

TRÁCH NHIỆM ĐẠO ĐỨC VÀ PHÁP LÝ VỀ ROBOT: MỘT SỐ VẤN ĐỀ LÝ LUẬN, THỰC TIỄN*

NCS. Hoàng Thị Bích Ngọc - PGS.TS Vũ Công Giao

Tóm tắt

Bài viết tổng hợp thông tin trong báo cáo phân tích về những thách thức về đạo đức và pháp lý đang đặt ra với nhân loại trong việc chế tạo và sử dụng người máy (robot) thế hệ mới có sự tích hợp trí tuệ nhân tạo (Report of COMEST on Robotics Ethics) của Hội đồng Thế giới của UNESCO về Tri thức khoa học và công nghệ (World Commission on the Ethics of Scientific Knowledge and Technology of UNESCO - COMEST) và một số tài liệu khác, trong đó cố gắng làm rõ các vấn đề sau: Khái niệm, lịch sử phát triển và ứng dụng của người máy trong đời sống của nhân loại; Những thách thức về đạo đức và pháp lý đang đặt ra với việc sử dụng người máy; Giải pháp cho những thách thức đạo đức đang đặt ra với nhân loại trong việc sử dụng người máy. Trên cơ sở đó, bài viết gợi mở một số định hướng chính sách ở Việt Nam trong thời gian tới nhằm tiệm cận với xu thế chung trên thế giới trong lĩnh vực này.

Dẫn đề

Công nghệ người máy (hay công nghệ robot - robotic technologies) đã được nhân loại phát triển và sử dụng cho các mục đích công nghiệp và quân sự từ giữa thế kỷ 20. Kể từ đó, công nghệ này ngày càng được mở rộng phạm vi áp dụng trong các lĩnh vực khác, như giao thông vận tải, chăm sóc sức khỏe, giáo dục, nông nghiệp và cả trong môi trường gia đình hay hoạt động giải trí.

Cùng với sự phát triển nhanh chóng của nó, công nghệ robot đã và đang làm mờ ranh giới giữa con người và máy móc, hay giữa chủ thể con người và các đối tượng công nghệ (human subjects and technological objects)¹⁷⁷. Điều này làm nảy sinh một loạt vấn đề/thách thức về đạo đức chưa từng có trước đó mà nhân loại cần phải giải quyết.

Khác với thời kỳ trước đây, hiện nay, các robot ngày nay đã và đang được tích hợp ngày càng nhiều hơn trí tuệ nhân tạo (AI), khiến cho chúng có được những khả năng cảm nhận của con người, có thể sử dụng ngôn ngữ, tương tác, tự giải quyết các vấn đề, học tập và thậm chí là sáng tạo. Robot không còn là những công cụ thuần túy của con người mà thực sự đã trở thành những “cỗ máy nhận thức”, mà các quyết định và hành động của chúng có thể trở lên không thể đoán trước được¹⁷⁸.

* Bài viết mới ở dạng bản thảo, sẽ còn được chỉnh sửa, bổ sung. Vì vậy, xin độc giả vui lòng chưa trích dẫn.

¹⁷⁷ <http://www.unesco.org/new/en/social-and-human-sciences/themes/comest/robotics-ethics/>

¹⁷⁸ <http://www.unesco.org/new/en/social-and-human-sciences/themes/comest/robotics-ethics/>

Nguy cơ con người không thể kiểm soát được robot là có thật và sự hiện diện ngày càng tăng của các loại robot thế hệ mới trong xã hội đang làm cho thách thức đó trở lên nghiêm trọng hơn.

Bối cảnh trên đặt ra câu hỏi và yêu cầu cấp thiết về việc xác định trách nhiệm giải trình về mặt đạo đức và pháp lý với những hành vi của robot. Đây là vấn đề bước đầu được COMEST phân tích trong bài viết này.

1. Khái niệm, lịch sử phát triển và ứng dụng của người máy trong đời sống của nhân loại

Khái niệm

Thuật ngữ Robot (thường được gọi phiên âm là Rô-bốt hoặc Rô-bô) có nguồn gốc xuất phát từ chữ *Robota* trong tiếng Séc và từ *robot* lần đầu tiên được sử dụng để biểu thị một nhân vật hư cấu trong vở kịch ‘*Các Robot toàn năng của Rosum*’ (tiếng Séc: *Rosumovi Univerzalni Roboti*) của nhà văn Séc Karel Capek vào năm 1920.¹⁷⁹

Hiện tại vẫn còn một vấn đề gây tranh cãi đó là: Một loại máy như thế nào thì đủ tiêu chuẩn để được gọi là một robot? Tuy nhiên, bài viết này sẽ phân tích đặc tính của robot dựa trên báo cáo¹⁸⁰ của COMEST, trong đó đề cập đến bốn đặc tính chủ chốt của rô-bốt, bao gồm: ‘(1) *tự động (mobility)*; (2) *tương tác (interactivity)*; (3) *giao tiếp (communication)* và (4) *tự chủ (autonomy)*’.

Một robot hiện đại điển hình thường thực hiện các hành vi tương tác thông qua các hình thức của *Trí tuệ nhân tạo (Artificial Intelligence-AL)*: như là thay thế nhận thức và trí thông minh của con người bằng hệ thống máy tính, từ đó tạo ra những cỗ máy có thể làm những công việc đòi hỏi một hình thức trí tuệ thông minh, cụ thể như khả năng nhận thức được những thay đổi trong môi trường và thực hiện các chức năng tương ứng để tương tác lại với môi trường. Trí thông minh nhân tạo rất quan trọng đối với sự hình thành chức năng *tự chủ* của robot, vì nó cho phép chúng thực hiện được các nhiệm vụ phức tạp trong một môi trường thường xuyên có sự thay đổi; ví dụ như robot có thể lái xe ô tô và tự thích nghi với các điều kiện trên đường mà không cần phải chịu sự điều khiển trực tiếp hoặc gián tiếp từ xa bởi con người.

Robot thực hiện nhiệm vụ thông qua các thuật toán mà xác lập các quy tắc hoặc hướng dẫn để giải quyết vấn đề. Có thể phân biệt hai loại thuật toán thông qua hai dạng: *thuật toán xác định* – robot thực hiện điều khiển các hành vi đã được lập trình sẵn (deterministic robot); Và *thuật toán ngẫu nhiên* – robot có khả năng học

¹⁷⁹ Ivan Margolius, *The Robot of Prague*, Newsletter, The Friends of Czech Heritage, Số 17, 2017. Trang 3-6. Trích dẫn từ: <https://czechfriends.net/images/RobotsMargoliusJul2017.pdf>

¹⁸⁰ UNESCO-COMEST, *Report of COMEST on Robotics Ethics*, Pháp, 2017.

tập và có nhận thức (cognitive robots). Một robot lập trình sẵn có thể thực hiện những hành vi mang tính phức tạp cao, đòi hỏi ít hoặc không có sự giám sát của con người do nó đã được lập trình sẵn quy trình thực hiện hành vi với mục đích và khả năng nhất định; ví dụ: các loại robot được sử dụng trong lắp ráp máy móc công nghiệp, nông nghiệp... Còn đối với robot có trí tuệ nhân tạo, robot có khả năng ghi nhớ và tự học từ những kinh nghiệm có được trong quá khứ và tự hiệu chỉnh thuật toán để điều chỉnh hành vi của chúng, ví dụ như những robot mang hình dáng người có tên Asimo¹⁸¹ và Sophia¹⁸², do đó hành vi của chúng khó có thể dự đoán trước và điều này gây ra những lo ngại về mặt đạo đức mà đang được thế giới quan tâm.

Lịch sử

Từ thời cổ xưa, con người đã mong muốn tạo ra những vật mang hình thái con người với mục đích phục vụ cho chính mình. Ví dụ như trong câu chuyện thần thoại Hy Lạp kể về người khổng lồ Promethe đúc ra con người từ đất sét và truyền cho họ sự sống; hay như câu chuyện về tên nô lệ khổng lồ Talos được tạo ra bởi Hephaestus (con trai của Thần Zeus và Hera). Trong câu chuyện này, nô lệ Talos được làm bằng đồng và được giao nhiệm vụ bảo vệ đảo Crete; cùng một số câu chuyện tương tự khác.

Đến năm 1920, như đã được đề cập ở trên, nhà biên kịch Karel Capek đã dùng từ robot để gọi một thiết bị lao công do con người tạo ra. Vào những năm 1940, nhà văn viễn tưởng người Nga Issac Asimov đã mô tả robot là một chiếc máy tự động, mang diện mạo của con người và được điều khiển bằng một hệ thần kinh positron, do con người lập trình. Asimov đặt tên cho ngành khoa học nghiên cứu về robot là *Robotics*, thuật ngữ này xuất hiện lần đầu trong tác phẩm *Liar* (1941) của ông. Đồng thời ông đã chỉ ra ba nguyên tắc cơ bản nổi tiếng về robotics trong tác phẩm *Runaround* (1942) của mình, bao gồm: Thứ nhất, robot không được xúc phạm con người và không gây tổn hại đến con người; Thứ hai, hoạt động của robot phải tuân theo các quy tắc do con người đặt ra. Các quy tắc này không được vi phạm nguyên tắc thứ nhất; Thứ ba, một robot cần phải bảo vệ sự sống của mình, nhưng không được vi phạm hai nguyên tắc trước. Các nguyên tắc mà Asimov đã trở thành nền tảng cho việc thiết kế robot sau này.

Trong thế kỉ 19, đã có những robot đầu tiên xuất hiện mang tên *Unimate* và *Shakey*. Robot *Unimate* được tạo ra vào năm 1954 bởi Joseph Engelberger và

¹⁸¹ Robot Asimo được phát triển bởi công ty Honda ở Nhật Bản vào năm 2000, Asimo được thiết kế có trí tuệ thông minh nhân tạo, giúp việc và hỗ trợ cho con người trong những công việc và hoạt động hằng ngày trong gia đình.

¹⁸² Robot Sophia được phát triển bởi công ty Hanson Robotic ở Hồng Kông vào năm 2015. Sophia được thiết kế cử động giống con người và có trí tuệ thông minh nhân tạo. Mục đích chế tạo Sophia là phát minh ra một robot có ý thức, có sự sáng tạo và có khả năng như bất kỳ con người nào để giúp con người trong các vấn đề cuộc sống thường ngày như để phục vụ chăm sóc sức khỏe, điều trị, giáo dục và các ứng dụng dịch vụ khách hàng.

George Devol cho tập đoàn General Motors ở bang New Jersey, Mỹ; Còn robot Shakey được chế tạo bởi Charles Rosen và các cộng sự từ năm 1966 đến 1972 tại bang California, Mỹ và được xem là robot có trí tuệ nhân tạo đầu tiên trên thế giới. Kể từ đó, công nghệ chế tạo robot trên thế giới đã có sự phát triển nhất định, số lượng robot và chất lượng robot được tạo ra ngày càng tăng lên, ví dụ như: robot Wabot-1 (1973), robot Manny (1989), robot P2 (1996), robot Asimo (2000), robot Sophia (2015). Theo thời gian, những Robot sau này dần dần đều có sự thay đổi và phát triển trong khả năng thực hiện các công việc đòi hỏi kỹ thuật cao, có sự tương tác với môi trường và có trí thông minh gần như con người. Điều này chứng minh robot càng ngày càng mang tính ứng dụng cao trong cuộc sống con người.

Ứng dụng

Ứng dụng của robot được thể hiện trong nhiều lĩnh vực của đời sống, có thể kể đến những lĩnh vực quan trọng sau:

Trong công nghiệp, hoạt động công nghiệp trên thế giới hơn một thập kỉ qua đã chứng minh rằng robot có thể thay thế công nhân trong nhiều vị trí việc làm và việc đó mang lại lợi ích to lớn về kinh tế, đồng thời nâng cao hiệu quả sản xuất trong xã hội. Tuy nhiên, việc điều khiển hay làm việc cùng robot đặt ra những yêu cầu mới về quy trình kỹ thuật cũng như các tiêu chuẩn an toàn mà đòi hỏi phải có sự huấn luyện. Đồng thời, ở cấp độ quốc tế, việc sử dụng robot một cách phổ biến trong hoạt động sản xuất cũng tạo ra những thử thách mới trong quan hệ quốc tế, bởi nó có thể đào sâu hơn khoảng cách giữa những quốc gia phát triển và đang phát triển.

Trong quân sự, robot trong lĩnh vực này thường được thể hiện dưới dạng thiết bị *máy bay không người lái (drones)* và các người máy chiến đấu hoặc vận chuyển hay gỡ mìn. Những thiết bị này được điều khiển từ xa bởi con người với các mục đích khác nhau như: do thám, tấn công, bảo vệ, hỗ trợ ...Hiệu quả của hoạt động quân sự được nâng cao, tổn thất về nhân mạng của binh sĩ được giảm thiểu khi robot có thể thay các quân nhân thực hiện những nhiệm vụ mang tính nguy hiểm và tốn nhiều công sức. Tuy nhiên, việc sử dụng robot trong hoạt động quân sự cũng tạo ra những lo ngại về vấn đề quyền con người và về việc tuân thủ các nguyên tắc của Luật Nhân đạo quốc tế (International Humanitarian Law-IHL). Sự quan ngại này càng ngày càng trở thành vấn đề lớn khi các hệ thống vũ khí tự động đang được phát triển từng ngày.

Trong giao thông, *phương tiện tự hành (the autonomous vehicle-AV)* đã sẵn sàng để được áp dụng vào thực tế một cách phổ biến. Phương tiện giao thông tự hành góp phần làm giảm tỷ lệ tai nạn giao thông và nâng cao hiệu quả giao thông trong xã hội như giảm tắc đường, giảm khí thải,... Tuy nhiên, điều này cũng tạo ra những vấn đề đạo đức liên quan đến quá trình *tự hành* của phương tiện. Làm thế nào để một chiếc xe tự hành được lập trình có thể đối phó với các tình huống tai nạn

bất ngờ? Và các vấn đề phát sinh từ tình huống này được điều chỉnh bởi pháp luật hay bởi các quy tắc ứng xử chung?

Trong công tác bảo vệ sức khỏe và đảm bảo an sinh xã hội, việc sử dụng robot trong y tế nói chung, cụ thể là trong quá trình thực hiện các công việc khám chữa bệnh như mổ nội soi, phục hồi chức năng, chăm sóc sức khỏe người già,... nói riêng ngày càng trở nên phổ biến. Việc sử dụng robot trong y tế mang lại độ chính xác cao trong các thao tác mà robot thực hiện, mặc dù ở thời điểm ban đầu việc áp dụng được điều này cũng đặt ra yêu cầu về chi phí đầu tư cao hơn bình thường đối với các cơ sở y tế. Thêm vào đó, bên cạnh những lợi ích mà robot mang lại, cũng có những câu hỏi về phương diện đạo đức về tính hợp lý khi sử dụng những công nghệ này vào công việc chăm sóc con người, như là: Liệu robot có an toàn cho người sử dụng? Sự ảnh hưởng của nó tác động lên nhận thức của người sử dụng như thế nào khi những người sử dụng công nghệ robot đó là người đang bị bệnh, già cả hoặc là đối tượng cần ưu tiên đặc biệt? Hay có tồn tại hay không mối quan hệ tình dục giữa robot và con người?... Những câu hỏi đó đặt ra yêu cầu việc chế tạo và áp dụng công nghệ robot cần phải cân nhắc một cách thấu đáo về sự ảnh hưởng của nó tới con người cả trên phương diện đạo đức và pháp luật.

Đối với lĩnh vực giáo dục và đời sống, robot có thể thực hiện hỗ trợ cá nhân người học trong các hoạt động giáo dục như phục vụ công tác nghiên cứu, giảng dạy; còn đối với đời sống, robot có thể giúp con người trong các hoạt động sinh hoạt hằng ngày như là lau dọn nhà cửa, là quần áo, nấu ăn,... Việc xuất hiện robot trong gia đình giúp cho cuộc sống con người trở nên tiện nghi hơn, đặc biệt là đối với những gia đình có người bệnh, người già hay những người bị suy giảm khả năng vận động thể chất. Tuy nhiên, việc cho phép robot sống chung với con người như những thành viên trong gia đình cũng tạo ra những quan ngại về vấn đề đạo đức. Sẽ ra sao nếu robot bị lỗi về mặt công nghệ? Nó có thể làm cho robot tiếp nhận và xử lý thông tin lỗi, điều này có thể dẫn tới những hậu quả mà trên thực tế con người có thể chưa lường trước được.

Và cuối cùng, trong lĩnh vực nông nghiệp, robot đang dần được áp dụng nhiều trong sản xuất lương thực, ví dụ như việc sử dụng máy bay không người lái để thu thập và phân tích thông tin mùa màng, sử dụng robot để phun thuốc bảo vệ thực vật,... Việc sử dụng công nghệ robot mang lại những lợi ích quan trọng như là tiết kiệm sức lao động cho con người, nâng cao hiệu quả về năng suất và chất lượng sản phẩm đồng đều hơn,...

Tóm lại, phát minh robot là một phát minh rất quan trọng đối với con người. Robot mang lại nhiều lợi ích về kinh tế, nâng cao chất lượng cuộc sống và cùng con người chạm tới những tinh hoa tri thức mới của nhân loại. Dù vậy, đồng thời với những lợi ích to lớn đó, robot cũng tạo ra những thách thức to lớn đối với con người

trong việc kiểm soát, sử dụng và quản lý những vấn đề phát sinh xung quanh việc áp dụng công nghệ này vào đời sống như đã đề cập ở trên. Mục dưới đây sẽ đi sâu tìm hiểu về hai vấn đề nổi bật liên quan đến mà robot đang được toàn thế giới quan tâm nhất hiện nay, đó là vấn đề về đạo đức và vấn đề về pháp lý.

2. Những thách thức về pháp lý và đạo đức đang đặt ra với việc chế tạo và sử dụng người máy

Ngành chế tạo robot và robot đã tạo ra những thử thách mới, chưa từng có tiền lệ liên quan đến những vấn đề về đạo đức về pháp lý. Với sự phức tạp trong thiết kế, chế tạo và lập trình robot, vấn đề trọng tâm về đạo đức ở đây được tạo ra từ là khả năng *truy xuất nguồn gốc (traceability)* của robot; khả năng truy xuất nguồn gốc là khả năng giúp cho robot có thể theo dõi được những nguyên nhân tạo ra tất cả những hành động trong quá khứ. Với những robots có khả năng tự quyết, khả năng đưa ra quyết định và khả năng học hỏi thì việc truy xuất nguồn gốc này chính là mối lo ngại, bởi những robot này không chỉ được lập trình để làm những công việc cụ thể, mà nó còn có thể tự học và tự phát triển bản thân trong sự tương tác với môi trường; Chính điều này tạo ra một lỗ hổng lớn trong khung pháp lý và đạo đức hiện hành đối với khả năng truy xuất nguồn gốc của robot.

Thách thức pháp lý

Vào năm 2017, được sự ủy quyền của Hội đồng Nghị viện của Ủy ban Châu Âu (PACE¹⁸³) Viện Rathenau đã công bố bản báo cáo có tên là ‘Quyền con người trong thời đại robot¹⁸⁴’. Bản báo cáo chỉ ra những ảnh hưởng xấu do công nghệ robot có thể tạo ra đối với các phương diện liên quan như: quyền con người, sự tôn trọng cuộc sống riêng tư, nhân phẩm, sở hữu, an ninh, trách nhiệm, tự do biểu đạt, chống phân biệt đối xử, tiếp cận công lý và tiếp cận xét xử công bằng. Bản báo cáo đã gợi ý bổ sung hai quyền con người mới: quyền không bị sắp đặt, phân tích và chỉ đạo (liên quan đến các hành vi lạm dụng trí tuệ nhân tạo, thu thập dữ liệu và Internet kết nối vạn vật); và quyền tiếp xúc với con người một cách có ý nghĩa (liên quan đến khả năng bị lạm dụng, có chủ ý hoặc vô ý, của robot chăm sóc).

Nhà nghiên cứu tên là Leenes và một nhóm tác giả đã chỉ ra bốn tình huống khó xử trong cuốn sách *Những thách thức pháp lý đối với công nghệ robot: Một số hướng dẫn giải quyết các vấn đề về pháp lý và đạo đức*¹⁸⁵, bao gồm sự tiến thoái lưỡng nan trong: (1) việc chạy đua với những tiên bộ của công nghệ; (2) việc tạo ra sự cân bằng giữa thúc đẩy, đổi mới và bảo vệ quyền và giá trị cơ bản của con người; (3) việc khẳng định các chuẩn mực xã hội hiện có hoặc làm lu mờ những chuẩn

¹⁸³ The Parliament Assembly of the Council of Europe, viết tắt là PACE.

¹⁸⁴ Bản báo cáo đầy đủ: <https://www.rathenau.nl/sites/default/files/2018-02/Human%20Rights%20in%20the%20Robot%20Age-Rathenau%20Instituut-2017.pdf>

¹⁸⁵ R. Leenes, E. Palmerini, B. Koops, P. Salvini, F. Lucivero (2017), Regulatory challenges of robotics: some guidelines for addressing legal and ethical issues, *Law, Innovation and Technology*, page 1-44.

mục đó theo hướng mới; (4) việc cân bằng giữa tính hiệu quả và tính hợp lý trong các quy định về công nghệ kỹ thuật.

Thách thức đạo đức

Công nghệ robot làm mờ đi ranh giới giữa chủ thể là con người với các đối tượng công nghệ. Câu hỏi được đặt ra là ai sẽ là người phải chịu *trách nhiệm (responsibility)* về đạo đức và pháp lý cho những *hành vi* mà robot thực hiện. Dường như trách nhiệm sẽ được “chia sẻ” giữa người thiết kế, kỹ sư, người lập trình, người sản xuất, nhà đầu tư, người bán và người mua. Không ai trong những nhóm người trên được mặc định là người chịu trách nhiệm cho những hành động cuối cùng của robot, bởi nếu xem mỗi người đều có một phần trách nhiệm trong tổng hòa một khối trách nhiệm thì sẽ không có ai là người đứng ra chịu trách nhiệm hoàn toàn cho hành vi của robot.

Vấn đề thứ hai là vấn đề về sự *tự chủ (agency)* với *địa vị đạo đức (moral status)*. Robot có khả năng tự đưa ra quyết định dựa trên chính tư duy của nó mà không cần phải được lập trình sẵn từ trước bởi con người. Mặc dù khả năng tự quyết giữa robot và con người rõ ràng là có sự khác biệt, vấn đề chính được đặt ra ở đây là vậy thì sự tự chủ của robot làm thay đổi những hành vi của con người như thế nào. Liệu robot có trở thành một chủ thể mang giá trị đạo đức, vượt lên cả những giá trị đơn thuần về một cái máy đã được lập trình sẵn hay không? Nếu như vậy thì những robot đó có được hưởng sự tôn trọng và bảo vệ hay không? Và chúng có được trao quyền trên phương diện đạo đức hay không?

Khía cạnh cuối cùng nhìn từ phương diện đạo đức đó là sự tác động của công nghệ robot đối với các *khuôn khổ đạo đức (moral framework)*. Chúng không chỉ có tác động tới xã hội về mặt đạo đức, mà nó còn ảnh hưởng đến các khuôn khổ đạo đức mà chúng ta có thể dùng để đánh giá chúng. Ví dụ, một robot làm nhiệm vụ chăm sóc (care robot) có thể thay đổi giá trị của con người trong nhiệm vụ chăm sóc. Hay trong giảng dạy, robot (teaching robot) có thể làm thay đổi những tiêu chí đối với một nền giáo dục tốt. Một robot tình dục (sex robot) có thể tác động đến những giá trị tình cảm và những mối quan hệ mật thiết mà con người đang có. Để đối phó với những thay đổi đó, con người phải có trách nhiệm trong việc dự đoán và thử nghiệm, bám sát những tác động mà công nghệ robot tạo ra lên các khung giá trị; thiết lập các chương trình thử nghiệm nhỏ để đánh giá tác động của robot, từ đó áp dụng vào việc thiết kế, chế tạo robot và hoạch định những chính sách hợp lý trong việc sản xuất và áp dụng công nghệ robot vào thực tiễn.

3. Giải pháp cho những thách thức đạo đức và pháp lý đang đặt ra với việc chế tạo và sử dụng robot

Để đưa ra những giải pháp phù hợp với những vấn đề đạo đức của công nghệ

robot, cần phải phân biệt giữa *robot lập trình* (deterministic robots) và *robot nhận thức* (cognitive robots). Đối với robot lập trình, hành vi của robot đã được lập trình sẵn và được dự báo trước, vì vậy việc thiết lập những quy định để quản lý hành vi của robot lập trình rõ ràng và dễ dàng hơn so với robot nhận thức. Bởi vì đối với robot nhận thức, không thể đoán trước được hành vi của robot, vì thế việc thiết lập những quy định để quản lý hành vi đối với robot nhận thức chính là vấn đề trọng tâm.

Dựa trên những yếu tố này, COMEST đã đề xuất những giải pháp tương ứng với từng loại robot. Đối với robot lập trình, COMEST gợi ý tập trung vào các công cụ pháp lý để quản lý việc sử dụng chúng; còn đối với robot nhận thức, ngoài việc sử dụng các công cụ pháp lý thì cần phải cân nhắc thêm việc ban hành các quy tắc thực hành, hướng dẫn đạo đức cho cả người sản xuất và người sử dụng robot. Bên cạnh đó, còn một lưu ý đối với việc sử dụng công nghệ máy móc tự động (ví dụ như xe không người lái, máy bay không người lái và một số vũ khí tự động khác,...), chúng ta cần cân nhắc xem nên cho phép máy móc tự chủ trong một phạm vi nhất định để tránh tạo ra những mối nguy hại khó lường, con người vẫn phải duy trì sự kiểm soát chính đối với những loại máy móc đó. Bản Báo cáo của COMEST đã chỉ ra những nguyên tắc đạo đức và những giá trị cần phải lưu ý¹⁸⁶, đồng thời đưa ra những giải pháp cho những thách thức về đạo đức và pháp lý, cụ thể bao gồm:

i. Thứ nhất, xây dựng bộ tiêu chuẩn đạo đức (codes of ethics) cho người máy phải được xem xét từ cả hai cấp độ là quốc gia và quốc tế, nó phải được phát triển, áp dụng và cập nhật theo nhiều cách khác nhau, đồng thời nghiên cứu sự ảnh hưởng của nó tới cuộc sống và môi trường sống của con người (năng lượng, chất thải công nghệ, sinh trắc học,...). Thứ hai, các chuyên ngành đặc biệt có khả năng sử dụng robot nhiều trong ứng dụng thực tiễn, từ kỹ sư điện và trí tuệ nhân tạo đến y học, động vật học, tâm lý học, vật lý,... cần phải tự xây dựng những bộ tiêu chuẩn đạo đức cụ thể đối với robot trong lĩnh vực mà họ đang làm. Thứ ba, tất cả những bộ tiêu chuẩn về đạo đức (codes of ethics) và bộ tiêu chuẩn về hành vi (codes of conduct) cùng với những tài liệu liên quan khác phải trở thành một trong những nội dung học tập chính thức đối với những ngành học đào tạo các chuyên gia về công nghệ robot.

ii. Yếu tố đạo đức phải trở thành một phần trong quá trình thiết kế robot. Và quá trình đó cần được xây dựng dựa trên phương pháp *Thiết kế giá trị nhạy cảm* (*Value Sensitive Design - VSD*)¹⁸⁷. VSD là phương pháp tiếp cận dựa trên lý thuyết

¹⁸⁶ Những giá trị đó bao gồm: *Phẩm giá của con người (Human dignity)*, *Giá trị của sự tự chủ (Value of autonomy)*, *Giá trị của sự riêng tư (Value of privacy)*, *Nguyên tắc không gây hại ('Do not harm' principle)*, *Nguyên tắc trách nhiệm (Vrinciple of responsibility)*, *Giá trị của sự lương thiện (Value of beneficence)*, *Giá trị của công lý (Value of justice)*.

¹⁸⁷ Được phát triển bởi Batya Friedman và Peter Kahn tại trường Đại học Washington trong giai đoạn những năm 1980 đến 1990.

về thiết kế công nghệ lấy con người làm giá trị căn cứ cho toàn bộ quy trình thiết kế. VSD sinh ra để giải quyết những vấn đề sinh ra giữa quy trình thiết kế hệ thống thông tin và khả năng tương tác của robot với con người bằng cách nhấn mạnh các giá trị đạo đức của các bên sở hữu (trực tiếp và gián tiếp). Thiết kế được xây dựng và phát triển thông qua việc điều tra ba phần (a tripartite investigation) bao gồm: *khái niệm (conceptual)*, *kinh nghiệm (empirical)* và *kỹ thuật (technological)*. Việc thực hiện các bước điều tra này giúp cho các nhà thiết kế có thể điều chỉnh thiết kế liên tục dựa trên những dữ liệu đã thu thập được. Việc điều tra về khái niệm giúp cho các bên sở hữu hiểu hơn về giá trị và giải quyết những mâu thuẫn có thể xuất hiện trong quá trình sử dụng công nghệ giữa các bên. Còn đối với việc điều tra kinh nghiệm, quy trình sẽ được tiến hành dựa trên những nghiên cứu (cả về số lượng lẫn chất lượng) trong thực tế, để từ đó các nhà thiết kế có thể dựa trên những thông tin đã thu thập để thiết kế ra một sản phẩm công nghệ phù hợp với nhu cầu, mục đích cũng như khả năng của người dùng. Và bước cuối cùng là điều tra công nghệ, bước này xem xét tính ứng dụng khi con người sử dụng công nghệ đó trên thực tế như thế nào và sử dụng những kết quả hữu ích thu thập được từ hai bước trước để từ đó kết hợp các yếu tố lại với nhau để tạo ra một sản phẩm công nghệ cuối cùng.

iii. Để nâng cao trách nhiệm khi giới thiệu robot đến với xã hội, yêu cầu công nghệ mới đó phải được nghiên cứu giới thiệu cẩn thận, trong một giới hạn nhất định, đảm bảo về mặt kiểm soát và phải được công khai. Những giá trị thu thập được thông qua việc ra mắt và thử nghiệm bởi người dùng sẽ được dùng để thay đổi các thiết kế của robot cho phù hợp, từ đó xây dựng bộ quy tắc sử dụng nghiêm ngặt đối với người dùng.

iv. Tổ chức các buổi thảo luận công khai về việc sử dụng công nghệ robot trong thực tế dựa trên những thước đo đa dạng trong xã hội và cuộc sống hằng ngày, ví dụ như khả năng tác động đến môi trường của việc tái chế phế phẩm công nghệ, để nâng cao nhận thức cho các nhà nghiên cứu chế tạo robot và các nhà nghiên cứu luật pháp.

v. Các quốc gia, các tổ chức chuyên ngành, các viện nghiên cứu giáo dục phải cân nhắc tới việc áp dụng robot vào công việc, bởi nó sẽ tạo ra nguy cơ mất việc làm của con người. Bên cạnh đó nó còn tạo ra những cơ hội việc làm với những yêu cầu về năng lực mới, điều này yêu cầu các quốc gia, các tổ chức phải có sự thay đổi trong việc đào tạo nguồn nhân lực có đủ khả năng đáp ứng trong môi trường công việc mới.

vi. Đối với các phương tiện tự hành (autonomous vehicles) - có khả năng tự điều chỉnh hoạt động dựa vào khả năng tự nhận thức của hệ điều hành, trách nhiệm của hành vi phải thuộc về con người.

vii. Việc sử dụng máy bay không người lái trong lực lượng vũ trang tạo ra

những vấn đề về đạo đức và pháp lý. Bên cạnh vấn đề về pháp lý liên quan đến Luật Nhân đạo quốc tế (IHL), luôn có một nguyên tắc đạo đức chống lại việc chế tạo những robot với mục đích giết người hoặc dùng trong lực lượng vũ trang hoặc các hoạt động chống đối. COMEST đề nghị các quốc gia cần cân nhắc vấn đề này một cách kỹ lưỡng.

viii. Còn đối với vũ khí tự hành, COMEST nhấn mạnh các quốc gia cần nắm giữ điều hành và kiểm soát việc sử dụng vì lý do pháp lý, đạo đức và các lý do quân sự liên quan. Báo cáo năm 2016 của Hội Chữ Thập đỏ quốc tế (ICRC)¹⁸⁸ yêu cầu ‘xác định cấp độ quản lý của con người đối với các vũ khí tự hành để đảm bảo việc vận hành vũ khí phù hợp với các điều kiện pháp lý và phù hợp với các yêu cầu về đạo đức và xã hội.

ix. Với các loại máy bay tự hành dùng trong các mục đích khác ngoài mục đích quân sự, các quốc gia phải có những chính sách sử dụng cụ thể. Cần xây dựng điều khoản sử dụng riêng cho các cơ quan cảnh sát cho việc sử dụng máy bay tự hành với mục đích đảm bảo an ninh, thu thập chứng cứ phạm tội,... và các điều khoản này phải được ban hành bởi các tổ chức đại diện chứ không phải bởi chính cơ quan cảnh sát nhằm mục đích đảm bảo tính khách quan, minh bạch, tránh lạm quyền.

x. Bên cạnh đó, đối với máy bay tự hành cá nhân phải được cấp phép sử dụng và chỉ được sử dụng trong một số phạm vi nhất định dưới sự quản lý của nhà nước.

xi. Phải có sự tuy tính cụ thể về vấn đề giới tính đối với tất cả các loại robot nói chung và một số loại robot cụ thể như robot đồ chơi, robot tình dục và những loại robot làm việc thay thế công việc của con người.

xii. Tương tự với các công nghệ khác, vấn đề về bảo vệ môi trường và phân tích quy trình tái chế cũng như xử lý rác thải công nghệ cũng cần phải xây dựng một quy trình cụ thể. Các nhà thiết kế và các tổ chức sản xuất công nghệ robot cần phải nghiên cứu sử dụng các chất liệu sản xuất thân thiện với môi trường và có cơ chế tái sử dụng hoặc tiêu hủy phù hợp, không gây hại cho môi trường sống của con người.

Tài liệu tham khảo:

- (1) UNESCO (2017) REPORT OF COMEST ON ROBOTICS ETHICS
- (2) UNESCO (2018) HUMANDECISIONSTHOUGHTS ON AI
- (3) Mathias Risse (2018), Human Rights and Artificial Intelligence, An Urgently Needed Agenda, Carr Center for Human Rights Policy

¹⁸⁸ Hội Chữ Thập đỏ quốc tế, Báo cáo thường niên năm 2016, <https://www.icrc.org/data/files/annual-report-2016/ICRC-2016-annual-report.pdf>