**Hành lang pháp lý đối với AI của Việt Nam**

**Với sự phát triển bùng nổ của trí tuệ nhân tạo (AI), việc quản trị rủi ro có thể phát sinh từ công nghệ này trở thành một thách thức lớn cho chính phủ nhiều quốc gia, bao gồm cả Việt Nam. Liệu chúng ta đã sẵn sàng xây dựng một khuôn khổ pháp lý phù hợp để vừa bảo vệ quyền riêng tư vừa thúc đẩy sự phát triển bền vững của công nghệ đột phá này?**

Trong thế giới của trò chơi điện tử “Watch Dogs”, người chơi nhập vai vào Aiden Pearce, một hacker có khả năng kiểm soát toàn bộ hệ thống thành phố qua một chiếc điện thoại thông minh. Trong trò chơi này, người chơi có thể truy cập và thao túng dữ liệu cá nhân của công dân, bao gồm cả dữ liệu sinh trắc học, thông qua một hệ thống AI giám sát toàn thành phố trong thời gian thực.

Từ việc điều khiển giao thông, theo dõi người dân thông qua hệ thống camera giám sát (CCTV) đến phá vỡ an ninh, Pearce dường như nắm trong tay quyền lực tối cao nhờ vào dữ liệu và công nghệ. Điều mà “Watch Dogs” đưa ra không chỉ là một viễn cảnh khoa học viễn tưởng mà còn là một lời cảnh báo về quyền lực của trí tuệ nhân tạo và dữ liệu trong thế giới thực.

Những gì từng chỉ tồn tại trong trò chơi giờ đây đang dần trở thành hiện thực. Dù chỉ là một sản phẩm giải trí, “Watch Dogs” đã đặt ra câu hỏi quan trọng về những gì có thể xảy ra trong thực tế nếu các hệ thống giám sát AI thời gian thực không được quản lý nghiêm ngặt.

Trải nghiệm từ “Watch Dogs” đã tạo ra một bức tranh sống động về những gì có thể xảy ra nếu AI không được quản lý đúng cách. Trong thực tế cũng như vậy, nếu thiếu một khung pháp lý chặt chẽ, các công nghệ AI tại Việt Nam có thể bị lạm dụng, gây ra hậu quả nghiêm trọng cho quyền riêng tư và an toàn xã hội.

Việc lựa chọn mô hình điều chỉnh ảnh hưởng lớn đến cách thức giải quyết các vấn đề liên quan đến đạo đức và tính hiệu quả trong quản lý và sử dụng AI. Hiểu rõ các ưu và nhược của các mô hình này giúp Việt Nam có thể lựa chọn một trong các mô hình hoặc phối hợp các mô hình để tiến hành xây dựng hành lang pháp lý đối với AI.

**Các nỗ lực lập pháp trong việc điều chỉnh AI ở các nước**

Nhìn ra thế giới, EU là khu vực tiên phong trong việc xây dựng một khung pháp lý toàn diện nhằm điều chỉnh và quản lý các hệ thống trí tuệ nhân tạo theo tốc độ phát triển nhanh chóng của công nghệ này. Đạo luật AI của EU (AI Act) được thiết kế để bảo vệ người dân trước những rủi ro tiềm ẩn từ các ứng dụng AI và thúc đẩy sự phát triển an toàn, minh bạch. Đạo luật này không chỉ định hướng cách thức phát triển, triển khai, và giám sát AI nhằm bảo vệ quyền lợi của người dân mà còn đưa ra các tiêu chuẩn nghiêm ngặt về bảo vệ dữ liệu cá nhân và đảm bảo tính minh bạch trong quá trình nghiên cứu, phát triển và sử dụng AI.

Được Ủy ban châu Âu đề xuất vào tháng 4/2021, Đạo luật AI của EU bao gồm 85 điều và 9 chương, cùng với 8 phụ lục (annexes), tập trung vào việc phân loại rủi ro và quy định các biện pháp kiểm soát đối với các ứng dụng AI có nguy cơ cao. Những quy định này đảm bảo tuân thủ các tiêu chuẩn về quyền riêng tư, an toàn, và tính minh bạch trong quá trình phát triển và triển khai công nghệ AI.

Tương tự, Anh đã đưa ra một khung pháp lý tập trung vào các nguyên tắc liên ngành và dựa trên kết quả, mặc dù chưa được pháp điển hóa. Khung pháp lý này bao gồm năm nguyên tắc chính: an toàn, bảo mật và độ bền vững; minh bạch và khả năng giải thích hợp lý; công bằng; trách nhiệm và quản trị; khả năng kháng nghị và bồi thường, với các quy định dự kiến sẽ có những can thiệp lập pháp cụ thể trong tương lai (2).

Trong khi đó, Hoa Kỳ đang phát triển khung pháp lý AI với trọng tâm là quản lý rủi ro và thúc đẩy đổi mới. Sắc lệnh Hành pháp về AI của Tổng thống Joe Biden vào năm 2023 tập trung vào đảm bảo sự phát triển AI an toàn và có trách nhiệm, với các quy định cụ thể về bảo vệ dữ liệu và trách nhiệm giải trình (3).

Ngay gần Việt Nam, Trung Quốc đã xây dựng một khung pháp lý nhằm cân bằng giữa phát triển công nghệ AI và đảm bảo an toàn pháp lý. Pháp luật AI của Trung Quốc bao gồm các luật chung như Luật An ninh mạng và các biện pháp cụ thể cho ứng dụng AI, tập trung vào các hướng dẫn đạo đức, tiêu chuẩn bảo vệ dữ liệu cá nhân, và cơ chế trách nhiệm giải trình (1).

Việc nghiên cứu các quy định từ EU và thực tiễn từ các quốc gia như châu Âu, Trung Quốc, Anh, và Hoa Kỳ mang lại cho Việt Nam những kinh nghiệm quan trọng.

**Giám sát theo thời gian thực**

Đạo luật AI của EU nhấn mạnh vai trò quan trọng của khái niệm “thời gian thực” (real-time) áp dụng đối với các hệ thống trí tuệ nhân tạo, đặc biệt là những hệ thống liên quan đến nhận diện sinh trắc học. Bởi thực tế hiện nay các hệ thống AI thời gian thực được thiết kế để xử lý và phân tích dữ liệu ngay khi dữ liệu được thu thập, chẳng hạn như video từ camera an ninh hoặc các thiết bị ghi hình tương tự. Điều này cho phép nhận diện và so sánh dữ liệu sinh trắc học ngay lập tức, từ đó đưa ra quyết định trong thời gian thực.

Một ví dụ khác điển hình là hệ thống giám sát công nghệ cao của Trung Quốc, nơi sử dụng AI và công nghệ nhận diện sinh trắc học để theo dõi và quản lý xã hội. Các hệ thống này có khả năng quét và nhận diện khuôn mặt từ hàng triệu camera trên khắp đất nước, sau đó so sánh với cơ sở dữ liệu của chính phủ trong thời gian thực để nhận diện cá nhân. Hệ thống này không chỉ giúp phát hiện các hành vi phạm tội mà còn được sử dụng để giám sát xã hội, bao gồm việc theo dõi hoạt động của các cá nhân và nhóm trong thời gian thực (4).

Trong bối cảnh AI, tự do của cá nhân đòi hỏi phải giảm thiểu sự can thiệp bất hợp pháp hoặc không công bằng, bao gồm cả sự giám sát không chính đáng, lừa đảo và thao túng không công bằng.

Trung Quốc đã phát triển một mạng lưới giám sát mạnh mẽ, kết hợp AI với dữ liệu sinh trắc học như nhận diện khuôn mặt, vân tay, và thậm chí cả giọng nói để tạo nên một hệ thống quản lý công dân toàn diện. Việc sử dụng AI trong hệ thống này không chỉ dừng lại ở nhận diện mà còn mở rộng sang dự đoán hành vi, giúp chính phủ có thể đưa ra các biện pháp ngăn chặn kịp thời (5).

**Lựa chọn mô hình điều chỉnh phù hợp**

Khi áp dụng các hệ thống trí tuệ nhân tạo, việc lựa chọn mô hình điều chỉnh (*regulating model*) phù hợp không chỉ ảnh hưởng đến hiệu quả mà còn liên quan chặt chẽ đến vấn đề đạo đức và quyền riêng tư. Theo nghiên cứu của Sheehy và Feaver (2025) (6), có ba mô hình chính trong cấu trúc hệ thống quy định: Tập trung (*Centralised*), Phi tập trung (*Decentred*), và Phân tán (*Distributed*).

– **Mô hình Tập trung (Centralised):** Trong mô hình này, quyền lực và trách nhiệm chính đều được tập trung vào một cơ quan hoặc tổ chức duy nhất. Nhà nước giữ phần lớn quyền kiểm soát và chỉ ủy quyền một số nhiệm vụ hành chính cho các cơ quan quản lý. Mô hình này đặt ra cơ chế kiểm soát đơn giản và cứng nhắc, đảm bảo sự nhất quán và tuân thủ quy định nghiêm ngặt, đồng thời dễ dàng theo dõi trách nhiệm giải trình.

– **Mô hình Phi tập trung (Decentred):** Quyền lực được phân phối cho nhiều cơ quan hoặc tổ chức khác nhau, bao gồm cả các cơ quan quản lý và các tổ chức ngoài công quyền. Mô hình này cho phép các tổ chức bên ngoài tham gia vào các chức năng hành chính hoặc thực thi, giúp phân phối trách nhiệm và cải thiện hiệu quả quản lý.

– **Mô hình Phân tán (Distributed):** Quyền lực và trách nhiệm được phân bổ cho nhiều đối tượng khác nhau, bao gồm cả các tổ chức tư nhân và các bên liên quan. Các tổ chức có thể phát triển các tiêu chuẩn, quy trình, và cơ chế đánh giá riêng của mình. Mô hình này khuyến khích sự tự quy định và đồng quy định, đáp ứng nhu cầu đa dạng và cải thiện hiệu quả quản lý trong các tình huống phức tạp.

Việc lựa chọn mô hình điều chỉnh ảnh hưởng lớn đến cách thức giải quyết các vấn đề liên quan đến đạo đức và tính hiệu quả trong quản lý và sử dụng AI. Mô hình Tập trung đảm bảo sự kiểm soát và trách nhiệm giải trình rõ ràng, mô hình phi tập trung cung cấp sự linh hoạt và phân phối trách nhiệm giữa nhiều bên, và mô hình phân tán khuyến khích sự tham gia và tự quy định. Hiểu rõ các ưu và nhược của các mô hình này giúp Việt Nam có thể lựa chọn một trong các mô hình hoặc phối hợp các mô hình để tiến hành xây dựng hành lang pháp lý đối với AI.

Ví dụ, Đạo luật AI của EU đã kết hợp các mô hình này bằng cách áp dụng từng mô hình cho các loại rủi ro khác khau. Điều này có nghĩa là các hệ thống AI được phân loại và điều chỉnh theo các mức độ rủi ro từ thấp, trung bình đến cao dựa trên khả năng gây hại hoặc ảnh hưởng của chúng đến quyền lợi con người và xã hội. Điều này giúp các nhà quản lý tập trung nguồn lực và các biện pháp giám sát vào những ứng dụng AI có rủi ro cao nhất, từ đó tối ưu hóa việc sử dụng các nguồn lực và đảm bảo hiệu quả trong việc bảo vệ lợi ích công cộng.

**Quyền riêng tư là trung tâm**

EU từ lâu đã coi trọng quyền riêng tư và bảo vệ dữ liệu cá nhân, thể hiện qua việc ban hành Quy định bảo vệ dữ liệu chung (GDPR). GDPR đã đặt nền móng cho một khung pháp lý toàn diện nhằm đảm bảo rằng dữ liệu cá nhân của người dùng được xử lý một cách an toàn và minh bạch. Trong bối cảnh công nghệ AI và dữ liệu lớn phát triển bùng nổ, khả năng xâm phạm quyền riêng tư ngày càng gia tăng, đặc biệt khi các hệ thống AI có thể thu thập và phân tích một lượng lớn dữ liệu cá nhân nhạy cảm.

Nhận thức được những rủi ro tiềm ẩn từ việc xử lý dữ liệu cá nhân, EU đã rất chú trọng đưa các nguyên tắc bảo vệ quyền riêng tư vào Đạo luật AI. Khoản 69 trong phần Mở đầu của đạo luật nhấn mạnh rằng quyền bảo vệ quyền riêng tư và dữ liệu cá nhân phải được đảm bảo trong suốt vòng đời của hệ thống AI. Các biện pháp bảo vệ có thể bao gồm không chỉ ẩn danh và mã hóa, mà còn sử dụng công nghệ cho phép thuật toán làm việc với dữ liệu mà không cần truyền tải hoặc sao chép dữ liệu giữa các bên.

Điều này nhấn mạnh sự cần thiết của việc bảo vệ dữ liệu cá nhân trong quá trình phát triển và triển khai AI, đảm bảo rằng công nghệ mới này không vi phạm các nguyên tắc đã được đặt ra trong GDPR. Bằng cách này, EU tiếp tục khẳng định cam kết của mình đối với việc đặt con người vào trung tâm của các tiến bộ công nghệ, bảo vệ quyền tự do cá nhân, nhân phẩm và quyền tự quyết trong thời đại kỹ thuật số.

**Nhấn mạnh đến các nguyên tắc đạo đức**

Trong nghiên cứu AI, có một thuật ngữ quan trọng “thuật toán đạo đức” (*ethical algorithm*). Dễ hiểu,  thuật toán là một chuỗi các hướng dẫn được xác định rõ ràng và chính xác để thực hiện một nhiệm vụ cụ thể, ví dụ như sắp xếp một dãy số từ nhỏ đến lớn. Mỗi bước của thuật toán phải rõ ràng và kết quả luôn đảm bảo đúng theo yêu cầu, bất kể đầu vào ban đầu (7). Theo Bộ nguyên tắc đạo đức do Nhóm chuyên gia cấp cao về AI (AI HLEG) của Ủy ban châu Âu phát triển vào năm 2019, có bảy nguyên tắc cơ bản để phát triển một hệ thống đáng tin cậy và có đạo đức, bao gồm: (i) Quyền tự quyết của con người và giám sát *(human agency and oversight)*; (ii) Tính an toàn và độ tin cậy kỹ thuật *(technical robustness and safety)*; (iii) Quyền riêng tư và quản trị dữ liệu *(privacy and data governance)*; (iv) Tính minh bạch (transparency); (v) Sự đa dạng, không phân biệt đối xử và công bằng *(diversity, non-discrimination, and fairness)*; (vi) Phúc lợi xã hội và môi trường *(environmental and societal well-being)*; (vii) Trách nhiệm giải trình *(accountability)*.

Lấy ví dụ, nguyên tắc quyền riêng tư và quản trị dữ liệu nhấn mạnh rằng AI phải được phát triển và sử dụng theo cách bảo vệ quyền riêng tư của người dùng, đảm bảo rằng dữ liệu cá nhân được thu thập, lưu trữ, và xử lý một cách minh bạch và an toàn. Điều này không chỉ giúp AI tuân thủ các quy định pháp luật hiện hành mà còn xây dựng một hệ thống AI nhân văn, tập trung vào con người. EU AI Act đã lồng ghép những giá trị này vào các điều khoản cụ thể, sử dụng phương pháp dẫn chiếu ngược lại các nguyên tắc đạo đức đã có trước đây, nhằm đảm bảo mọi hoạt động liên quan đến AI đều được thực hiện dưới sự giám sát chặt chẽ về quyền riêng tư và đạo đức.

Bảo vệ quyền riêng tư còn liên quan mật thiết đến tính minh bạch và trách nhiệm của AI. Khi các hệ thống AI hoạt động mà không có sự giám sát đầy đủ, nguy cơ lạm dụng dữ liệu và vi phạm quyền riêng tư càng cao. Quy định AI của EU đặt ra các yêu cầu rõ ràng về bảo vệ dữ liệu, yêu cầu các nhà phát triển AI phải tuân thủ các tiêu chuẩn về quyền riêng tư và minh bạch. Điều này bao gồm việc chỉ định rõ ràng trách nhiệm của các bên liên quan và tạo ra các cơ chế giám sát để đảm bảo sự tuân thủ. Bằng cách này, EU AI Act không chỉ ngăn chặn những rủi ro tiềm ẩn mà còn tăng cường sự tin tưởng của công chúng vào AI, từ đó thúc đẩy sự phát triển bền vững của công nghệ này.

Bên cạnh nguyên tắc vừa nêu trên, yếu tố sự tự quyết của cá nhân (*freedom of the individual*) trong bộ nguyên tắc đạo đức của AI HLEG cũng nhấn mạnh rằng con người cần được tự do đưa ra các quyết định cho chính mình mà không bị can thiệp từ bên ngoài, bao gồm cả các hệ thống AI. Điều này đòi hỏi sự tự do khỏi sự ép buộc và giám sát không cần thiết, cũng như các mối đe dọa đến quyền tự trị tinh thần và sức khỏe tâm thần. Trong bối cảnh AI, tự do của cá nhân đòi hỏi phải giảm thiểu sự can thiệp bất hợp pháp hoặc không công bằng, bao gồm cả sự giám sát không chính đáng, lừa đảo và thao túng không công bằng. Điều này liên quan mật thiết đến quyền riêng tư, vì quyền tự do cá nhân bao gồm cả quyền được bảo vệ sự riêng tư và đời sống cá nhân. EU AI Act đã quy định rõ ràng các yêu cầu này, đảm bảo rằng các hệ thống AI phải hoạt động một cách minh bạch và có trách nhiệm, từ đó bảo vệ quyền tự do và quyền riêng tư của cá nhân.

**Hỗ trợ các doanh nghiệp công nghệ**

Ở các quốc gia, việc đặt ra rất nhiều nguyên tắc như vậy để đảm bảo quyền con người trong bối cảnh phát triển quá nhanh chóng của công nghệ đều song song với việc có các chính sách hỗ trợ các doanh nghiệp công nghệ. Tương tự, sự phát triển của trí tuệ nhân tạo đang mở ra nhiều cơ hội cho các doanh nghiệp tại Việt Nam. Do vậy, một trong những cơ hội quan trọng mà Luật Công nghệ số mang lại, khi được ban hành, là các chính sách ưu đãi về thuế. Chính phủ có thể xem xét việc giảm thuế thu nhập doanh nghiệp hoặc áp dụng các mức thuế ưu đãi đặc biệt cho các doanh nghiệp hoạt động trong lĩnh vực công nghệ số, đặc biệt là AI. Điều này không chỉ giúp giảm gánh nặng tài chính cho doanh nghiệp mà còn khuyến khích họ đầu tư nhiều hơn vào nghiên cứu và phát triển các sản phẩm AI tiên tiến.

Bên cạnh các chính sách thuế, Dự thảo Luật Công nghệ số còn đề cập đến việc hỗ trợ các doanh nghiệp trong việc tiếp cận đất đai và cơ sở hạ tầng. Các khu công nghệ cao, trung tâm dữ liệu và các khu công nghiệp chuyên dụng có thể được phát triển để cung cấp cơ sở hạ tầng hiện đại và thuận lợi cho các doanh nghiệp AI. Điều này sẽ giúp các doanh nghiệp tiết kiệm chi phí đầu tư ban đầu và tập trung vào việc phát triển sản phẩm và dịch vụ AI có giá trị gia tăng cao.

Ngoài ra, Chính phủ có thể tạo ra các chương trình hỗ trợ đào tạo và phát triển nguồn nhân lực trong lĩnh vực AI. Việc đào tạo và phát triển kỹ năng cho người lao động trong ngành công nghệ số sẽ giúp các doanh nghiệp Việt Nam nâng cao năng lực cạnh tranh, đồng thời đảm bảo rằng họ có đủ nguồn nhân lực chất lượng cao để đáp ứng nhu cầu ngày càng tăng của ngành công nghiệp AI.

\*\*\*

Với sự phát triển nhanh chóng của trí tuệ nhân tạo, Việt Nam đang đứng trước những thách thức lớn trong việc quản lý và kiểm soát các công nghệ mới này. Như đã thấy từ các ví dụ trong trò chơi “Watch Dogs” và thực tế tại các quốc gia tiên tiến như Trung Quốc, việc quản lý AI đòi hỏi một khung pháp lý đủ mạnh để bảo vệ quyền riêng tư và ngăn chặn sự lạm dụng công nghệ.

Việt Nam có thể học hỏi từ kinh nghiệm lập pháp của một số quốc gia hay khu vực để xây dựng một khung pháp lý phù hợp. Việc triển khai các quy định giám sát chặt chẽ không chỉ giúp ngăn chặn các rủi ro tiềm ẩn mà còn tạo nền tảng cho sự phát triển bền vững của AI trong tương lai. Tất nhiên, giới nghiên cứu về AI, ở nhiều khía cạnh khác nhau, cần tiếp tục thảo luận cụ thể từng điểm trên để đảm bảo các quy định sao cho cân bằng được giữa đổi mới và lợi ích của người dân, bảo vệ quyền lợi của người dân và doanh nghiệp nhưng cũng không cản trở sự sáng tạo trong nghiên cứu phát triển công nghệ mới.

———

(\*) Các tác giả Tô Kiến Lương, Mai Nguyễn Dũng đang công tác tại Khoa Luật, CELG, Đại học UEH.

**Tài liệu tham khảo:**

**1.** Wang, Z. (2023). *Legal regulation of artificial intelligence: Experience of Chin*a. ResearchGate. https://www.researchgate.net/publication/379147422\_Legal\_Regulation\_of\_Artificial\_Intelligence\_Experience\_of\_China#pf5

2. Gallo, V., & Nair, S. (2024). “The UK’s framework for AI regulation: Agility is prioritised, but future legislation is likely to be needed”. *Deloitte*. https://www.deloitte.com/uk/en/Industries/financial-services/blogs/the-uks-framework-for-ai-regulation.html

3. The White House. (2023). *Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence*.

*Bài đăng Tia Sáng số 18/2024*

https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence

4. Dai J. (2022). “China’s surveillance state should not be a model for the rest of the world”, *The Guardian.*

https://www.theguardian.com/commentisfree/2022/jul/25/china-surveillance-state-model-world accessed 20 August 2024.

5. Mozur P, Kessel J and Chan M. (2018). “Inside China’s Dystopian Dreams: A.I., Shame and Lots of Cameras”. *The New York Times.*

https://www.nytimes.com/2018/07/08/business/china-surveillance-technology.html accessed 20 August 2024.

6. Sheehy, B., & Feaver, D. (2015). *Designing Effective Regulation: A Positive Theory*. University of New South Wales Law Journal, 38(3), 961.

7. Kearns, M., & Roth, A. (2019). *The ethical algorithm: The science of socially aware algorithm design.* Oxford University Press.

Tác giả: Tô Kiến Lương, Mai Nguyễn Dũng,

Nguồn: https://tiasang.com.vn/doi-moi-sang-tao/hanh-lang-phap-ly-doi-voi-ai-cua-viet-nam/