẩy hợp tác quốc tế trong việc quản lý AI. Cuốn sách "Sóng thần công nghệ" đã phân tích nguy cơ "chạy đua vũ trang" AI, sự phân mảnh quyền lực và khó khăn trong việc quản lý AI trên phạm vi toàn cầu. Khi AI ngày càng trở nên mạnh mẽ và phổ biến, việc hợp tác quốc tế là điều kiện tiên quyết để "kiềm tỏa" làn sóng công nghệ, đảm bảo cho sự phát triển và ứng dụng AI diễn ra một cách an toàn, có kiểm soát và có lợi cho toàn nhân loại.

Để thúc đẩy hợp tác quốc tế trong việc quản lý AI, Dự thảo Luật CNCNS cần bổ sung và hoàn thiện các quy định. Việc xây dựng các quy định và tiêu chuẩn chung về AI là bước đi quan trọng đầu tiên. Việt Nam cần tham gia tích cực vào các diễn đàn quốc tế về AI, đóng góp vào việc xây dựng các quy định và tiêu chuẩn chung về AI, đảm bảo tính thống nhất và hiệu quả trong việc quản lý AI trên phạm vi toàn cầu. Đồng thời, Luật cũng cần khuyến khích các tổ chức, DN và cá nhân Việt Nam tham gia vào các hoạt động tiêu chuẩn hóa quốc tế về AI. Bên cạnh đó, việc chia sẻ thông tin, kinh nghiệm và hợp tác nghiên cứu về AI cũng là yếu tố then chốt.

Luật cần quy định việc thiết lập cơ chế chia sẻ thông tin, kinh nghiệm và hợp tác nghiên cứu về AI với các quốc gia khác, đặc biệt là các quốc gia tiên tiến trong lĩnh vực AI. Hỗ trợ các hoạt động trao đổi học thuật, đào tạo và chuyển giao công nghệ về AI giữa Việt Nam và các quốc gia khác cũng cần được chú trọng.

Cuối cùng, việc xây dựng các cơ chế hợp tác quốc tế trong việc giám sát và kiểm soát rủi ro AI cũng là yếu tố quan trọng. Việt Nam cần tham gia vào các sáng kiến quốc tế về an ninh mạng, phòng chống deepfake và kiểm soát vũ khí tự động. Hợp tác với các quốc gia khác trong việc xây dựng các cơ chế cảnh báo sớm, chia sẻ thông tin và ứng phó với các sự cố liên quan đến AI cũng cần được đẩy mạnh.

**Kết luận**

Làn sóng công nghệ mới, với AI và sinh học là trọng tâm, đang ập đến với tốc độ chóng mặt, mang theo cả cơ hội và thách thức chưa từng có. Trong bối cảnh đó, việc hoàn thiện Dự thảo Luật CNCNS là nhiệm vụ cấp bách, mang tính sống còn, nhằm đảm bảo cho Việt Nam có thể "cưỡi" được làn sóng này, khai thác tối đa tiềm năng của công nghệ, đồng thời kiểm soát hiệu quả những rủi ro tiềm ẩn.

Để Luật thực sự phát huy hiệu quả, cần có sự chung tay của tất cả các bên liên quan, từ nhà nước, DN, nhà khoa học đến người dân. Chỉ có sự đồng lòng và nỗ lực chung mới có thể giúp Việt Nam xây dựng một tương lai công nghệ an toàn, bền vững và thịnh vượng.

**Tài liệu tham khảo**

[1] “Sóng thần công nghệ: Lời cảnh tỉnh và kêu gọi hành động.” [https://vietnamnet.vn/song-tha...](https://vietnamnet.vn/song-than-cong-nghe-loi-canh-tinh-va-keu-goi-hanh-dong-2265539.html) (accessed Jul. 13, 2024).

[2] “Glossary - Artificial Intelligence.” [https://www.coe.int/en/web/art...](https://www.coe.int/en/web/artificial-intelligence/glossary) (accessed Jul. 13, 2024).

[3] “Reasoning through arguments against taking AI safety seriously - Yoshua Bengio.” [https://yoshuabengio.org/2024/...](https://yoshuabengio.org/2024/07/09/reasoning-through-arguments-against-taking-ai-safety-seriously/?fbclid=IwZXh0bgNhZW0CMTAAAR2c249AQTdk2WDfXhZkIQS4pdO4GMBWxa44kd7OecrLwjtvPslA4Dz208w_aem_zEyBKxHi5lFHdsZLZuswEA) (accessed Jul. 13, 2024)./.