Một số vấn đề an ninh thông tin trọng yếu trong kỷ nguyên AI (Phần 2)

3. Trí tuệ nhân tạo càng mạnh, yếu tố con người càng quan trọng

Trong thập kỷ qua, AI đã phát triển nhanh chóng và đạt được nhiều bước đột phá khiến cho chúng vượt qua khả năng của con người ở nhiều lĩnh vực và nhiệm vụ khác nhau (Henshall, 2023), kể cả trong một số khía cạnh của việc đảm bảo an ninh thông tin. Mặc dù năng lực tính toán của AI ngày càng tăng, các công năng hay sản phẩm được tạo ra bởi AI vẫn là do con người định hướng và quyết định (thông qua quá trình đào tạo mô hình và ra lệnh trực tiếp cho AI). Nói cách khác, AI càng phát triển nhanh chóng, chúng sẽ càng khuếch đại năng lực và sức mạnh của người sử dụng chúng (hay người mà chúng phục vụ). Điều này có khả năng dẫn đến hai vấn đề dưới đây.

Đầu tiên, sức mạnh mà AI mang lại có thể giúp cho người sử dụng có nhiều lựa chọn hơn. Những việc trước đây họ không làm được do giới hạn về kiến thức, khả năng, sức lực, và thời gian, thì giờ đây đã có AI bổ trợ. Con người có thể chuyên tâm vào nhiệm vụ tăng cường vốn tri thức, học cách điều khiển AI một cách hiệu quả. Thế nhưng, tại sao điều này lại là vấn đề? Vấn đề nằm ở chỗ họ có thể lựa chọn sử dụng AI cho mục đích xấu, như thực hiện các hành vi tấn công lừa đảo, tấn công tống tiền, tạo các mã độc, v.v.. Điều này góp phần tạo ra nhiều rủi ro về an ninh thông tin trong tương lai, khi mà một người ban đầu thiếu chuyên môn về bảo mật lại có thể trở thành một *hacker* mũ đen nhanh chóng nếu như họ biết cách điều khiển AI cho mục tiêu tấn công mạng. Hậu quả sẽ còn tệ hơn nếu như những *hacker* này đột ngột xuất hiện từ trong nội bộ của một tổ chức hay công ty.

Bên cạnh đó, sức mạnh AI cung cấp cho người sử dụng càng lớn, thì tác động của họ lên mọi thứ xung quanh cũng sẽ trở nên lớn hơn. Như đã trình bày ở trên, cho dù AI có năng lực bảo mật cực kỳ mạnh mẽ thì rủi ro từ các lỗ hổng do lỗi con người tạo ra sẽ luôn tồn tại. Nếu con người mắc lỗi khi vận hành AI, thì hậu quả tạo ra từ lỗi đấy sẽ còn nghiêm trọng hơn gấp nhiều lần. Ví dụ, do lỗi con người vô tình hay cố ý mà khiến dữ liệu huấn luyện của mô hình phân loại của AI trở nên kém chính xác, rủi ro bảo mật tạo ra bởi sự kém chính xác đấy sẽ tăng lên gấp nhiều lần do nguyên lý tự động vận hành liên tục của AI trong một thời gian dài và con người khó có khả năng can thiệp vào quá trình đấy (nếu có thì cũng rất tốn kém). Nếu AI tiếp tục được phát triển với tốc độ vượt tất cả các dự đoán như hiện nay, việc nó được tích hợp vào trong mọi mặt cuộc sống hàng ngày của mỗi cá nhân, doanh nghiệp, và quốc gia cả về độ sâu và rộng sẽ có khả năng sớm xảy ra (Henshall, 2023; Stacey & Milmo, 2023). Đặc biệt, với sự xuất hiện của các hệ thống thực ảo (*cyberphysical systems*), như lưới điện thông minh, hệ thống ô tô tự động, giám sát y tế, hệ thống điều khiển công nghiệp, hệ thống robot, v.v.., khoảng cách giữa thế giới thực và thế giới ảo sẽ ngày càng bị thu hẹp. Như thế, các tác động trong thế giới ảo cũng sẽ có khả năng ảnh hưởng trực tiếp đến thế giới thực. Chỉ một lỗi sai sót ban đầu tưởng tư không quá nghiêm trọng trong quá trình vận hành tự động của AI do con người gây ra (đặc biệt là trong vấn đề bảo mật) thì cũng có thể dẫn đến những hậu quả không lường trước được.

Cả hai vấn đề trên đều đến từ việc thay đổi cấu trúc sức mạnh trong xã hội (Suleyman, 2023). Sức mạnh ở đây có thể hiểu theo "khả năng tạo ra hoặc ngăn chặn sự thay đổi" (Green, 1998). Vì thế, để đạt được an ninh thông tin trong kỷ nguyên AI, chúng ta cần nhận thức được rõ hơn yếu tố con người và xã hội trong quá trình phát triển và vận hành AI. Đặc biệt là các vấn đề liên quan đến sự tự do cá nhân, sức mạnh, trách nhiệm, vai trò của tổ chức quản lý và điều tiết của nhà nước, và trách nhiệm của các công ty công nghệ.

4. Sự chuyển dịch cấu trúc xã hội và ý niệm tự do trong kỷ nguyên AI

Để hiểu rõ hơn vai trò của yếu tố con người và cấu trúc xã hội đối với an ninh thông tin trong kỷ nguyên AI, chúng ta cần xem xét từ gốc độ thành phần cơ bản nhất trong cấu trúc xã hội: suy nghĩ, quyết định, và hành vi cá nhân. Vì bối cảnh cấu trúc xã hội đang có sự chuyển dịch từ giai đoạn không có AI sang giai đoạn AI được tích hợp vào mọi mặt trong cuộc sống, nên lý thuyết *Mindsponge* được sử dụng để giúp làm rõ vấn đề nhờ khả năng lý giải linh hoạt của hệ lý thuyết xoay quanh tương tác với thông tin.

Lý thuyết *Mindsponge* cho rằng mỗi cá nhân là một hệ thống lưu trữ và xử lý thông tin sinh học có khả năng đưa ra quyết định và hành vi nhằm tương tác với môi trường xung quanh (bao gồm cả môi trường tự nhiên, xã hội, văn hóa, chính trị, và công nghệ) (Vuong, 2023). Sự vận hành của hệ thống xử lý thông tin bao gồm quá trình đánh giá chi phí và lợi ích với mục tiêu tối ưu hóa các lợi ích nhận thức được và giảm thiểu các chi phí cảm nhận được (Vuong *et al.*, 2022). Các đánh giá chi phí và lợi ích này bị ảnh hưởng bởi các mục tiêu và ưu tiên của hệ thống, cũng như tuân theo nguyên tắc tiết kiệm năng lượng (nguyên lý bảo tồn năng lượng của sinh vật). Mục đích hay ưu tiên cơ bản nhất của hệ thống là đảm bảo kéo dài sự tồn tại của hệ thống bằng cách này hay cách khác, bao gồm sinh tồn, phát triển, và sinh sản (Vuong, 2023). Thông qua lăng kính xử lý thông tin *Mindsponge*, ta có thể hình dung nhận thức về sức mạnh (nhận thức về khả năng tạo ra hoặc ngăn chặn sự thay đổi/ảnh hưởng) của mỗi các nhân là sản phẩm của quá trình xử lý thông tin và tương tác với môi trường xung quanh của cá nhân đấy. Các nhận thức về sức mạnh tồn tại giới hạn do các quan sát từ thực tế khách quan và các đánh giá chủ quan của bản thân liên quan đến kiến thức, năng lực, sức lực, tài sản, địa vị xã hội, và thời gian (Nguyen *et al.*, 2023).

Khi AI bắt đầu xuất hiện và được ứng dụng trong xã hội, các cá nhân sẽ dần quan sát được các lợi ích mà AI mang lại từ thực tế khách quan mà lựa chọn sử dụng chúng. Thông qua quá trình tương tác, trao đổi thông tin qua lại với AI, các nhận thức ban đầu (lúc trước khi biết đến AI) sẽ dần bị biến đổi. Những nhận thức đấy bao gồm cả những nhận thức về giới hạn kiến thức, kỹ năng, sức lực, và

thời gian của bản thân. Dựa vào AI, cá nhân đã có khả năng thực hiện những việc mà trước đây bản thân họ không thể làm được hoặc không nghĩ tới do bị giới hạn bởi kiến thức, kỹ năng, sức lực, và thời gian khách quan. Ví dụ, một người trước đây chưa từng biết vẽ hoặc chưa từng biết về lập trình máy tính thì giờ đây đã có thể dễ dàng tạo ra những bức ảnh đầy tính nghệ thuật hoặc các đoạn mã lập trình máy tính nhờ vận dụng AI. Chưa kể, AI Deepfake hiện nay còn cho họ sức mạnh để tạo ra các nội dung giả mạo trông như thật, như những hình ảnh, video giả mạo khuôn mặt, giọng nói người khác, một cách nhanh chóng và dễ dàng.

Khi sức mạnh khách quan (hay khả năng tạo ra hoặc ngăn chặn sự thay đổi) (Green, 1998) của cá nhân được gia tăng nhanh chóng dựa trên sự giúp đỡ của AI, điều đấy đồng nghĩa tập hợp các hành động có thể xảy ra với cá nhân đấy cũng đang tăng lên. Hay nói cách khác, sự tự do tổng thể của cá nhân gia tăng (Pansardi, 2012). Nếu như không có các cơ chế quản lý chính xác, điều này sẽ làm cho các rủi ro về an ninh thông tin gia tăng đáng kể (như giải thích ở Phần 3).

Trên thực tế, sự tự do tổng thể của một cá nhân trong xã hội bị hạn chế bởi các hệ thống xã hội. Mặc dù cá nhân đó có khả năng thực hiện một tập hợp các hành động, nhưng do sự ngăn cản hoặc tác động của các cá nhân hoặc nhiều cá nhân khác trong xã hội (thông qua luật lệ, văn hóa, hoặc đạo đức), nên họ không thực hiện một số hành động trong số các hành động họ có khả năng thực hiện (Kramer, 2008; Pansardi, 2012). Theo góc nhìn của lý thuyết *Mindsponge*, thì cá nhân đó có khả năng thực hiện các hành động một cách khách quan, nhưng lại không thực hiện do các đánh giá chủ quan của họ về mặt chi phí (tạo ra bởi các cá nhân khác thông qua luật lệ, văn hóa, hoặc đạo đức) (Nguyen *et al.*, 2023; Vuong, 2023).

Hiện nay, khi sự xuất hiện của AI trong đời sống còn quá mới và tương lai phát triển của nó vẫn chưa thể xác định chính xác, các chuẩn mực văn hóa hay đạo đức khi sử dụng AI vẫn còn tồn tại nhiều tranh cãi và chưa thể định hình. Trong khi đó, đạo luật về quản lý trí tuệ nhân tạo đầu tiên trên thế giới cũng chỉ mới vừa được phê chuẩn bởi Liên minh Châu Âu vào ngày 13/3/2024 (Liaropoulos, 2020). Vì thế, chúng ta cần hiểu rõ hơn về sự chuyển dịch của cấu trúc xã hội, do sự thay đổi sức mạnh và mức độ tự do mà AI mang lại trên diện rộng, để có thể triển khai các cơ chế kiểm soát sức mạnh và điều tiết tự do một cách phù hợp.

Thông thường các cơ chế kiểm soát sức mạnh này sẽ được quản lý và triển khai bởi nhà nước. Nhưng tại sao các cá nhân lại đồng ý mất đi một phần tự do, hay nói cách khác, để nhà nước giới hạn sức mạnh của bản thân?

Điều này có thể lý giải thông qua Lý thuyết khế ước xã hội (*social contract*) (Hobbes, 1894; Locke, 1967; Rousseau, 2016). Lý thuyết này cho rằng các cá nhân cùng nhau hình thành một tập thể có cơ quan thẩm quyền (ví dụ như nhà nước) và từ bỏ một phần (thậm chí toàn bộ trong một số điều kiện đặc biệt nghiêm trọng với sinh tồn của tập thể xã hội) của quyền tự do của họ cho cơ quan này quản lý và thực hiện trách nhiệm của họ như được mô tả trong pháp luật. Đổi lại, cơ quan thẩm quyền phải mang lại cho những cá nhân trong tập thể đấy các lợi ích của trật tự chính trị và xã hội, ví dụ như sự ổn định, an toàn của bản thân, và tài sản (Bierens *et al.*, 2017; Boucher & Kelly, 2003; Liaropoulos, 2020). Với sự xuất hiện của các công ty tư nhân vào Thế kỷ 20, một bên thứ ba đã được thêm vào hợp đồng xã hội (Liaropoulos, 2020). Họ được xem là một thực thể pháp lý trong một quốc gia có mục tiêu tối ưu hóa lợi nhuận, thông qua đấy tạo ra động lực phát triển cho xã hội (tạo ra việc làm, của cải, thúc đẩy đổi mới sáng tạo, v.v..). Tuy nhiên, các công ty tư nhân không được gây hại cho khế ước xã hội giữa các công dân và nhà nước, vì thế nhà nước được quyền áp dụng các luật lệ và quy định riêng biệt lên các công ty tư nhân trong khi vẫn xem xét các yếu tố khác như lực cạnh tranh thị trường giữa các công ty và người dân. Nếu công ty trở thành độc quyền, hoặc tựa độc quyền, các luật lệ và quy định của chính phủ cần được nhà nước tăng cường để kiểm soát (Bierens *et al.*, 2017; Liaropoulos, 2020).

Tuy nhiên, cuộc cách mạng thông tin, và gần đây nhất là sự xuất hiện của AI, khiến cho thế giới trở

nên siêu kết nối và thay đổi cấu trúc sức mạnh trong xã hội thông qua việc tăng cường sức mạnh cho những ai có khả năng tiếp cận và vận dụng AI. Việc này dẫn đến câu hỏi là liệu chính quyền của các quốc gia có đủ sức kiểm soát và giữ được trật tự chính trị và ổn định xã hội hay không? Nếu có thì ở mức độ và phạm vi nào, vì không gian ảo là không gian gần như không có biên giới? Ở chiều ngược lại, khi các cá nhân đã nắm giữ được sức mạnh không tưởng của AI, nghĩa là mức độ tự do tổng thể của họ tăng lên trên diện rộng, liệu họ có còn sẵn sàng đánh đổi sự tự do đấy để có được sự ổn định xã hội như trước hay không? Nếu có thì họ sẵn sàng đánh đổi bao nhiêu tự do để tối ưu hóa lợi ích mà họ nhận thức được? Điều gì xảy ra nếu các quy tắc ứng xử cộng đồng chia sẻ nền tảng thông tin trở nên xung đột với các khế ước xã hội đặc hữu, gây xói mòn hệ thống đạo đức, và trở thành siêu quy tắc có khả năng gây nên xung đột *siêu văn hóa* diện rộng?

Ngoài ra, hiện nay các chính quyền đều chưa thể có ngay công cụ hữu hiệu để hạn chế sức mạnh của người sử dụng được nhân lên bội phần bởi AI và các công nghệ thông tin khác, vì các nhà cung cấp chính các dịch vụ này là các tập đoàn công nghệ đa quốc gia hàng đầu thế giới, như Microsoft, Meta, Google, v.v.. Sâu xa hơn, các tập đoàn này đều đang nắm giữ phần lớn tài sản kỹ thuật số (dữ liệu, phần mềm) và các cơ sở hạ tầng để vận hành các công nghệ kỹ thuật số và AI (Nilekani, 2018). Phần lớn dữ liệu tìm kiếm trên Internet đều được lưu giữ bởi Google, trong khi Meta (hay Facebook trước đây) thống trị mạng xã hội với hơn 2 tỷ người dùng. Với số lượng người dùng và dữ liệu khổng lồ thu được từ họ, mặc dù các tập đoạn này không sở hữu nhiều tài sản vật lý, không có cảnh sát, tòa án hoặc các cơ sở tương tự nhà nước, nhưng họ vẫn có khả năng kiểm soát các nguồn thông tin, tác động vào ý kiến, và thậm chí là thao túng tâm lý và hành vi của lượng lớn người dùng (Shadmy, 2019).

Trong thời đại bùng nổ công nghệ thông tin và AI, sự thay đổi trong cấu trúc sức mạnh của các thành phần trong xã hội đang diễn ra. Sự biến chuyển, thậm chí là nâng cấp, trong hợp đồng xã hội là cần thiết để xã hội thích nghi, thậm chí là tiến hóa, nhưng vẫn đảm bảo được sự ổn định chính trị và xã hội, mà trong đấy an ninh thông tin là một phần thiết yếu. Các hợp đồng xã hội chỉ có sự tham gia của các chính phủ riêng lẻ thì rất khó có khả năng được đảm bảo. Vì thế, hợp đồng xã hội cần có sự phối hợp và kết nối giữa các bên thông qua hợp tác giữa các chính phủ, các tổ chức siêu quốc gia, các đối tác công tư, với công dân, các tổ chức phi chính phủ, và các công ty tư nhân (đặc biệt là các tập đoàn công nghệ) (Liaropoulos, 2020).

5. Nhận thức, đầu tư về an ninh mạng, và một số kiến nghị cho Việt Nam

An ninh và an toàn thông tin đóng vai trò quan trọng trong các hoạt động kinh tế xã hội hiện đại và bảo vệ quốc gia (Nash-Hoff, 2012). Trong bối cảnh toàn cầu hóa và hội nhập kinh tế đang diễn ra, mối quan hệ giữa kinh tế, đặc biệt là thương mại điện tử, và an ninh quốc gia ngày càng trở nên gắn bó với nhau (Okhrimenko *et al.*, 2023). Khi công nghệ tiến bộ, không gian và thời gian sống trong thế giới ảo của con người càng gia tăng, các hệ thống thực ảo được triển khai và vận hành rộng rãi hơn trong nền kinh tế toàn cầu và các hoạt động xã hội, thì việc bảo vệ thông tin sẽ trở thành một nhu cầu thiết yếu để đảm bảo không chỉ vấn đề an ninh thông tin, mà còn là sự phát triển bền vững và an ninh quốc gia.

Sự phát triển nhanh chóng của AI vừa cho thấy các tiềm năng vượt trội trong lĩnh vực an ninh thông tin, nhưng cũng mang đến các nỗi lo đáng kể vì các hacker cũng có thể sử dụng AI cho mục đích tấn công mạng hoặc lừa đảo. Điều này vô hình chung tạo nên một cuộc chạy đua giữa hai phe: phe phòng thủ và phe tấn công. Bên nào có thể phát triển được AI tốt hơn, nhanh hơn, và vận dụng hiệu quả hơn thì sẽ có nhiều lợi thế hơn. Chính vì thế, tập trung nguồn lực vào phát triển AI dùng trong lĩnh vực an ninh mạng sẽ cần được đầu tư trọng điểm để đảm bảo tài sản (thông tin) của quốc gia, doanh nghiệp, và cá nhân không bị thất thoát và khai thác trái phép cho những mục tiêu ác ý hoặc gián điệp. Tuy nhiên, các vấn đề liên quan đến tính hiệu quả đầu tư cũng cần phải được cân nhắc cẩn thận để tránh đầu tư không hiệu quả và lãng phí (Vuong, 2018).

Việc đầu tư các mô hình AI mới sẽ rất tốn kém và vượt quá khả năng chi trả của phần lớn doanh nghiệp, đặc biệt là ở các nước đang phát triển như Việt Nam. Chưa kể, AI là một hệ thống học máy, nên sẽ cần được liên tục huấn luyện và cập nhật các tính năng, thuật toán mới để đảm bảo hệ thống có khả năng ứng đối với các phương thức tấn công được thiết kế ngày càng tinh vi và tùy chỉnh theo từng mục tiêu và hoàn cảnh của tội phạm mạng. Việc sử dụng các mô hình AI được phát triển bởi các tập đoàn công nghệ lớn, như Microsoft, sẽ giúp giảm rõ rệt chi phí cho mục đích bảo mật. Tuy nhiên, cách tiếp cận này sẽ làm lộ tất cả các điểm yếu bảo mật cho các công ty cung cấp dịch vụ AI. Nếu việc này xảy ra ở trên quy mô lớn, thì nó có khả năng dẫn đến các rủi ro về hoạt động gián điệp và thao túng ở quy mô quốc gia. Do đó, chính phủ cần có các chính sách và chương trình hỗ trợ đặc biệt để phối hợp với các doanh nghiệp an ninh mạng trong nước để tự phát triển các hệ thống AI bảo mật, song song với việc dùng dịch vụ cung cấp ngoài cho các loại dữ liệu và hệ thống thông tin không ảnh hưởng nhiều đến an ninh quốc gia. Đứng trước thách thức an ninh quốc gia, việc hợp tác này về cơ bản phải loại bỏ được xung đột lợi ích thương mại thuần túy, trong khi vẫn phải đảm bảo lợi ích quyền lợi hợp pháp với tài sản trí tuệ.

Hiện nay, an toàn và an ninh thông tin tại Việt Nam đang chứng kiến những bước tiến quan trọng. Trong năm 2023, Việt Nam đặt mục tiêu trở thành "cường quốc an toàn thông tin mạng" vào năm 2025, với việc phát triển và xuất khẩu sản phẩm, dịch vụ an toàn thông tin mạng. Nước này cũng đang tập trung vào việc xây dựng nguồn nhân lực chất lượng cao trong lĩnh vực này (Anh, 2024). Việt Nam có các công ty và tổ chức an ninh mạng có khả năng cung cấp các dịch vụ an ninh và an toàn thông tin chuyên nghiệp như của Viettel, CTCP Công nghệ An ninh Không gian Mạng Việt Nam, CTCP Dịch vụ Công nghệ Tin học HPT, CTCP Tập đoàn Công nghệ CMC... Hơn nữa, Việt Nam còn tham gia Liên minh Xác thực trực tuyến thế giới (FIDO), giúp tiếp cận với các xu hướng và giải pháp công nghệ xác thực không mật khẩu tiên tiến, và tổ chức Hội thảo - Triển lãm quốc tế Ngày An toàn thông tin Việt Nam 2023 (Tạp chí An toàn thông tin, 2023). Đây là những điều kiện tiền đề rất hữu ích để Việt Nam triển khai phát triển một mô hình AI dùng riêng cho bảo mật.

Ở khía cạnh doanh nghiệp, nhận thức về an ninh và an toàn thông tin trong các doanh nghiệp chưa thực sự sâu sắc. Các doanh nghiệp Việt Nam đa phần vẫn sử dụng đội ngũ công nghệ thông tin "kiêm nhiệm" vừa phát triển hệ thống vừa làm công tác bảo mật thay vì thuê các đơn vị chuyên nghiệp. Công tác ngăn chặn thông tin độc hại còn yếu và nhiều thiếu sót. Mức độ trưởng thành về an ninh mạng trong doanh nghiệp vẫn chưa tương xứng với mối đe dọa về an toàn thông tin. Một cuộc khảo sát về mức độ trưởng thành về an ninh mạng của McKinsey năm 2021 với hơn 100 công ty thuộc các lĩnh vực và ngành khác nhau cho thấy mối quan hệ tương quan giữa mức độ trưởng thành về an ninh mạng và tỷ suất lợi nhuận. Điều này thể hiện các chiến lược an ninh mạng hiệu quả có thể đóng góp vào sức khỏe tài chính tổng thể của công ty (Eiden *et al*., 2021). Vì thế, các công ty Việt Nam cần nghiêm túc hơn trong việc đầu tư vào các biện pháp an ninh mạng, đặc biệt là trong kỷ nguyên AI, tội phạm mạng có thể phát triển nhanh chóng cả về số lượng và chất lượng.

Khi AI càng mạnh và đa năng, yếu tố con người trở nên cực kỳ quan trọng, vì nó sẽ góp phần quyết định mức độ hiệu quả vận dụng AI và chống chọi với các rủi ro an ninh thông tin. Cho nên, ngoài việc đầu tư phát triển mô hình AI cho mục tiêu bảo mật, các hoạt động nâng cao nhận thức về tầm quan trọng của thông tin, rủi ro bị khai thác thông tin, cũng như huấn luyện và đào tạo người dân, doanh nghiệp, và các cơ quan chính phủ cách bảo vệ thông tin và hệ thống thông tin cũng cần được chú trọng. Bằng cách này, các cá nhân, doanh nghiệp, và các cơ quan chính phủ tham gia vào các hoạt động trên không gian mạng sẽ có được ý thức và năng lực tự bảo vệ mình trước các rủi ro bảo mật, từ đấy góp phần vào sự bền vững của không gian thông tin quốc gia. Thật vậy, các vấn đề an ninh thông tin đã và đang xuất ngày càng nhiều ở Việt Nam. Ví dụ như các vụ lừa đảo mới diễn ra trên các ứng dụng như Zalo và Telegram, hay sự xuất hiện của công nghệ *deepfake* trong các vụ lừa đảo (Sơn, 2023).

Cho dù hệ thống bảo mật tích hợp AI có hoàn thiện đến đâu đi nữa, nó sẽ luôn có khả năng để lọt các lỗ hổng tạo ra do lỗi con người. Một trong những vấn đề đáng chú ý là sự thiếu tuân thủ đầy đủ

các quy định về an toàn thông tin từ phía một số cơ quan chính phủ. Điều này được thể hiện rõ qua việc chỉ cần tìm kiếm các từ khóa như "cá độ" hay "bóng đá" trên các tên miền của nhiều cơ quan cũng có thể phát hiện ra sự xâm nhập của *hacker* và sự xuất hiện của nội dung không mong muốn. Những sự cố bảo mật này không chỉ gây ra hậu quả về việc phát tán thông tin không phù hợp, mà còn tiềm ẩn nguy cơ lớn nếu *hacker* lợi dụng để đưa ra thông tin sai lệch, gửi tới xã hội các thông điệp mạo danh, hoặc thực hiện các hành vi lừa đảo, gây ra hậu quả nghiêm trọng. Vấn đề này cần có biện pháp khắc phục kịp thời để nâng cao độ an toàn và bảo mật cho các hệ thống thông tin của các cơ quan chính phủ.

Việt Nam cũng cần nhận thức rõ tầm quan trọng của việc phát triển nguồn nhân lực trong lĩnh vực công nghệ thông tin và an ninh mạng (Vuong *et al.*, 2019). Hiện đã có một số trường đại học tích cực đưa môn học An toàn Thông tin vào chương trình giảng dạy. Tuy nhiên, cả về số lượng và chất lượng của các khóa học này vẫn chưa thật sự đáp ứng được nhu cầu, và vẫn ở giai đoạn sơ khởi. Để theo kịp với sự phát triển nhanh chóng của công nghệ, nội dung giảng dạy cùng với đội ngũ giảng viên cần được cập nhật liên tục, nhằm đáp ứng những tiến bộ công nghệ mới và thích ứng với xu hướng thời đại. Điều này không chỉ giúp cung cấp kiến thức và kỹ năng cần thiết cho sinh viên, mà còn góp phần nâng cao năng lực tổng thể của ngành công nghệ thông tin và an ninh mạng ở Việt Nam. Thêm vào đó, chính phủ và các trường đại học cần định hướng và thúc đẩy các nghiên cứu khoa học xã hội, tâm lý, và hành vi liên quan tới vấn đề an ninh mạng, vì con người vẫn là mắt xích quan trọng nhất, và cũng có nhiều tồn tại nhất, để có thể đạt được mục tiêu đảm bảo an ninh thông tin một cách toàn diện. Hiện tại, số lượng nghiên cứu về vấn đề này vẫn còn hạn chế (Maalem Lahcen *et al.*, 2020; Payne & Hadzhidimova, 2018).

Tóm lại, Việt Nam đang nỗ lực trong việc nâng cao an toàn và an ninh thông tin, nhưng vẫn còn đối mặt với nhiều thách thức. Việc tăng cường tuân thủ các quy định an toàn thông tin, đối phó với các hình thức lừa đảo mới, và cải thiện nhận thức về an ninh thông tin trong cộng đồng doanh nghiệp là những điểm cần được chú trọng. Đồng thời, việc phát triển nguồn nhân lực chất lượng cao trong lĩnh vực này, đặc biệt là thông qua các chương trình giảng dạy đại học cập nhật và chuyên sâu, sẽ là chìa khóa để Việt Nam tự cường về an ninh kinh tế - xã hội, tiếp tục tiến xa hơn trên con đường trở thành một quốc gia hàng đầu về an toàn thông tin mạng./.

(Hết)

Vương Quân Hoàng [1], Lã Việt Phương [1], Nguyễn Hồng Sơn [2], Nguyễn Minh Hoàng [1]

[1] Trung tâm nghiên cứu ISR, Trường Đại học Phenikaa

[2] Nguyên Chánh văn phòng, Hội đồng Lý luận Trung ương

Tài liệu tham khảo

Anh, N. (2024). *Hướng tới mục tiêu "cường quốc an toàn thông tin mạng"*. VnEconomy. Retrieved March 19 from https://vneconomy.vn/huong-toi-muc-tieu-cuong-quoc-an-toan-thong-tin-mang.htm

Bierens, R., Klievink, B., & van Den Berg, J. (2017). A social cyber contract theory model for understanding national cyber strategies. Electronic Government: 16th IFIP WG 8.5 International Conference, St. Petersburg, Russia.

Biggio, B., Corona, I., Maiorca, D., Nelson, B., Šrndić, N., Laskov, P., . . . Roli, F. (2013). Evasion attacks against machine learning at test time. Machine Learning and Knowledge Discovery in Databases: European Conference, Prague, Czech Republic.

Biggio, B., Fumera, G., & Roli, F. (2013). Security evaluation of pattern classifiers under attack. *IEEE Transactions on Knowledge and Data Engineering*, *26*(4), 984-996. https://doi.org/10.1109/TKDE.2013.57

Boucher, D., & Kelly, P. (2003). *The social contract from Hobbes to Rawls*. Routledge.

Chính, P. M., & Hoàng, V. Q. (2009). *Kinh tế Việt Nam: Thăng trầm và đột phá*. Nxb Chính trị quốc gia-Sự thật.

Clifford, C. (2018). *Google CEO: A.I. is more important than fire or electricity*. CNBC. Retrieved March 18 from https://www.cnbc.com/2018/02/01/google-ceo-sundar-pichai-ai-is-more-important-than-fire-electricity.html

Eiden, K., Kaplan, J., Kazimierski, B., Lewis, C., & Telford, K. (2021). *Organizational cyber maturity: A survey of industries*. https://www.mckinsey.com/capabilities/risk-and-resilience/our-insights/organizational-cyber-maturity-a-survey-of-industries

Green, L. (1998). Power. In *Routledge Encyclopedia of Philosophy*: Taylor and Francis.

Greenberg, A. (2015). *Hackers remotely kill a Jeep on the highway—With me in it*. WIRED. Retrieved March 17 from https://www.wired.com/2015/07/hackers-remotely-kill-jeep-highway/

Handa, A., Sharma, A., & Shukla, S. K. (2019). Machine learning in cybersecurity: A review. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *9*(4), e1306. https://doi.org/10.1002/widm.1306

Henshall, W. (2023). *4 charts that show why AI progress is unlikely to slow down*. Time. Retrieved March 18 from https://time.com/6300942/ai-progress-charts/

Ho, M.-T., & Vuong, Q.-H. (2023). Disengage to survive the AI-powered sensory overload world. *AI and Society*. https://doi.org/10.1007/s00146-023-01714-0

Hobbes, T. (1894). *Leviathan: Or, the matter, form, and power of a commonwealth ecclesiastical and civil* (Vol. 21). G. Routledge and sons.

Hu, K. (2023). *ChatGPT sets record for fastest-growing user base - analyst note*. Reuters. Retrieved 11 May from https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/

Insiders, C. (2023). *2023 insider threat report*. https://istari-global.com/insights/spotlight/2023-insider-threat-report/

Keck, M., Gillani, S., Dermish, A., Grossman, J., & Rühmann, F. (2022). *The role of cybersecurity and data security in the digital economy*. https://policyaccelerator.uncdf.org/all/brief-cybersecurity-digital-economy

Kramer, M. H. (2008). *The quality of freedom*. Oxford University Press.

Lappin, S. (2023). Assessing the strengths and weaknesses of Large Language Models. *Journal of Logic, Language and Information*, *33*, 9-20. https://doi.org/10.1007/s10849-023-09409-x

Liaropoulos, A. (2020). A social contract for cyberspace. *Journal of Information Warfare*, *19*(2), 1-11. https://www.jstor.org/stable/27033617

Locke, J. (1967). *Two treatises of government*. Cambridge university press.

Lu, Y., & Da Xu, L. (2018). Internet of Things (IoT) cybersecurity research: A review of current research topics. *IEEE Internet of Things Journal*, *6*(2), 2103-2115. https://doi.org/10.1109/JIOT.2018.2869847

Maalem Lahcen, R. A., Caulkins, B., Mohapatra, R., & Kumar, M. (2020). Review and insight on the behavioral aspects of cybersecurity. *Cybersecurity*, *3*, 1-18. https://doi.org/10.1186/s42400-020-00050-w

Mantello, P., Ho, M.-T., Nguyen, M.-H., & Vuong, Q.-H. (2023). Machines that feel: behavioral determinants of attitude towards affect recognition technology—upgrading technology acceptance theory with the mindsponge model. *Humanities and Social Sciences Communications*, *10*, 430. https://doi.org/10.1057/s41599-023-01837-1

Marr, B. (2021). *How much data do we create every day? The mind-blowing stats everyone should read*. Bernard Marr & Co. Retrieved March 18 from https://bernardmarr.com/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/

Martínez Torres, J., Iglesias Comesaña, C., & García-Nieto, P. J. (2019). Machine learning techniques applied to cybersecurity. *International Journal of Machine Learning and Cybernetics*, *10*(10), 2823-2836. https://doi.org/10.1007/s13042-018-00906-1

Morgan, S. (2022). *Cybercrime to cost the world 8 trillion annually in 2023*. Cybercrime Magazine. Retrieved March 18 from https://cybersecurityventures.com/cybercrime-to-cost-the-world-8-trillion-annually-in-2023/

Musa, S. (2018). Smart cities-a road map for development. *IEEE Potentials*, *37*(2), 19-23. https://doi.org/10.1109/MPOT.2016.2566099

Nash-Hoff, M. (2012). *What does the economy have to do with national security?* IndustryWeek. Retrieved March 19 from https://www.industryweek.com/finance/software-systems/article/21954333/what-does-the-economy-have-to-do-with-national-security

Nguyen, M.-H., Le, T.-T., & Vuong, Q.-H. (2023). Ecomindsponge: A novel perspective on human psychology and behavior in the ecosystem. *Urban Science*, *7*(1), 31. https://doi.org/10.3390/urbansci7010031

Nilekani, N. (2018). Data to the people: India's inclusive internet. *Foreign Affairs*, *97*(5), 19-27.

Novet, J. (2023). *Microsoft introduces an A.I. chatbot for cybersecurity experts*. CNBC. Retrieved March 18 from https://www.cnbc.com/2023/03/28/microsoft-launches-security-copilot-in-private-preview.html

Novet, J. (2024). *Microsoft says new AI security chatbot pricing model lets customers 'buy what they need'*. CNBC. Retrieved March 18 from https://www.cnbc.com/2024/03/13/microsoft-uses-compute-units-to-charge-customers-for-security-copilot.html

Okhrimenko, I., Stepenko, V., Chernova, O., & Zatsarinnaya, E. (2023). The impact of information sphere in the economic security of the country: case of Russian realities. *Journal of Innovation and Entrepreneurship*, *12*(1), 67. https://doi.org/10.1186/s13731-023-00326-8

Paiho, S., Tuominen, P., Rökman, J., Ylikerälä, M., Pajula, J., & Siikavirta, H. (2022). Opportunities of collected city data for smart cities. *IET Smart Cities*, *4*(4), 275–291. https://

doi.org/10.1049/smc2.12044

Pansardi, P. (2012). Power and freedom: opposite or equivalent concepts? *Theoria*, *59*(132), 26-44. https://www.jstor.org/stable/41802526

Payne, B. K., & Hadzhidimova, L. (2018). Cyber security and criminal justice programs in the United States: Exploring the intersections. *International Journal of Criminal Justice Sciences*, *13*(2). https://doi.org/10.5281/zenodo.2657646

Rao, Vikram Singh. (2021). *Best AI-based cyber security tools for improved safety*. Echnotification. Retrieved March 19 from https://www.technotification.com/2021/06/best-ai-based-cyber-security-tools.html

Rieck, K., Trinius, P., Willems, C., & Holz, T. (2011). Automatic analysis of malware behavior using machine learning. *Journal of Computer Security*, *19*(4), 639-668. https://doi.org/10.5555/2011216.2011217

RiskXchange. (2023). *Cybersecurity statistics you should know in 2023*. RiskXchange. Retrieved March 19 from https://riskxchange.co/1006415/cybersecurity-statistics-2023/

Rousseau, J.-J. (2016). The social contract. In R. Blaug & J. Schwarzmantel (Eds.), *Democracy: A Reader* (pp. 43-51). Columbia University Press.

Shadmy, T. (2019). The new social contract: Facebook's community and our rights. *Boston University International Law Journal*, *37*, 307.

Sharma, A., & Sahay, S. K. (2014). Evolution and detection of polymorphic and metamorphic malwares: A survey. *International Journal of Computer Applications*, *90*(2), 7–11. https://doi.org/10.48550/arXiv.1406.7061

Sơn, M. (2023). *An ninh mạng: Những xu hướng đáng chú ý trong 6 tháng cuối năm 2023*. VietnamPlus. Retrieved March 19 from https://www.vietnamplus.vn/an-ninh-mang-nhung-xu-huong-dang-chu-y-trong-6-thang-cuoi-nam-2023-post869804.vnp

Stacey, K., & Milmo, D. (2023). *AI developing too fast for regulators to keep up, says Oliver Dowden*. The Guardian. Retrieved March 18 from https://www.theguardian.com/technology/2023/sep/22/ai-developing-too-fast-for-regulators-to-keep-up-oliver-dowden

Stanton, J. M., Stam, K. R., Mastrangelo, P., & Jolton, J. (2005). Analysis of end user security behaviors. *Computers and Security*, *24*(2), 124-133. https://doi.org/10.1016/j.cose.2004.07.001

Suleyman, M. (2023). *How the AI revolution will reshape the world*. Time. Retrieved March 18 from https://time.com/6310115/ai-revolution-reshape-the-world/

Tạp chí An toàn thông tin. (2023). *An toàn thông tin 10 dấu ấn nổi bật trong lĩnh vực bảo mật và an ninh, an toàn thông tin tại Việt Nam năm 2023*. Trung tâm Công nghệ Thông tin và Truyền thông Nghệ An. Retrieved March 19 from https://naict.tttt.nghean.gov.vn/attt/an-toan-thong-tin-10-dau-an-noi-bat-trong-linh-vuc-bao-mat-va-an-ninh-an-toan-thong-tin-tai-viet-nam-nam-2023-597.html

Vailshery, L. S. (2023). *Number of IoT connected devices worldwide 2019-2023, with forecasts to 2030*. Statista. Retrieved March 18 from https://www.statista.com/statistics/1183457/iot-connected-devices-worldwide/

Vanian, J., & Leswing, K. (2023). *ChatGPT and generative AI are booming, but the costs can be extraordinary*. CNBC. Retrieved March 18 from https://www.cnbc.com/2023/03/13/chatgpt-and-generative-ai-are-booming-but-at-a-very-expensive-price.html

Vuong, Q.-H. (2018). The (ir) rational consideration of the cost of science in transition economies. *Nature Human Behaviour*, *2*(1), 5.

Vuong, Q.-H. (2023). *Mindsponge Theory*. Walter de Gruyter GmbH. https://www.amazon.com/dp/B0C3WHZ2B3/

Vuong, Q.-H., *et al*. (2019). Artificial intelligence vs. natural stupidity: Evaluating AI readiness for the vietnamese medical information system. *Journal of Clinical Medicine*, *8*(2), 168. https://doi.org/10.3390/jcm8020168

Vuong, Q.-H., *et al*. (2023a). AI's humanoid appearance can affect human perceptions of Its emotional capability: Evidence from self-reported data in the US. *International Journal of Human–Computer Interaction*, 1-12. https://doi.org/10.1080/10447318.2023.2227828

Vuong, Q.-H., *et al*. (2023b). How AI's self-prolongation influences people's perceptions of its autonomous mind: The case of US residents. *Behavioral Sciences*, *13*(6), 470. https://doi.org/10.3390/bs13060470

Vuong, Q.-H., *et al*. (2023). Are we at the start of the artificial intelligence era in academic publishing? *Science Editing*, *10*(2), 158-164. https://doi.org/10.6087/kcse.310

Vuong, Q.-H., Nguyen, M.-H., & La, V.-P. (2022). *The mindsponge and BMF analytics for innovative thinking in social sciences and humanities*. Walter de Gruyter GmbH. https://www.amazon.com/dp/B0C4ZK3M74/

Wise, J. (2023). *How many Google searches per minute in 2024?* EarthWeb. Retrieved March 18 from https://earthweb.com/how-many-google-searches-per-minute/

World Economic Forum. (2023). *The global risks report 2023*. https://www3.weforum.org/docs/WEF_Global_Risks_Report_2023.pdf