**Luật về trí tuệ nhân tạo và tầm nhìn chiến lược của Trung Quốc**

Trung Quốc đang tiến tới giai đoạn định hình đạo luật về trí tuệ nhân tạo (AI) sau một kế hoạch bài bản hướng đến việc cân bằng giữa chính sách khuyến khích phát triển và điều tiết rủi ro. Đây có thể là một trong những đạo luật đầu tiên trên thế giới về AI (bên cạnh đạo luật của EU đã được ban hành) và sẽ có tác động rất lớn đến chính sách, pháp luật toàn cầu nhằm điều chỉnh các hoạt động liên quan đến AI.

**Kế hoạch rõ ràng, hành động nhanh chóng**

Ngày 20-7-2017, Quốc vụ viện Trung Quốc lần đầu tiên công bố “Kế hoạch phát triển AI thế hệ mới” tầm nhìn đến 2030 với nhiều tham vọng và thời gian biểu rõ ràng theo lộ trình: bắt kịp các nước phương Tây vào năm 2023, vượt qua các nước phương Tây vào năm 2025, dẫn đầu thế giới vào năm 2030. Ngay sau đó, hàng loạt chính sách đã được triển khai từ nhiều cấp độ để tập trung nguồn lực, động lực cho các hoạt động nghiên cứu, phát triển và ứng dụng AI.

Chỉ sau hơn năm năm, nước này đã có những thành tựu đáng kể và đang từng bước hiện thực hóa bản kế hoạch nhiều tham vọng của mình. Theo Báo cáo chỉ số AI (Artificial Intelligence Index Report) của Đại học Stanford công bố năm 2024, Trung Quốc đang nằm trong nhóm các quốc gia dẫn đầu về đầu tư, nghiên cứu, sáng chế và ứng dụng AI. Theo đó, từ năm 2022, nước này đã thống trị bảng xếp hạng toàn cầu về ngành robot công nghiệp (với số lượng lắp đặt chiếm tỷ lệ 52,4%) và số lượng bằng sáng chế liên quan đến AI (chiếm 61,1%). Năm 2023, Trung Quốc là nước đứng thứ 2 thế giới về giá trị các khoản đầu tư vào AI (sau Mỹ).

Bên cạnh đó, theo một nghiên cứu được thực hiện bởi MacroPolo (có trụ sở tại Viện Paulson, Chicago, Mỹ), gần một nửa số nhà nghiên cứu hàng đầu về AI thuộc tốp 20% thế giới đến từ Trung Quốc.

**Tận dụng cơ hội, ứng phó với thách thức**

Theo đuổi chiến lược phát triển AI dường như là nhiệm vụ không thể đảo ngược, không chỉ riêng đối với Trung Quốc. Đầu năm 2023, sự xuất hiện của ChatGPT một lần nữa giúp nhân loại nhận ra AI đã sớm không còn là câu chuyện của tương lai mà đã và đang phát triển đến một trình độ vượt trội, có thể thay thế con người trong nhiều lĩnh vực và tạo ra rất nhiều thách thức kể cả về đạo đức, pháp luật hay xã hội.

Quan sát quá trình quản lý và điều tiết các hoạt động liên quan đến AI của Trung Quốc, có thể nhận thấy khá rõ ràng sự vận động tuần tự của nước này, đặc biệt trong việc lựa chọn cách thức tiếp cận pháp lý phù hợp với từng giai đoạn.

Tại “Dự thảo kiến nghị bởi chuyên gia”, sự phân bổ trách nhiệm rõ ràng hơn trong chuỗi giá trị AI cũng được nêu bật nhằm ngăn cản các nhà phát triển mô hình AI trốn tránh trách nhiệm thông qua ngôn ngữ hợp đồng hoặc với lý do bảo vệ bí mật kinh doanh.

Ngay trong bản “Kế hoạch phát triển AI thế hệ mới” ban hành năm 2017, Quốc vụ viện nước này cũng đã đưa ra lộ trình điều tiết và điều chỉnh pháp luật đối với AI. Theo đó, đến năm 2025, Trung Quốc sẽ thiết lập hệ thống quy phạm điều chỉnh AI bao gồm đạo luật về AI, bộ quy tắc đạo đức, hệ thống chính sách thông tin về AI để hình thành khả năng đánh giá và kiểm soát an ninh, rủi ro đối với AI.

Từ năm 2017 đến nay, các quy tắc ứng xử ban đầu với AI đã được ban hành bởi Ủy ban chuyên môn quản trị AI thế hệ mới(1). Đầu tiên phải kể đến “Bộ nguyên tắc quản trị AI thế hệ mới: Phát triển AI có trách nhiệm”, được ban hành vào tháng 6-2019 và theo sau đó là “Chuẩn mực đạo đức cho AI thế hệ mới”, được ban hành vào tháng 9-2021. Hai văn bản này đều mang tính chất hướng dẫn, định hướng hành vi của các chủ thể nên không kèm theo yêu cầu bắt buộc và các chế tài. Tuy vậy, những quy tắc xử sự trong hai văn bản này là nền móng quan trọng cho việc định hình khung pháp lý điều chỉnh AI.

Bước ngoặt đáng kể nhất phải kể đến động thái ban hành văn bản pháp lý về “Các biện pháp tạm thời để quản lý các dịch vụ AI” vào tháng 7-2023 bởi bảy cơ quan quản lý dẫn đầu bởi Cục Quản lý không gian mạng Trung Quốc.

Tiếp ngay sau đó, dự thảo Luật về AI cũng đã được liệt kê trong danh sách đính kèm Kế hoạch lập pháp. Đây được cho là nỗ lực thứ hai, chỉ sau EU trong việc ban hành khung pháp lý riêng nhằm điều chỉnh các vấn đề liên quan đến AI. Những nỗ lực này nhằm đảm bảo các điều kiện tốt hơn cho việc phát triển và sử dụng công nghệ, nhưng cũng kiểm soát và ngăn ngừa những mặt trái mà nó có thể gây ra đối với cuộc sống của con người.

Trong khi chờ dự thảo luật được hoàn thiện, công bố và lấy ý kiến, ngày 18-3-2024 vừa qua, nhóm nghiên cứu gồm bảy trường đại học kết hợp Viện Nghiên cứu Công nghệ thông tin và Truyền thông Trung Quốc (CAICT) đã công bố một “Dự thảo kiến nghị bởi chuyên gia” (Expert Proposal) nhằm phục vụ công tác góp ý, tham mưu cho dự thảo luật sắp tới. Dự thảo do các chuyên gia soạn thảo được công bố như một bản luật mẫu kiến nghị bao gồm chín chương và 96 điều, nội dung bao gồm các nguyên tắc chung, thúc đẩy phát triển, bảo vệ quyền lợi người dùng, nghĩa vụ và chuẩn mực, quản lý giám sát, lĩnh vực đặc biệt, hợp tác quốc tế và trách nhiệm pháp lý.

**Xác định nguyên tắc tiếp cận pháp lý**

Thông qua hệ thống văn bản hướng dẫn, các quy định pháp lý tại “Các biện pháp tạm thời để quản lý các dịch vụ AI” năm 2023 cũng như “Dự thảo kiến nghị bởi chuyên gia” cho luật về AI năm 2024, quan điểm lập pháp của giới học giả và quản lý nước này đối với khung pháp lý điều chỉnh AI trong tương lai dường như đã bắt đầu định hình và thể hiện. Theo đó, người viết cho rằng Trung Quốc đang xây dựng khung pháp lý dựa trên bốn nguyên tắc tiếp cận pháp lý sau đây:

Thứ nhất, việc lựa chọn phương thức tiếp cận phải đảm bảo lấy con người làm trung tâm. Cụ thể, các quyền con người và quyền dân sự cơ bản phải là trọng tâm xem xét trong quá trình xây dựng khung pháp lý cho AI. Ngoài ra, cần phải đặc biệt lưu tâm đến quyền và lợi ích của các nhóm đặc thù, nhóm yếu thế trong tiếp cận không gian số (ví dụ như người già, người khuyết tật...) nhằm đảm bảo quyền được giải thích cũng như quyền được từ chối sử dụng AI cho những nhóm người này.

Thứ hai, nguyên tắc tuân thủ là một trong những nguyên tắc quan trọng nhất. Cụ thể, việc cung cấp và sử dụng AI phải đảm bảo tuân thủ nghiêm ngặt các quy định pháp luật, tôn trọng quy chuẩn và đạo đức xã hội, nghiêm cấm tạo ra các nội dung kích động, gây nguy hiểm cho an ninh, lợi ích quốc gia, làm tổn hại đến hình ảnh đất nước, nghiêm cấm nội dung bạo lực, khiêu dâm, phân biệt đối xử, thông tin sai lệch... Căn cứ quan trọng để đối chiếu, xác minh là các quy định pháp luật hiện hành, cũng như “Chuẩn mực đạo đức cho AI thế hệ mới” ban hành năm 2021.

Thứ ba, nghĩa vụ công khai và minh bạch phải đặc biệt được nhấn mạnh. Giống như các nền tảng cung cấp dịch vụ Internet, nhà cung cấp dịch vụ AI cũng đóng vai trò người gác cổng, đảm bảo dịch vụ AI được cung cấp sẽ trở nên minh bạch hơn và nội dung được tạo ra chính xác và đáng tin cậy hơn. Đối với người dùng, nhà cung cấp có nghĩa vụ làm rõ và công bố đối tượng người dùng mục tiêu, các trường hợp và cách sử dụng dịch vụ, hướng dẫn người dùng hiểu và sử dụng dịch vụ AI một cách khoa học và hợp lý.

Thứ tư, nghĩa vụ tôn trọng quyền sở hữu trí tuệ và bảo mật thông tin người dùng cũng phải được ưu tiên quy định cụ thể. Đặc biệt là trong quá trình thu thập và xử lý thông tin (training-data processing), chú thích dữ liệu (data Annotation), thiết lập và thu thập thông tin hồ sơ sử dụng của người dùng...

Nhìn chung, thay vì đặt ra giới hạn và hạn chế sự phát triển của AI, khung pháp lý hiện tại về AI của Trung Quốc tập trung khá nhiều vào việc tìm kiếm điểm cân bằng giữa khuyến khích phát triển và phòng ngừa rủi ro, giữa đổi mới và an toàn. Đây là một góc nhìn có thể còn gây tranh cãi nhưng chắc chắn tốt hơn việc chỉ nhìn nhận và lên án AI như một “thiên thạch” đang bay đến và đe dọa tồn vong của nhân loại. Đây cũng là cơ sở quan trọng để xây dựng các nguyên tắc phát triển bền vững, thúc đẩy giám sát chặt chẽ hơn các loại AI quan trọng, trong tình huống đặc biệt và xây dựng khung pháp lý thử nghiệm (sandbox) phù hợp.

Tại “Dự thảo kiến nghị bởi chuyên gia”, sự phân bổ trách nhiệm rõ ràng hơn trong chuỗi giá trị AI cũng được nêu bật nhằm ngăn cản các nhà phát triển mô hình AI trốn tránh trách nhiệm thông qua ngôn ngữ hợp đồng hoặc với lý do bảo vệ bí mật kinh doanh.

Ngoài ra, một mục tiêu lập pháp nữa cũng được thể hiện thống nhất là đảm bảo sự chủ động của con người đối với AI thông qua cả giám sát trực tiếp và phương tiện kỹ thuật, kể cả khi AI hoạt động tự chủ. Do đó, nhóm chuyên gia cũng cho rằng cách tiếp cận theo Danh sách tiêu cực (Negative list), nhiều tương đồng với phương pháp phân loại AI theo rủi ro để quản lý tại đạo luật AI của EU, là phù hợp với Trung Quốc. Bên cạnh đó, các chuyên gia cũng đề xuất thành lập một cơ quan mới là Văn phòng AI quốc gia (đầu tiên trên toàn thế giới) để tập trung điều phối và giám sát việc quản lý công nghệ AI, tránh trường hợp phân tán giám sát hoặc chồng chéo quản lý và trách nhiệm hành chính.

(1) Thành lập vào tháng 2-2019, trực thuộc Bộ Khoa học và Công nghệ Trung Quốc. Xem thêm: [*http://www.ircip.cn/web/1044764-1044764.html?id=26645&newsid=1118169*](http://www.ircip.cn/web/1044764-1044764.html?id=26645&newsid=1118169)

Tác giả: Nguyễn Ngọc Phương Hồng - Lưu Minh Sang, Trường Đại học Kinh tế - Luật, ĐHQG TPHCM.

Nguồn: https://thesaigontimes.vn/luat-ve-tri-tue-nhan-tao-va-tam-nhin-chien-luoc-cua-trung-quoc/