BỘ KHOA HỌC VÀ CÔNG NGHỆ
CỤC THÔNG TIN KHOA HỌC VÀ CÔNG NGHỆ QUỐC GIA

TỔNG LUẬN | KHOA HỌC CÔNG NGHỆ KINH TẾ

ISSN 0866 - 7721

Số 4 - 2021

# NHỮNG CÂN NHẮC CHÍNH SÁCH THÚC ĐẨY PHÁT TRIỂN TRÍ TUỆ NHÂN TẠO

Hà Nội, 4/2021

**CỤC THÔNG TIN VÀ KHOA HỌC CÔNG NGHỆ QUỐC GIA**
Địa chỉ: 24, Lý Thường Kiệt, Hoàn Kiếm, Hà Nội.
Tel: (024) 38262718, Fax: (024) 39349127
**BAN BIÊN TẬP**
**TS. Trần Đắc Hiến** (Trưởng ban); ThS. Trần Thị Thu Hà (Phó Trưởng ban)
KS. Nguyễn Mạnh Quân; ThS. Nguyễn Lê Hằng; ThS. Phùng Anh Tiến

# MỤC LỤC

**LỜI MỞ ĐẦU ……...….1**

**I. TỔNG QUAN HỆ THỐNG TRÍ TUỆ NHÂN TẠO ……...2**

1.1. Khái quát hệ thống TTNT ……2

1.2. Ứng dụng của TTNT…….…5

**II. CÁC CÂN NHẮC CHÍNH SÁCH CÔNG ……...9**

2.1. Yêu cầu chính sách đối với TTNT ……9

2.2. Chính sách thúc đẩy sự phát triển của TTNT…….15

**III. CHÍNH SÁCH VÀ SÁNG KIẾN TTNT …….26**

3.1. Nguyên tắc TTNT trong xã hội……..26

3.2. Sáng kiến quốc gia……..28

3.3. Các sáng kiến liên chính phủ ……....36

3.4. Sáng kiến của các bên liên quan tư nhân……40

**KẾT LUẬN ……..46**

# LỜI MỞ ĐẦU

Trí tuệ nhân tạo (TTNT) đang định hình lại nền kinh tế, hứa hẹn làm tăng năng suất, cải thiện hiệu quả và giảm chi phí. TTNT góp phần làm cho cuộc sống tốt đẹp hơn và giúp mọi người đưa ra dự báo chính xác và quyết định sáng suốt hơn. Dù các công nghệ TTNT đang ở giai đoạn phát triển ban đầu, nhưng vẫn có triển vọng giải quyết những thách thức toàn cầu, thúc đẩy đổi mới sáng tạo và tăng trưởng.

Tuy nhiên, TTNT cũng gây lo ngại về khía cạnh đạo đức. Nhiều vấn đề đặt ra liên quan đến mức độ đáng tin cậy của hệ thống TTNT, bao gồm nguy cơ làm gia tăng các thành kiến hiện nay về giới tính và chủng tộc hoặc vi phạm các quyền và giá trị của con người như quyền riêng tư. Lo ngại hệ thống TTNT làm trầm trọng tình trạng bất bình đẳng, biến đổi khí hậu, sự tập trung thị trường và sự phân chia kỹ thuật số, ngày càng lớn. Chưa một quốc gia hay chủ thể nào đưa ra được giải pháp ứng phó với những thách thức này. Do đó, cần có sự hợp tác quốc tế và phản ứng của nhiều bên liên quan để định hướng việc phát triển và sử dụng TTNT vì lợi ích chung lớn lao.

Tổng luận "***Những cân nhắc chính sách thúc đẩy phát triển trí tuệ nhân tạo***" đề **cập đến bối cảnh phát triển TTNT và nêu bật các vấn đề chính sách quan trọng. Mục tiêu** hướng đến là cung cấp kiến thức về TTNT và nhấn mạnh đến việc đẩy mạnh đối thoại rộng rãi các vấn đề chính sách quan trọng như phát triển thị trường lao động và nâng cao kỹ năng trong kỷ nguyên số; quyền riêng tư; trách nhiệm giải trình của các quyết định được đưa ra dựa vào TTNT; và những vấn đề về trách nhiệm, bảo mật và an toàn nảy sinh từ TTNT.

Trân trọng giới thiệu.

**CỤC THÔNG TIN KHOA HỌC**
**VÀ CÔNG NGHỆ QUỐC GIA**

## I. TỔNG QUAN HỆ THỐNG TRÍ TUỆ NHÂN TẠO

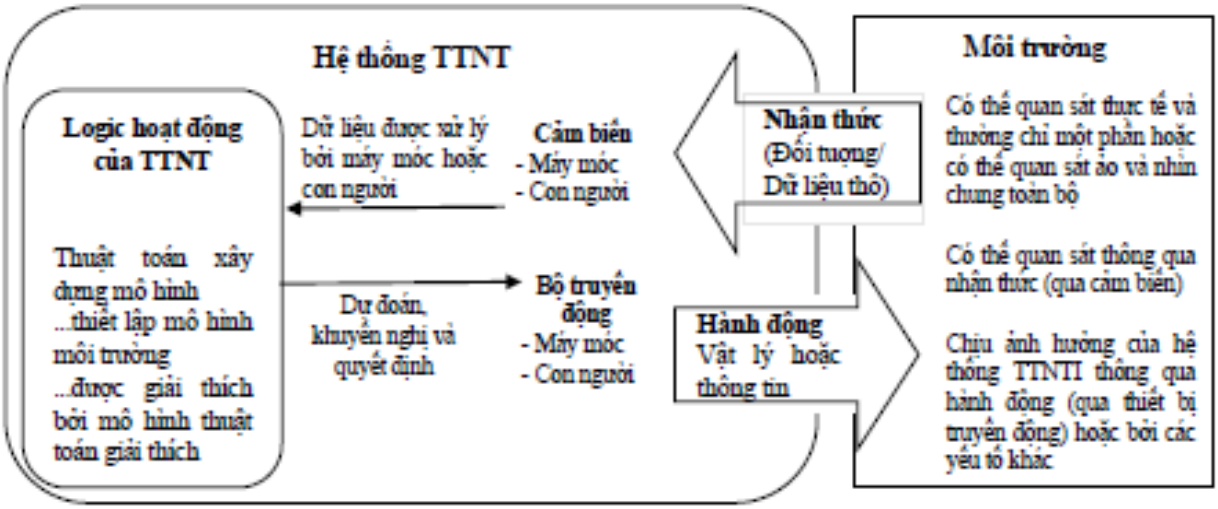
### *1.1. Khái quát hệ thống TTNT*

Đến nay chưa có định nghĩa thống nhất về trí tuệ nhân tạo (TTNT). Tháng 11/2018, Nhóm chuyên gia TTNT của OECD (AIGO) đã đưa ra mô tả hệ thống TTNT theo hướng dễ hiểu, chính xác về mặt kỹ thuật, trung tính về công nghệ (technology-neutral) và có thể áp dụng trước mắt và lâu dài. Phạm vi mô tả hệ thống đủ rộng để bao trùm nhiều định nghĩa TTNT, thường được sử dụng bởi các cộng đồng khoa học, doanh nghiệp và chính sách.

***Quan điểm về khái niệm hệ thống TTNT***

Mô tả hiện nay về hệ thống TTNT dựa trên quan điểm về khái niệm TTNT được trình bày chi tiết trong cuốn sách "*Trí tuệ nhân tạo: Phương pháp tiếp cận hiện đại*" do hai tác giả Stuart J. Russell và Peter Norvig viết và được phát hành năm 2009. Quan điểm này phù hợp với định nghĩa TTNT được sử dụng rộng rãi, đó là TTNT "nghiên cứu các tính toán giúp nó có thể nhận thức, suy luận và hành động" (Winston, 1992 [10]) và với các định nghĩa chung tương tự (Gringsjord và Govindarajulu, 2018 [ 11]).

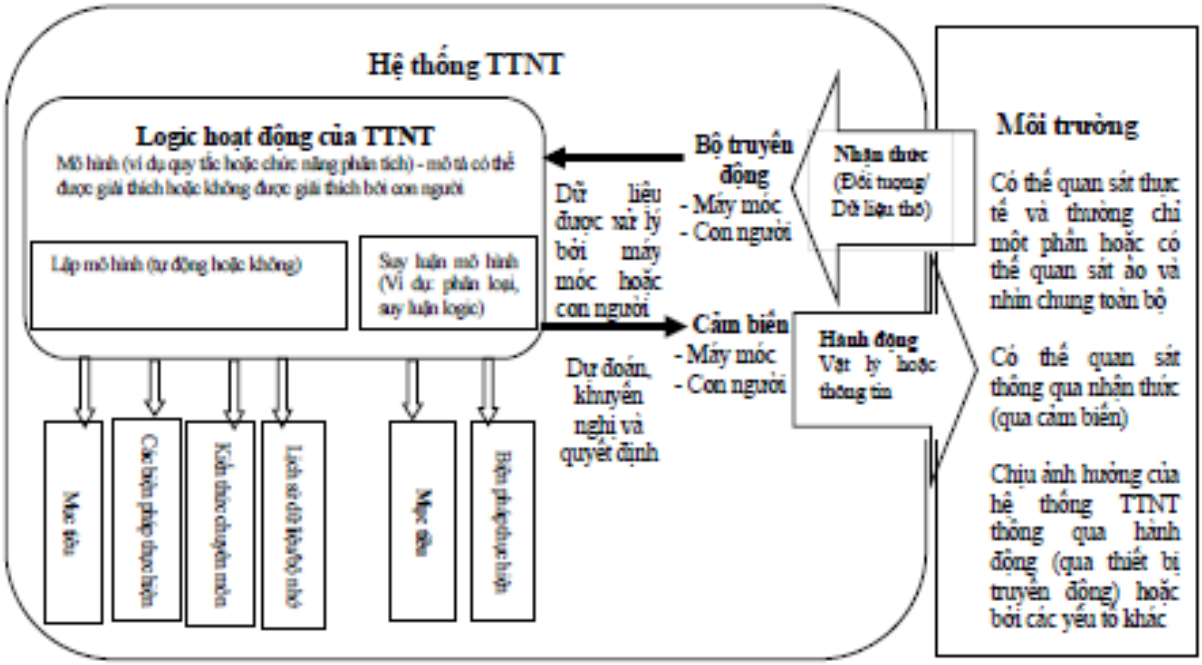
Quan điểm về khái niệm TTNT lần đầu tiên được trình bày dưới dạng cấu trúc cấp cao của hệ thống TTNT chung (còn được gọi là "tác nhân thông minh") (Hình 1). Hệ thống TTNT bao gồm ba yếu tố chính: cảm biến, logic hoạt động và bộ truyền động. Bộ cảm biến thu thập dữ liệu thô từ môi trường, trong khi bộ truyền động hoạt động để thay đổi trạng thái của môi trường. Sức mạnh chính của hệ thống TTNT nằm ở logic hoạt động của nó. Đối với một tập hợp các mục tiêu nhất định và dựa vào dữ liệu đầu vào từ các cảm biến, logic hoạt động cung cấp đầu ra cho bộ truyền động. Kết quả đầu ra tồn tại dưới dạng các khuyến nghị, dự đoán hoặc quyết định có thể ảnh hưởng đến hiện trạng môi trường.




***Hình 1.*** Quan điểm khái niệm cấp cao về hệ thống TTNT

*Nguồn: AIGO (2/2019)*

Cấu trúc chi tiết thể hiện các yếu tố chính liên quan đến các khía cạnh chính sách của hệ thống TTNT (Hình 2). Để thể hiện các loại hệ thống TTNT và tình huống khác nhau, sơ đồ phân tách quá trình xây dựng mô hình (như máy học) khỏi chính mô hình. Việc xây dựng mô hình cũng tách biệt với quá trình diễn giải mô hình, sử dụng mô hình để đưa ra các dự đoán, khuyến nghị và quyết định; bộ truyền động sử dụng các đầu ra này để tác động đến môi trường.




***Hình 2.*** Mô tả khái niệm chi tiết về Hệ thống TTNT

*Nguồn: AIGO (2/2019)*

*Môi trường*

Môi trường liên quan đến hệ thống TTNT là không gian có thể quan sát được thông qua nhận thức (qua cảm biến) và bị ảnh hưởng thông qua hành động (qua bộ truyền động). Cảm biến và bộ truyền động là máy móc hoặc con người. Môi trường hoặc là thực (như vật chất, xã hội, tinh thần) và thường chỉ có thể quan sát bán phần hoặc là ảo (như trò chơi trên bàn cờ) và thường có thể quan sát toàn phần.

*Hệ thống TTNT*

Hệ thống TTNT là hệ thống dựa vào máy móc, phụ thuộc vào một tập hợp các mục tiêu nhất định do con người xác định để đưa ra các dự đoán, khuyến nghị hoặc quyết định ảnh hưởng đến môi trường thực hoặc ảo.

Để làm được điều đó, hệ thống TTNT sử dụng các đầu vào từ máy móc và/hoặc con người để: i) nhận thức môi trường thực và/hoặc ảo; ii) trừu tượng hóa những nhận thức đó thành các mô hình thông qua phân tích tự động (như với máy học hoặc theo cách thủ công); và iii) sử dụng suy luận mô hình để đưa ra các lựa chọn thông tin hoặc hành động. Hệ thống TTNT được thiết kế để hoạt động với nhiều mức độ tự chủ khác nhau.

*Mô hình TTNT, xây dựng mô hình và giải thích mô hình*

Cốt lõi của hệ thống TTNT là mô hình TTNT đại diện cho toàn bộ hoặc một phần môi trường bên ngoài của hệ thống mô tả cấu trúc và/hoặc động lực của môi trường. Một mô hình có thể dựa vào tri thức chuyên môn và/hoặc dữ liệu, bởi con người và/hoặc các công cụ tự động (như thuật toán máy học). Các mục tiêu (như các biến đầu ra) và các thước đo hiệu quả (như độ chính xác, tính đại diện của bộ dữ liệu) định hướng quá trình xây dựng. Suy luận mô hình là quá trình con người và/hoặc các công cụ tự động thu được kết quả từ mô hình, có thể ở dạng khuyến nghị, dự đoán hoặc quyết định. Mục tiêu và các thước đo hiệu quả định hướng việc thực hiện. Trong một số trường hợp (như các quy tắc xác định), mô hình đưa ra một khuyến nghị duy nhất. Trong các trường hợp khác (như mô hình xác suất), mô hình đưa ra nhiều khuyến nghị. Ví dụ, các khuyến nghị này được liên kết với các thước đo hiệu quả ở cấp độ khác nhau như mức độ tin cậy, độ chắc chắn hoặc rủi ro.

***Minh họa hệ thống TTNT***

*Hệ thống chấm điểm tín dụng*

Hệ thống chấm điểm tín dụng minh họa một hệ thống dựa vào máy móc ảnh hưởng đến môi trường của nó (dù khách hàng có được cho vay hay không). Hệ thống đưa ra các khuyến nghị (điểm tín dụng) cho một tập hợp các mục tiêu nhất định (giá trị tín dụng), bằng cách sử dụng cả đầu vào dựa trên máy móc (dữ liệu lịch sử về hồ sơ của khách hàng và khả năng chi trả các khoản vay) và đầu vào từ con người (một bộ quy tắc). Với hai tập hợp đầu vào này, hệ thống nhận biết môi trường thực tế (liệu khách hàng có đang trả nợ liên tục hay không). Hệ thống tự động tóm tắt những nhận thức đó thành các mô hình. Ví dụ, một thuật toán cho điểm tín dụng có thể sử dụng một mô hình thống kê. Cuối cùng, nó sử dụng suy luận mô hình (thuật toán chấm điểm tín dụng) để đưa ra đề xuất (điểm tín dụng) về các lựa chọn cho kết quả (cung cấp hoặc từ chối một khoản vay).

*Trợ lý cho người khiếm thị*

Trợ lý cho người khiếm thị minh họa cách một hệ thống dựa vào máy móc ảnh hưởng đến môi trường. Hệ thống đưa ra các khuyến nghị (như cách để người khiếm thị tránh chướng ngại vật hoặc băng qua đường) cho một nhóm mục tiêu nhất định (đi từ nơi này đến nơi khác). Hệ thống làm điều này bằng cách sử dụng đầu vào dựa vào máy móc và/hoặc con người (cơ sở dữ liệu hình ảnh quy mô lớn được gắn thẻ về các vật thể, chữ viết và thậm chí cả khuôn mặt người) theo ba hướng. Đầu tiên, hệ thống nhận diện hình ảnh của môi trường (máy ảnh chụp lại hình ảnh về những gì ở phía trước của một người và gửi đến ứng dụng). Thứ hai, hệ thống trừu tượng hóa những nhận thức đó thành các mô hình một cách tự động (thuật toán nhận dạng vật thể có thể nhận ra đèn giao thông, ô tô hoặc chướng ngại vật trên vỉa hè). Thứ ba, hệ thống sử dụng suy luận mô hình để đề xuất các lựa chọn cho kết quả (cung cấp mô tả bằng âm thanh về các vật thể được phát hiện trong môi trường) để người đó quyết định cách hành động nên sẽ ảnh hưởng đến môi trường.

*AlphaGo Zero*

AlphaGo Zero là hệ thống TTNT chơi cờ vây tốt hơn bất cứ người chơi cờ vây chuyên nghiệp nào. Môi trường của trò chơi trên bàn cờ là ảo và hoàn toàn có thể quan sát được. Các vị trí trong trò chơi bị ràng buộc bởi các mục tiêu và luật chơi. AlphaGo Zero là hệ thống sử dụng cả đầu vào dựa vào con người (các quy tắc của cờ

vậy) và đầu vào dựa vào máy móc (học dựa vào việc chơi lặp đi lặp lại với chính nó, bắt đầu chơi hoàn toàn ngẫu nhiên). Hệ thống trừu tượng hóa dữ liệu thành mô hình hành động (ngẫu nhiên) ("di chuyển" trong trò chơi) được đào tạo thông qua học tăng cường. Cuối cùng, hệ thống sử dụng mô hình để đề xuất nước đi mới dựa vào trạng thái ván cờ.

*Hệ thống lái xe tự động*

Hệ thống lái xe tự động minh họa một hệ thống dựa vào máy móc, có thể ảnh hưởng đến môi trường của nó (dù ô tô tăng tốc, giảm tốc hay rẽ). Hệ thống đưa ra dự đoán (cho dù một đối tượng hoặc một dấu hiệu là chướng ngại vật hay chỉ dẫn) và/hoặc đưa ra quyết định (tăng tốc, phanh…) cho một nhóm mục tiêu nhất định (đi từ điểm A đến điểm B an toàn)

### 1.2. Ứng dụng của TTNT

***Ứng dụng TTNT trong giao thông cho xe tự lái***

Hệ thống TTNT đang xuất hiện trong toàn bộ nền kinh tế. Tuy nhiên, một trong những chuyển biến mạnh mẽ nhất là trong ngành giao thông với quá trình chuyển đổi sang xe tự lái hay xe tự hành.

Hiện trạng giao thông đang thay đổi là do ba yếu tố tác động của thị trường gần đây: sự phát triển của hệ thống xe tự động, sử dụng dịch vụ đi chung xe và chuyển sang dùng xe điện. Các nhà sản xuất ô tô truyền thống gặp khó khăn trong việc xác định chiến lược phát triển khi phải đối mặt với hai xu hướng. Thứ nhất, dịch vụ đi chung xe làm gia tăng các lựa chọn giao thông khả thi cho người dùng, đặc biệt là thế hệ trẻ. Thứ hai là những vấn đề về khả năng sở hữu lâu dài ô tô truyền thống. Các nhà sản xuất cao cấp đã và đang thử nghiệm các mô hình kinh doanh mới như dịch vụ đăng ký. Ví dụ, khi sử dụng dịch vụ "Truy cập bằng BMW", "Bộ sưu tập Mercedes" và "Hộ chiếu Porsche", người dùng phải trả một khoản phí cố định hàng tháng và được đổi xe khi họ muốn.

Các công ty công nghệ từ công ty đa quốc gia quy mô lớn đến công ty khởi nghiệp nhỏ, đang chuyển sang hệ thống xe tự lái, dịch vụ đi chung xe hoặc dùng xe điện hoặc một số kết hợp cả ba lựa chọn. Do tính phức tạp của hệ thống xe tự động, các công ty có xu hướng tập trung vào các lĩnh vực chuyên môn cụ thể và sau đó hợp tác với các công ty chuyên về các lĩnh vực khác. Waymo là một trong những công ty hàng đầu trong lĩnh vực xe tự động nhờ chuyên môn hóa trong các bộ dữ liệu lớn và máy học. Tuy nhiên, công ty này không tự sản xuất ô tô mà dựa vào các đối tác như General Motors (GM) và Jaguar. Bên cạnh đó, các nhà sản xuất ô tô lớn cũng đã hợp tác với các công ty khởi nghiệp quy mô nhỏ để tiếp cận công nghệ tiên tiến.

Ở cấp độ cơ bản, xe tự động được gắn hệ thống cảm biến mới với khả năng xử lý gây khó khăn cho quá trình trích xuất, biến đổi và tải hệ thống dữ liệu của xe. Hoạt động đổi mới đang gia tăng mạnh mẽ trong bối cảnh đầu tư cho xe tự lái ở mức cao. Ví dụ, các hệ thống phát hiện ánh sáng và định vị giá rẻ giúp quan sát môi trường xung quanh. Ngoài ra, các công nghệ thị giác máy tính mới theo dõi mắt và mức độ tập trung của lái xe và xác định thời điểm lái xe bị phân tâm. Giờ đây, sau khi thu thập và xử lý dữ liệu, TTNT đang tiến thêm thêm một bước nữa, đó là đưa ra quyết định hoạt động chỉ trong tích tắc.

*Ứng dụng TTNT trong nông nghiệp*

Cải thiện độ chính xác của các công nghệ điện toán nhận thức như nhận dạng hình ảnh đang làm thay đổi ngành nông nghiệp. Trước đây, ngành nông nghiệp dựa vào đôi mắt và bàn tay của người nông dân có kinh nghiệm để xác định loại cây trồng đến kỳ thu hoạch. Nhưng giờ rôbốt được trang bị công nghệ TTNT và dữ liệu từ camera và cảm biến đưa ra quyết định thu hoạch trong thời gian thực. Rôbốt này có thể thực hiện được những nhiệm vụ mà trước đây cần sức lao động và tri thức của con người.

Các công ty khởi nghiệp công nghệ đưa ra những giải pháp sáng tạo khai thác TTNT trong nông nghiệp, được phân thành ba loại sau:

- Rôbốt nông nghiệp xử lý các nhiệm vụ thiết yếu như thu hoạch. So với người lao động, rôbốt làm việc nhanh và năng suất hơn.

- Giải pháp theo dõi cây trồng và đất đai sử dụng thị giác máy tính và các thuật toán học sâu để theo dõi tình trạng của cây trồng và đất. Hoạt động này đã được cải thiện do có sẵn dữ liệu vệ tinh.

- Phân tích dự báo sử dụng mô hình máy học để theo dõi và dự đoán tác động của các yếu tố môi trường đến năng suất cây trồng.

*Ứng dụng TTNT trong dịch vụ tài chính*

Trong lĩnh vực tài chính, các công ty lớn như JPMorgan, Citibank, State Farm và Liberty Mutual đang nhanh chóng áp dụng TTNT. Các công ty khởi nghiệp như Zest Finance, Insurify, WeCash, CreditVidya và Aire cũng theo xu hướng này. Các công ty dịch vụ tài chính còn kết hợp nhiều phương thức máy học khác nhau. Ví dụ, công ty khởi nghiệp công nghệ QuantCube của Pháp phân tích hàng tỷ điểm dữ liệu được thu thập từ hơn 40 quốc gia. Đơn vị này sử dụng các phương tiện như xử lý ngôn ngữ, học sâu, lý thuyết đồ thị để triển khai các giải pháp TTNT hỗ trợ các công ty tài chính đưa ra quyết định.

Ứng dụng TTNT trong lĩnh vực tài chính mang lại nhiều lợi ích to lớn như cải thiện trải nghiệm của khách hàng, xác định các cơ hội đầu tư thông minh và tạo lòng tin cho khách hàng thông qua cung cấp các điều kiện tốt hơn. Tuy nhiên, vẫn còn có những vấn đề chính sách liên quan đến việc đảm bảo tính chính xác và ngăn ngừa phân biệt đối xử, cũng như tác động rộng lớn của tự động hóa đến việc làm.

*Ứng dụng TTNT trong tiếp thị và quảng cáo*

TTNT đang ảnh hưởng đến hoạt động tiếp thị và quảng cáo theo nhiều cách. Về bản chất, TTNT cho phép cá nhân hóa trải nghiệm trực tuyến, giúp hiển thị nội dung mà người dùng quan tâm nhiều nhất. Sự phát triển của máy học cùng với số lượng lớn dữ liệu được tạo ra, cho phép các nhà quảng cáo hiện thực hóa mục tiêu của các chiến dịch quảng cáo. Các nhà quảng cáo có thể phân phối quảng cáo động và được cá nhân hóa cho người tiêu dùng trên quy mô chưa từng có. Quảng cáo được cá nhân hóa mang lại lợi ích đáng kể cho doanh nghiệp và người tiêu dùng. Đối với doanh nghiệp, quảng cáo làm tăng doanh số bán hàng và lợi tức đầu tư của các chiến dịch tiếp thị. Đối với người tiêu dùng, các dịch vụ trực tuyến được tài trợ bởi doanh thu quảng cáo, thường được cung cấp miễn phí cho người dùng cuối và làm giảm đáng

kể chi phí nghiên cứu của người tiêu dùng.

***Ứng dụng TTNT trong khoa học***

Thách thức toàn cầu hiện nay rất đa dạng từ biến đổi khí hậu đến tình trạng kháng kháng sinh của vi khuẩn. Giải pháp xử lý những thách thức này đòi hỏi sự gia tăng tri thức khoa học. TTNT có thể làm tăng năng suất khoa học vào thời điểm mà một số học giả tuyên bố những ý tưởng mới ngày càng khó tìm kiếm. TTNT cũng hứa hẹn cải thiện năng suất nghiên cứu ngay cả khi áp lực lên ngân sách nghiên cứu công gia tăng. Hiểu biết khoa học phụ thuộc vào việc rút ra kiến thức từ khối lượng lớn dữ liệu khoa học được tạo ra bởi các thiết bị khoa học mới. Trong bối cảnh đó, không thể thiếu việc ứng dụng TTNT trong khoa học. Hơn nữa, tri thức TTNT sẽ được bổ sung cho các nhà khoa học thông qua số lượng lớn các bài báo khoa học đang tăng nhanh.

Việc sử dụng TTNT trong khoa học cũng tạo thuận lợi cho các hình thức khám phá mới và nâng cao khả năng tái tạo của nghiên cứu khoa học. Ứng dụng của TTNT trong khoa học và ngành công nghiệp ngày càng nhiều và có ý nghĩa. Ví dụ, TTNT đã dự đoán hành vi của các hệ thống hỗn loạn, giải quyết các vấn đề tính toán phức tạp trong di truyền học, cải thiện chất lượng hình ảnh thiên văn và giúp khám phá các quy tắc tổng hợp hóa học. Ngoài ra, TTNT đang được triển khai trong các chức năng từ phân tích bộ dữ liệu lớn, đưa ra giả thuyết, hiểu và phân tích tài liệu khoa học cho đến hỗ trợ thu thập dữ liệu, thiết kế thử nghiệm và tự thử nghiệm.

***Ứng dụng TTNT trong y học***

Ứng dụng TTNT trong chăm sóc sức khỏe và dược phẩm giúp phát hiện sớm tình trạng sức khỏe, cung cấp các dịch vụ phòng ngừa, tối ưu hóa việc ra quyết định lâm sàng và khám phá các liệu pháp và thuốc mới. Ứng dụng TTNT tạo điều kiện thuận lợi cho chăm sóc sức khỏe được cá nhân hóa và y học chính xác, đồng thời hỗ trợ các công cụ tự theo dõi, ứng dụng và giám sát. TTNT trong chăm sóc sức khỏe mang lại những lợi ích tiềm năng về chất lượng và chi phí chăm sóc, nhưng làm nảy sinh nhiều vấn đề như liên quan đến quyền truy cập vào dữ liệu (sức khỏe) và quyền riêng tư.

Ở khía cạnh nào đó, lĩnh vực y tế là nền tảng lý tưởng cho các hệ thống TTNT và là minh họa hoàn hảo cho những tác động tiềm tàng của nó. Ngành công nghiệp thâm dụng tri thức phụ thuộc vào dữ liệu và phân tích để cải thiện các liệu pháp. Đã có sự phát triển vượt bậc về phạm vi thông tin được thu thập, bao gồm dữ liệu lâm sàng, di truyền, hành vi và môi trường. Hàng ngày, các chuyên gia chăm sóc sức khỏe, nhà nghiên cứu y sinh và bệnh nhân tạo ra lượng lớn dữ liệu từ rất nhiều thiết bị, bao gồm hồ sơ sức khỏe điện tử, máy giải trình tự bộ gen, hình ảnh y tế độ phân giải cao, ứng dụng điện thoại thông minh và cảm biến thông dụng, cũng như các thiết bị IoT theo dõi sức khỏe bệnh nhân.

Nếu được đưa vào sử dụng, dữ liệu TTNT có giá trị lớn đối với công tác nghiên cứu và chăm sóc sức khỏe. Thật vậy, ngành y tế tại các quốc gia đang trải qua sự chuyển đổi sâu sắc khi khai thác các cơ hội do công nghệ thông tin và truyền thông mang lại. Các mục tiêu chính định hình quá trình chuyển đổi này, bao gồm cải thiện hiệu quả, năng suất và chất lượng chăm sóc.

***Ứng dụng TTNT trong tư pháp hình sự***

TTNT có tiềm năng cải thiện khả năng tiếp cận công lý và thúc đẩy quá trình xét xử công bằng và hiệu quả. Tuy nhiên, vẫn có lo ngại về những thách thức của hệ thống TTNT đối với sự tham gia của công dân, tính minh bạch, phẩm giá, quyền riêng tư và tự do.

TTNT được sử dụng ngày càng rộng rãi trong các giai đoạn khác nhau của thủ tục hình sự. Những phạm vi này bao gồm dự đoán nơi tội phạm có thể xuất hiện và kết quả của một thủ tục hình sự để thực hiện đánh giá rủi ro đối với bị cáo, cũng như góp phần quản lý hiệu quả quy trình. Dù nhiều ứng dụng TTNT vẫn đang được thử nghiệm, nhưng một số sản phẩm tiên tiến có khả năng dự báo đã được sử dụng trong quản lý tư pháp và thực thi pháp luật. AI có thể cải thiện khả năng tạo kết nối, phát hiện mẫu, ngăn chặn và giải quyết vấn đề tội phạm.

***Ứng dụng TTNT trong an ninh***

TTNT có triển vọng giải quyết những thách thức phức tạp về an ninh vật lý (physical security) và an ninh kỹ thuật số. Năm 2018, chi tiêu quốc phòng toàn cầu ước tính đạt 1,67 nghìn tỷ USD, tăng 3,3% so với năm 2017. Tuy nhiên, chi tiêu cho an ninh không chỉ giới hạn ở khu vực công. Khu vực tư nhân trên toàn thế giới dự kiến sẽ chi 96 tỷ USD để ứng phó với các rủi ro an ninh trong năm 2018, tăng 8% so với năm 2017. Các cuộc tấn công an ninh kỹ thuật số quy mô lớn gần đây đã nâng cao nhận thức của xã hội về an ninh kỹ thuật số. Vi phạm dữ liệu có thể gây ra những hậu quả sâu rộng về kinh tế, xã hội và an ninh quốc gia. Trong bối cảnh đó, phía nhà nước và tư nhân đều sử dụng các công nghệ TTNT để điều chỉnh cho phù hợp với bối cảnh an ninh đang thay đổi trên toàn thế giới. Hai lĩnh vực liên quan đến an ninh đang được tiếp nhận đặc biệt nhanh chóng: đó là an ninh kỹ thuật số và giám sát.

TTNT đã được sử dụng rộng rãi trong các ứng dụng an ninh kỹ thuật số như an ninh mạng, phát hiện bất thường, tự động hóa hoạt động an ninh và phát hiện mối đe dọa. Đồng thời, những hành vi sử dụng TTNT ác ý được cho là sẽ gia tăng, bao gồm xác định các lỗ hổng phần mềm để tấn công hệ thống, mạng và dữ liệu. Kết quả sẽ gây rủi ro an ninh kỹ thuật số.

Hạ tầng kỹ thuật số đang phát triển tại các thành phố, đặc biệt trong lĩnh vực giám sát ở đó, nhiều công cụ sử dụng TTNT, đang được cài đặt để tăng cường an ninh chung. Ví dụ, camera thông minh có thể phát hiện một cuộc chiến. Bộ định vị súng tự động báo cáo các cảnh quay đã ghi và cung cấp vị trí chính xác.

***Ứng dụng TTNT trong khu vực công***

Tiềm năng của TTNT đối với các cơ quan hành chính nhà nước rất đa dạng. Sự phát triển của công nghệ TTNT đã tác động đến cách thức hoạt động của khu vực công và thiết kế các chính sách để phục vụ người dân và doanh nghiệp. Các ứng dụng liên quan đến các lĩnh vực như y tế, giao thông vận tải và dịch vụ an ninh.

Chính phủ các nước OECD đang thử nghiệm và thực hiện các dự án khai thác TTNT để đáp ứng tốt hơn nhu cầu của người sử dụng dịch vụ công. Ngoài ra, các nước OECD cũng tăng cường khả năng quản lý nguồn lực của mình (như tiết kiệm thời gian công chức dành cho việc hỗ trợ khách hàng và các công việc hành chính). Các công cụ TTNT có thể nâng cao hiệu quả và chất lượng của nhiều thủ tục trong

khu vực công. Nếu được thiết kế và triển khai chính xác, các công nghệ TTNT có thể được tích hợp vào toàn bộ quá trình hoạch định chính sách, hỗ trợ cải cách khu vực công và cải thiện năng suất khu vực công.

Chính phủ một số nước đã triển khai hệ thống TTNT để tăng cường các chương trình phúc lợi xã hội. Ví dụ, TTNT giúp đạt được mức độ thống kê tối ưu tại các địa điểm cung cấp dịch vụ xã hội và y tế, thông qua các công nghệ máy học phân tích dữ liệu giao dịch và đưa ra các dự đoán bổ sung ngày càng chính xác. Điều này sẽ tạo điều kiện thuận lợi cho việc dự báo và xây dựng chính sách. Một ví dụ khác là, các thuật toán TTNT đang giúp chính phủ Vương quốc Anh phát hiện gian lận trong các yêu cầu trợ cấp xã hội.

***Ứng dụng TTNT sử dụng thực tế ảo và tăng cường***

Các công ty đang sử dụng công nghệ TTNT và các tác vụ nhận dạng hình ảnh cấp cao như phân loại hình ảnh và phát hiện đối tượng để phát triển phần cứng thực tế ảo và tăng cường. Các lợi ích được tạo ra bao gồm cung cấp trải nghiệm phong phú, đào tạo và giáo dục, hỗ trợ người khuyết tật và cung cấp dịch vụ giải trí. Thực tế ảo và tăng cường đã phát triển đáng kể từ khi Ivan Sutherland phát triển nguyên mẫu tai nghe thực tế ảo đầu tiên vào năm 1968 để xem hình ảnh 3D. Mẫu tai nghe quá nặng để có thể đeo nên đã được gắn trên trần nhà. Các công ty thực tế ảo hiện cung cấp trải nghiệm phát video $360^0$ bằng tai nghe nhẹ hơn nhiều. Pokemon GO đã thu hút sự chú ý của người tiêu dùng đến thực tế tăng cường vào năm 2016 và kỳ vọng vẫn ở mức cao. Các ứng dụng tích hợp TTNT đã có mặt trên thị trường. IKEA cung cấp một ứng dụng di động cho phép khách hàng xem một một món đồ nội thất trông ra sao và có phù hợp với không gian nhất định hay không với độ chính xác lên đến 1 mm. Một số công ty công nghệ đang phát triển các ứng dụng dành cho người khiếm thị.

## II. CÁC CÂN NHẮC CHÍNH SÁCH CÔNG

### 2.1. Yêu cầu chính sách đối với TTNT

#### *2.1.1. TTNT lấy con người làm trung tâm*

TTNT có ảnh hưởng ngày càng lớnkhi công nghệ được phổ biến. Tác động tiềm tàng từ các dự báo, khuyến nghị hoặc quyết định của TTNT đưa ra đến cuộc sống của con người, cũng gia tăng. Các cộng đồng kỹ thuật, kinh doanh và chính sách đang tích cực tìm cách để TTNT đặt con người vào vị trí trung tâm và trở nên đáng tin cậy, tăng tối đa lợi ích, giảm thiểu rủi ro và được xã hội chấp nhận rộng rãi.

***Hộp 1***. **Hệ thống "hộp đen" TTNT đặt ra những thách thức mới bắt nguồn từ những tiến bộ công nghệ trước đây**

Mạng nơ-ron thường được gọi là "hộp đen". Dù hành vi của các hệ thống này có thể được giám sát, nhưng thuật ngữ "hộp đen" phản ánh sự khác biệt lớn giữa khả năng giám sát các công nghệ trước đây với mạng nơ-ron. Mạng nơ-ron lặp lại dữ liệu mà chúng được đào tạo. Mạng nơ-ron phát hiện các mối tương quan xác suất phức tạp, đa biến, là một phần của mô hình mà hệ thống này xây dựng. Tuy nhiên, mạng nơ-ron không chỉ ra cách dữ liệu liên kết với nhau. Dữ liệu quá phức tạp để trí óc con người có thể hiểu được. TTNT có đặc điểm khác biệt so với những tiến bộ công nghệ trước đây và ảnh hưởng đến tính minh bạch và trách nhiệm giải trình bao gồm:

• *Khả năng phát hiện*: Các thuật toán dựa vào quy tắc có thể được đọc và kiểm tra

theo từng quy tắc để tìm ra vài loại lỗi tương đối đơn giản. Ngược lại, một số loại hệ thống máy học, đặc biệt là mạng nơ-ron, chỉ đơn giản là các mối quan hệ toán học trừu tượng giữa các thừa số. Chúng có thể rất phức tạp và khó hiểu ngay cả với các chuyên gia lập trình và đào tạo.

• *Bản chất phát triển*: Một số hệ thống máy học lặp lại và phát triển theo thời gian và thậm chí có thể thay đổi hành vi của chính chúng theo những cách bất ngờ.

• *Không dễ dàng lặp lại*: Một dự báo hoặc quyết định cụ thể chỉ có được khi hệ thống máy học được cung cấp các điều kiện và dữ liệu cụ thể, không nhất thiết phải lặp lại.

• Căng thẳng gia tăng trong việc bảo vệ dữ liệu cá nhân và nhạy cảm:

o *Tham khảo*: Ngay cả khi không có dữ liệu nhạy cảm hoặc được bảo vệ, hệ thống TTNT vẫn có thể suy ra các dữ liệu và mối tương quan này từ các biến proxy không mang tính cá nhân hoặc nhạy cảm như lịch sử mua hàng hoặc vị trí.

o *Biến proxy không phù hợp*: Các phương pháp tiếp cận chính sách và kỹ thuật với quyền riêng tư và không phân biệt có xu hướng giảm thiểu dữ liệu được thu thập, cấm sử dụng một số dữ liệu nhất định hoặc xóa dữ liệu để ngăn chặn sử dụng chúng. Nhưng hệ thống TTNT có thể đưa ra dự đoán dựa trên dữ liệu proxy có quan hệ chặt chẽ với dữ liệu bị cấm và không được thu thập. Hơn nữa, cách duy nhất để phát hiện những proxy này là thu thập dữ liệu cá nhân hoặc nhạy cảm như chủng tộc. Nếu những dữ liệu đó được thu thập, thì phải đảm bảo chúng chỉ được sử dụng theo những cách phù hợp.

o *Nghịch lý dữ liệu-quyền riêng tư*: Đối với nhiều hệ thống TTNT, nhiều dữ liệu đào tạo có thể cải thiện độ chính xác của các dự đoán bằng TTNT và giảm nguy cơ sai lệch từ các mẫu bị sai lệch. Tuy nhiên, càng nhiều dữ liệu được thu thập, thì rủi ro về quyền riêng tư đối với những người có dữ liệu được thu thập càng lớn.

Một số loại TTNT, thường được gọi là "hộp đen", đặt ra những thách thức mới so với những tiến bộ công nghệ trước đây (Hộp 1). Vì thế, OECD đã xác định những ưu tiên chính để TTNT lấy con người làm trung tâm. Trước hết, TTNT phải góp phần thúc đẩy tăng trưởng và phúc lợi bền vững và toàn diện. Thứ hai, TTNT phải tôn trọng các giá trị tập trung vào con người và sự công bằng. Thứ ba, việc sử dụng TTNT và cách thức hoạt động của các hệ thống TTNT phải minh bạch. Thứ tư, hệ thống TTNT phải mạnh mẽ và an toàn. Thứ năm là trách nhiệm giải trình về kết quả dự báo của TTNT và các quyết định tiếp theo.

### *2.1.2. Tăng trưởng và phúc lợi bền vững và toàn diện*

*TTNT có tiềm năng lớn thúc đẩy chương trình hướng tới thực hiện Mục tiêu Phát triển bền vững*

TTNT có thể được khai thác vì lợi ích xã hội và thúc đẩy thực hiện các Mục tiêu Phát triển bền vững của LHQ (SDGs) trong các lĩnh vực như giáo dục, y tế, giao thông, nông nghiệp và các thành phố bền vững. Nhiều tổ chức công và tư, bao gồm Ngân hàng thế giới, một số cơ quan của LHQ và OECD, đang sử dụng TTNT để thúc đẩy các Mục tiêu phát triển bền vững.

*Ưu tiên phát triển TTNT một cách công bằng và toàn diện*

Đây là ưu tiên đặc biệt đúng đắn trong bối cảnh lo ngại TTNT làm trầm trọng thêm tình trạng bất bình đẳng hoặc gây gia tăng chia rẽ hiện có trong và giữa các nước phát triển và nước đang phát triển. Nguyên nhân là do nguồn lực TTNT (công nghệ, kỹ năng, bộ dữ liệu và công suất tính toán của TTNT) tập trung tại một số công

ty và quốc gia. Ngoài ra, còn có một số lo ngại TTNT gây tác động khác nhau đến các nhóm dân cư dễ bị tổn thương và thiểu số, bao gồm những người có trình độ học vấn thấp, tay nghề thấp, phụ nữ và người cao tuổi, đặc biệt là tại các quốc gia có thu nhập thấp và trung bình. Trung tâm Nghiên cứu phát triển quốc tế của Canada gần đây đã khuyến nghị lập quỹ phát triển TTNT toàn cầu. Từ đó sẽ thành lập các Trung tâm TTNT xuất sắc tại các quốc gia có thu nhập thấp và trung bình để hỗ trợ thiết kế và thực hiện chính sách toàn diện dựa vào bằng chứng. Mục tiêu là đảm bảo phân bổ đồng đều các lợi ích của TTNT và tạo sự bình đẳng hơn cho xã hội. Các sáng kiến TTNT toàn diện nhằm đảm bảo chia sẻ rộng rãi lợi ích kinh tế từ TTNT trong xã hội.

TTNT bền vững và toàn diện là lĩnh vực trọng tâm của các quốc gia như Ấn Độ, các công ty như Microsoft và các nhóm nghiên cứu như Trung tâm Berkman Klein tại Harvard. Ví dụ, Microsoft đã khởi động các dự án như ứng dụng di động See AI hỗ trợ người khiếm thị. Ứng dụng này quét và nhận diện tất cả các yếu tố xung quanh một người và cung cấp mô tả bằng lời nói. Microsoft cũng đang đầu tư 2 triệu USD cho các sáng kiến đủ điều kiện để khai thác TTNT nhằm giải quyết các thách thức bền vững như đa dạng sinh học và biến đổi khí hậu.

### *2.1.3. Các giá trị lấy con người làm trung tâm và công bằng*

#### *Quyền con người và quy tắc đạo đức*

TTNT có thể hỗ trợ thực hiện quyền con người, nhưng lại gây ra những rủi ro mới do quyền con người bị cố ý hoặc vô tình vi phạm. Luật nhân quyền, cùng với các cấu trúc pháp lý và thể chế khác, cũng là một trong những công cụ đảm bảo TTNT lấy con người làm trung tâm.

Với tiềm năng ứng dụng rộng rãi, TTNT hứa hẹn sẽ thúc đẩy việc bảo vệ và thực hiện các quyền con người. Ví dụ, TTNT được sử dụng để phân tích các mô hình khan hiếm thực phẩm để chống nạn đói, cải thiện khả năng chẩn đoán và điều trị trong y học hoặc làm cho các dịch vụ y tế trở nên phổ biến và dễ tiếp cận, đồng thời làm sáng tỏ tình trạng phân biệt đối xử.

TTNT cũng đặt ra một số thách thức về quyền con người, thường được phản ánh trong các cuộc thảo luận về TTNT và đạo đức trên phạm vi rộng. Hệ thống TTNT cụ thể có thể được sử dụng để vi phạm nhân quyền một cách vô tình hoặc cố ý. Các tác động tình cờ được tập trung nhiều hơn. Ví dụ, các thuật toán máy học dự báo tái phạm có thể sai lệch mà không bị phát hiện. Tuy nhiên, các công nghệ TTNT còn liên quan đến việc cố ý vi phạm nhân quyền. Chẳng hạn, việc sử dụng công nghệ TTNT để tìm ra những người bất đồng chính kiến và hạn chế quyền tự do ngôn luận của cá nhân hoặc khả năng tham gia vào đời sống chính trị. Trong những trường hợp này, bản thân vi phạm thường không chỉ xảy ra với việc sử dụng TTNT, nhưng có thể trầm trọng hơn do sự tinh vi và hiệu quả của TTNT.

Sử dụng TTNT cũng gây khó khăn đặc biệt trong các tình huống trong đó các tác động đến quyền con người không cố ý hoặc khó phát hiện. Lý do có thể là vì sử dụng dữ liệu đào tạo kém chất lượng, thiết kế hệ thống hoặc các tương tác phức tạp giữa hệ thống TTNT và môi trường của nó. Ví dụ về thuật toán làm trầm trọng thêm lời nói căm thù hoặc kích động bạo lực trên mạng hay việc vô tình khuếch đại tin tức giả tác động đến quyền tham gia vào các vấn đề chính trị và công vụ.

*Khung nhân quyền được bổ sung bởi các quy tắc đạo đức TTNT*

Các quy tắc đạo đức có thể xử lý nguy cơ TTNT không hoạt động theo cách lấy con người làm trung tâm hoặc phù hợp với các giá trị của con người. Cả các công ty tư nhân và chính phủ đều đã áp dụng rất nhiều quy tắc đạo đức TTNT.

Ví dụ, DeepMind thuộc sở hữu của Google đã thành lập đơn vị DeepMind Ethics & Society vào tháng 10/2017 để giúp các chuyên gia công nghệ hiểu được ý nghĩa đạo đức trong công việc của họ và giúp xã hội quyết định cách TTNT mang lại lợi ích. Đơn vị này cũng sẽ tài trợ cho các nghiên cứu bên ngoài về sự cố sai lệch thuật toán, tương lai việc làm… Bản thân Google đã công bố bộ nguyên tắc đạo đức định hướng nghiên cứu, phát triển sản phẩm và đưa ra các quyết định kinh doanh. Google đã xuất bản sách trắng về quản trị TTNT, xác định các vấn đề cần làm sáng tỏ với các chính phủ và xã hội dân sự.

Luật nhân quyền, cùng với các cơ chế thể chế và cấu trúc rộng lớn, cung cấp định hướng và nền tảng để đảm bảo việc phát triển và sử dụng TTNT trong xã hội có đạo đức và lấy con người làm trung tâm.

***Bảo vệ dữ liệu cá nhân***

TTNT có thể kết nối ngày càng nhiều bộ dữ liệu. Dữ liệu lưu trữ riêng biệt đã từng được coi không mang tính cá nhân (hoặc đã được bỏ mã định dạng cá nhân, tức là "không xác định danh tính"). Tuy nhiên, với TTNT, dữ liệu phi cá nhân liên quan đến dữ liệu khác và trùng khớp với các cá nhân cụ thể, nên mang tính cá nhân (hoặc "được nhận dạng lại"). Do đó, tương quan thuật toán làm giảm sự khác biệt giữa dữ liệu cá nhân và dữ liệu khác. Dữ liệu phi cá nhân được sử dụng ngày càng nhiều để xác định lại các cá nhân hoặc suy ra thông tin nhạy cảm về họ, ngoài những thông tin đã bị cố ý tiết lộ từ đầu. Ví dụ, năm 2007, các nhà nghiên cứu đã sử dụng dữ liệu ẩn danh được báo cáo để liên kết danh sách phim cho thuê của Netflix với các bài đánh giá được đăng trên IMDB. Theo cách đó, họ xác định được các khách hàng cá nhân và truy cập vào toàn bộ lịch sử xem phim của họ. Nhờ có nhiều dữ liệu được thu thập và những cải tiến công nghệ, có thể thực hiện được ngày càng nhiều các liên kết này. Vì thế, rất khó đánh giá dữ liệu nào được xem xét không mang tính cá nhân.

Ngày càng khó phân biệt giữa dữ liệu nhạy cảm và không nhạy cảm trong Quy định chung về bảo vệ dữ liệu của Liên minh châu Âu (GDPR). Một số thuật toán có thể suy ra thông tin nhạy cảm từ dữ liệu "không nhạy cảm", như đánh giá trạng thái cảm xúc của cá nhân dựa vào kiểu gõ bàn phím. Việc sử dụng TTNT để xác định hoặc xác định lại dữ liệu ban đầu không phải là dữ liệu cá nhân hoặc không được xác định danh tính, cũng là vấn đề pháp lý. Các khuôn khổ bảo vệ như Khuyến nghị của Hội đồng OECD liên quan đến các Hướng dẫn về việc bảo vệ quyền riêng tư và luồng dữ liệu cá nhân xuyên biên giới, áp dụng cho dữ liệu cá nhân. Tuy nhiên, chưa rõ liệu chúng có áp dụng cho dữ liệu trong một số trường hợp có thể nhận dạng hay không.

***Công bằng và đạo đức***

Cho đến nay, các sáng kiến chính sách TTNT có đặc trưng nổi bật là đạo đức, công bằng và/hoặc công lý. Tuy nhiên, nhiều lo ngại là các thuật toán máy học có xu hướng phản ánh và lặp lại các thành kiến tiềm ẩn trong dữ liệu đào tạo của chúng như thành kiến về chủng tộc và các liên kết rập khuôn. Vì các sản phẩm công nghệ thường thể hiện các giá trị xã hội, nên các cuộc thảo luận về sự công bằng cần xác định rõ công nghệ phục vụ cho xã hội nào, ai nên được bảo vệ và với những giá trị cốt lõi nào.

### *2.1.4. Tính minh bạch và khả năng giải thích*

Ý nghĩa về mặt kỹ thuật và chính sách của thuật ngữ "minh bạch" có sự khác biệt. Đối với các nhà hoạch định chính sách, tính minh bạch thường tập trung vào cách đưa ra quyết định, ai tham gia vào quá trình và các yếu tố được sử dụng để đưa ra quyết định. Từ góc độ này, các biện pháp minh bạch có thể tiết lộ cách TTNT đang được sử dụng trong dự đoán, đề xuất hoặc ra quyết định. Đối với các chuyên gia công nghệ, tính minh bạch của hệ thống TTNT tập trung chủ yếu vào vấn đề quy trình. Nó cho phép mọi người hiểu cách một hệ thống TTNT được phát triển, đào tạo và triển khai, cũng như các yếu tố tác động đến một dự đoán hoặc quyết định cụ thể. Nó thường không bao gồm việc chia sẻ mã hoặc bộ dữ liệu cụ thể. Trong nhiều trường hợp, hệ thống TTNT quá phức tạp để các yếu tố này tạo ra sự minh bạch. Hơn nữa, việc chia sẻ mã hoặc bộ dữ liệu cụ thể có thể tiết lộ bí mật thương mại hoặc dữ liệu người dùng nhạy cảm. Nhìn chung, nhận thức và hiểu biết về các quy trình lập luận của TTNT được coi là quan trọng để TTNT trở nên phổ biến và hữu ích.

### *2.1.5. Tính mạnh mẽ, bảo mật và an toàn*

***Tính mạnh mẽ và bảo mật***

Tính mạnh mẽ được hiểu là khả năng chịu đựng hoặc vượt qua các điều kiện bất lợi, bao gồm cả các rủi ro bảo mật kỹ thuật số. Hệ thống TTNT an toàn được hiểu là các hệ thống không gây rủi ro an toàn bất hợp lý trong điều kiện sử dụng bình thường hoặc có thể dự báo hoặc sử dụng sai trong suốt vòng đời của chúng. Các vấn đề về tính mạnh mẽ và an toàn của TTNT có liên quan với nhau. Ví dụ, bảo mật kỹ thuật số có thể ảnh hưởng đến sự an toàn của sản phẩm nếu các sản phẩm được kết nối như ô tô không người lái hoặc thiết bị gia dụng được TTNT hỗ trợ không đủ an toàn; tin tặc có thể kiểm soát chúng và thay đổi cài đặt từ xa.

Tác hại tiềm tàng của hệ thống TTNT cần được cân bằng với chi phí tăng tính minh bạch và trách nhiệm giải trình trong hệ thống TTNT. Các tác hại có thể bao gồm rủi ro với quyền con người, quyền riêng tư, sự công bằng và mạnh mẽ. Nhưng không phải mọi hành vi sử dụng TTNT đều gây rủi ro như nhau.

Các tổ chức sử dụng quản lý rủi ro để xác định, đánh giá, ưu tiên và xử lý các rủi ro tiềm ẩn có thể ảnh hưởng xấu đến hành vi và kết quả của hệ thống TTNT. Cách tiếp cận này cũng có thể được sử dụng để xác định rủi ro cho các bên liên quan khác nhau và xác định cách xử lý các rủi ro này trong suốt vòng đời của hệ thống TTNT.

Các chủ thể - người đóng vai trò tích cực trong vòng đời của hệ thống TTNT - đánh giá và giảm thiểu rủi ro trong toàn bộ hệ thống TTNT, cũng như trong từng giai đoạn của vòng đời hệ thống. Quản lý rủi ro bao gồm các bước sau, có mức độ liên quan khác nhau tùy thuộc vào giai đoạn của vòng đời hệ thống TTNT: xây dựng mục tiêu; xác định các bên liên quan và chủ thể; đánh giá rủi ro; giảm thiểu nguy cơ; áp dụng chiến lược giảm thiểu nguy cơ; theo dõi, đánh giá và phản hồi.

***Sự an toàn***

Các sản phẩm được tích hợp công nghệ TTNT đang phát triển nhanh chóng và đa dạng từ rôbốt và ô tô không người lái đến các sản phẩm và dịch vụ tiêu dùng hàng ngày như thiết bị thông minh và hệ thống an ninh cho nhà thông minh. Các sản phẩm này mang lại lợi ích an toàn to lớn, nhưng lại đặt ra những thách thức thực tế và pháp lý mới đối với các khuôn khổ an toàn của sản phẩm. Các khuôn khổ an toàn có xu hướng điều chỉnh các sản phẩm phần cứng "đã hoàn thiện" hơn là phần mềm, trong khi một số sản phẩm phần mềm TTNT học hỏi và phát triển trong suốt vòng đời của chúng. Các sản phẩm AI cũng có thể "tự chủ" hoặc bán tự chủ, tức là đưa ra và thực hiện các quyết định mà không cần hoặc ít cần sự can thiệp của con người.

Các ứng dụng TTNT khác nhau sẽ cần có các phản ứng chính sách phù hợp. Cụ thể, hệ thống TTNT cần có bốn sự cân nhắc. Đầu tiên là xem xét phương thức tốt nhất để đảm bảo an toàn cho các sản phẩm. Cụ thể, các sản phẩm không được gây rủi ro an toàn bất hợp lý trong điều kiện sử dụng bình thường hoặc theo dự báo hoặc sử dụng sai trong toàn bộ vòng đời của chúng. Thứ hai nên xem xét ai phải chịu trách nhiệm pháp lý và ở mức độ nào đối với tác hại do hệ thống TTNT gây ra. Đồng thời nên xem xét bên nào có thể đóng góp vào sự an toàn của máy móc tự động. Các bên có thể là người dùng, nhà sản xuất sản phẩm và cảm biến, nhà sản xuất phần mềm, nhà thiết kế, nhà cung cấp hạ tầng và các công ty phân tích dữ liệu. Thứ ba nên cân nhắc lựa chọn các nguyên tắc về trách nhiệm pháp lý, bao gồm trách nhiệm pháp lý nghiêm ngặt, trách nhiệm pháp lý do lỗi và vai trò của bảo hiểm. Thứ tư nên xem xét cách thực thi luật, yếu tố "bị lỗi" trong một sản phẩm TTNT, nghĩa vụ chứng minh và những biện pháp khắc phục có sẵn.

### *2.1.6. Trách nhiệm giải trình*

Trách nhiệm giải trình tập trung gắn trách nhiệm cho các tổ chức hoặc cá nhân để hệ thống TTNT hoạt động đúng cách. Tiêu chí về trách nhiệm giải trình bao gồm tôn trọng các nguyên tắc giá trị con người và công bằng, tính minh bạch, tính mạnh mẽ và an toàn. Trách nhiệm giải trình phụ thuộc vào vai trò cá nhân của các chủ thể TTNT, bối cảnh và mức độ phát triển. Đối với các nhà hoạch định chính sách, trách nhiệm giải trình phụ thuộc vào các cơ chế thực hiện một số chức năng. Các cơ chế xác định bên chịu trách nhiệm về một đề xuất hoặc quyết định cụ thể. Bên chịu trách nhiệm sẽ điều chỉnh khuyến nghị hoặc quyết định trước khi nó được thực hiện. Ngoài ra, họ cũng có thể phản đối hoặc kháng nghị quyết định hoặc thậm chí thách thức hệ thống chịu trách nhiệm đưa ra quyết định

Trên thực tế, trách nhiệm giải trình của hệ thống TTNT thường phụ thuộc vào phương thức hoạt động của hệ thống so với các chỉ số về mức độ chính xác hoặc hiệu

quả. Các thước đo bao gồm các chỉ số cho mục tiêu công bằng, an toàn và mạnh mẽ. Tuy nhiên, các chỉ số này có xu hướng ít được sử dụng hơn các thước đo về hiệu quả hoặc độ chính xác.

### 2.2. Chính sách thúc đẩy sự phát triển của TTNT

Các chính sách quốc gia thúc đẩy phát triển hệ thống TTNT đáng tin cậy, có thể mang lại kết quả có lợi và công bằng cho con người và hành tinh, đặc biệt là trong các lĩnh vực triển vọng nhưng chưa nhận được các khoản đầu tư được định hướng bởi thị trường. Xây dựng môi trường chính sách thuận lợi cho hệ thống TTNT đáng tin cậy thúc đẩy đầu tư công và tư cho NC&PT TTNT và trang bị cho con người những kỹ năng cần thiết để thành công trong công việc. Dưới đây là bốn lĩnh vực chính sách quan trọng thúc đẩy và phát triển hệ thống TTNT đáng tin cậy.

#### *2.2.1. Đầu tư nghiên cứu và phát triển TTNT*

TTNT đòi hỏi các nhà hoạch định chính sách phải xem xét lại sự tham gia của chính phủ ở mức độ phù hợp vào nghiên cứu TTNT để giải quyết các thách thức xã hội. Ngoài ra, các tổ chức nghiên cứu trong mọi lĩnh vực sẽ cần đến hệ thống TTNT có khả năng duy trì năng lực cạnh tranh, đặc biệt là trong các lĩnh vực khoa học y sinh và khoa học đời sống. Các công cụ mới như nền tảng chia sẻ dữ liệu và hạ tầng siêu máy tính hỗ trợ nghiên cứu TTNT và cần có các khoản đầu tư mới. Ví dụ, Nhật Bản đầu tư hơn 120 triệu USD mỗi năm để xây dựng hạ tầng máy tính hiệu năng cao cho các trường đại học và trung tâm nghiên cứu công.

TTNT được coi là công nghệ đa dụng với tiềm năng tác động đến nhiều ngành công nghiệp. Ngoài ra, TTNT cũng được coi là "phát minh của phương pháp phát minh" được sử dụng rộng rãi bởi các nhà khoa học và nhà phát minh để thúc đẩy ĐMST. Các ngành công nghiệp hoàn toàn mới có thể được tạo ra trên nền tảng của những đột phá khoa học có sự hỗ trợ của TTNT. Điều đó nhấn mạnh tầm quan trọng của nghiên cứu cơ bản và xem xét những phạm vi lâu dài trong chính sách nghiên cứu.

#### *2.2.2. Hỗ trợ hệ sinh thái kỹ thuật số cho TTNT*

***Công nghệ và hạ tầng TTNT***

Trong những năm gần đây, công nghệ TTNT đã đạt được những thành tựu lớn là nhờ độ chín muồi của các kỹ thuật mô hình thống kê như mạng nơ-ron và đặc biệt là mạng nơ-ron sâu (được gọi là học sâu). Nhiều công cụ quản lý và sử dụng TTNT tồn tại dưới dạng tài nguyên mã nguồn mở trong khu vực công. Chúng được sử dụng để chỉnh sửa lỗi phần mềm. Các công cụ này bao gồm TensorFlow (Google), Michelangelo (Uber) và Bộ công cụ nhận thức (Microsoft). Một số công ty và nhà nghiên cứu cũng chia sẻ công khai bộ dữ liệu đào tạo và công cụ đào tạo được quản lý để hỗ trợ phổ biến công nghệ TTNT.

TTNT có được những thành tựu gần đây một phần là do tốc độ máy tính tăng theo cấp số nhân theo định luật Moore (tức là mật độ bóng bán dẫn trong một mạch tích hợp tăng gấp đôi sau khoảng 2 năm). Hai yếu tố này kết hợp cho phép các thuật toán TTNT xử lý nhanh khối lượng lớn dữ liệu. Khi các dự án TTNT chuyển từ khái niệm sang ứng dụng thương mại, thường sẽ cần có nguồn tài nguyên là đơn vị xử lý đồ họa và điện toán đám mây chuyên biệt và đắt đỏ. Xu hướng trong các hệ thống

TTNT tiếp tục cho thấy sự phát triển phi thường về năng lực tính toán cần thiết. Theo ước tính, thử nghiệm AlphaGo Zero lớn nhất gần đây cần có năng lực tính toán mạnh gấp 300.000 lần thử nghiệm lớn chỉ cách đây 6 năm. Thành tựu của AlphaGo Zero trong trò chơi cờ vua và cờ vây liên quan đến năng lực tính toán ước tính vượt quá 10  siêu máy tính mạnh nhất thế giới gộp lại.

***Truy cập và sử dụng dữ liệu***

*Truy cập và chia sẻ dữ liệu có thể tăng tốc hoặc cản trở tiến bộ của TTNT*

Các công nghệ máy học hiện nay cần nguồn dữ liệu chính xác và có chọn lọc để đào tạo và phát triển. Do đó, quyền truy cập vào bộ dữ liệu chất lượng cao rất quan trọng đối với sự phát triển của TTNT. Các yếu tố liên quan đến việc truy cập và chia sẻ dữ liệu có thể tăng tốc hoặc cản trở sự tiến bộ của TTNT bao gồm:

- *Tiêu chuẩn*: Cần có các tiêu chuẩn cho phép tương tác và tái sử dụng dữ liệu trên các ứng dụng để thúc đẩy truy cập và đảm bảo dữ liệu có thể được tìm thấy, lập danh mục và/hoặc tìm kiếm và sử dụng lại.

- *Rủi ro*: Rủi ro đối với cá nhân, tổ chức và quốc gia chia sẻ dữ liệu bao gồm vi phạm bảo mật và quyền riêng tư, rủi ro đối với quyền SHTT và lợi ích thương mại, rủi ro an ninh quốc gia tiềm ẩn và rủi ro an ninh kỹ thuật số.

- *Chi phí dữ liệu*: Việc thu thập, truy cập, chia sẻ và sử dụng lại dữ liệu đòi hỏi phải đầu tư trước và theo dõi. Ngoài thu thập dữ liệu, cần đầu tư thêm để làm sạch dữ liệu, quản lý dữ liệu, bảo trì siêu dữ liệu, lưu trữ và xử lý dữ liệu cũng như hạ tầng CNTT an toàn.

- *Khuyến khích*: Phương pháp tiếp cận dựa vào thị trường có thể đưa ra các khuyến khích cấp quyền truy cập và chia sẻ dữ liệu với thị trường dữ liệu và các nền tảng thương mại hóa dữ liệu và cung cấp các dịch vụ giá trị gia tăng như hạ tầng thanh toán và trao đổi dữ liệu.

- *Sự không chắc chắn về quyền sở hữu dữ liệu*: Các khuôn khổ pháp lý gồm có quyền SHTT, luật hình sự (mạng), luật cạnh tranh và luật bảo vệ quyền riêng tư, kết hợp với nhiều bên tham gia xây dựng dữ liệu đã dẫn đến sự không chắc chắn xoay quanh vấn đề về "quyền sở hữu dữ liệu".

- *Trao quyền cho người dùng, bao gồm các tác nhân được TTNT hỗ trợ*: Trao quyền cho người sử dụng dữ liệu và thúc đẩy khả năng chuyển dữ liệu cũng như có được sự đồng thuận và lựa chọn hiệu quả của các chủ thể dữ liệu, có thể khuyến khích các cá nhân và doanh nghiệp chia sẻ dữ liệu cá nhân hoặc doanh nghiệp. Một số nhấn mạnh cách các tác nhân được sự hỗ trợ của TTNT biết sở thích của các cá nhân, có thể giúp họ đàm phán chia sẻ dữ liệu phức tạp với các hệ thống TTNT khác.

- *Bên thứ ba đáng tin cậy*: Bên thứ ba có thể tạo niềm tin và thúc đẩy việc chia sẻ và sử dụng lại dữ liệu giữa tất cả các bên liên quan. Bên trung gian dữ liệu hoạt động như đơn vị cấp chứng nhận. Các nền tảng chia sẻ dữ liệu đáng tin cậy cung cấp dữ liệu chất lượng cao. Hội đồng xét duyệt thể chế đảm bảo tôn trọng lợi ích hợp pháp của bên thứ ba.

- *Tính đại diện của dữ liệu*: Hệ thống TTNT đưa ra dự báo dựa vào các mẫu

được xác định trong bộ dữ liệu đào tạo. Trong trường hợp này, cả về độ chính xác và công bằng, các bộ dữ liệu đào tạo phải mang tính toàn diện, đa dạng và đại diện để không đề cập thiếu hoặc xuyên tạc các nhóm cụ thể.

*Chính sách tăng cường truy cập và chia sẻ dữ liệu để phát triển TTNT*

Các cách tiếp cận chính sách để tăng cường truy cập và chia sẻ dữ liệu bao gồm:

- *Cung cấp quyền truy cập dữ liệu khu vực công,* bao gồm dữ liệu khu vực công, dữ liệu mở của chính phủ, dữ liệu địa lý (như bản đồ) và dữ liệu giao thông.

- *Thúc đẩy chia sẻ dữ liệu trong khu vực tư nhân*, thường là trên cơ sở tự nguyện hoặc đối với các chính sách bắt buộc, việc chia sẻ dữ liệu bị hạn chế với những người dùng đáng tin cậy. Các lĩnh vực trọng tâm đặc biệt bao gồm "dữ liệu được công chúng quan tâm", dữ liệu trong các ngành công nghiệp mạng lưới như giao thông và năng lượng với khả năng tương tác dịch vụ và chuyển dữ liệu cá nhân.

- *Phát triển năng lực thống kê/phân tích dữ liệu* thông qua thành lập các trung tâm công nghệ hỗ trợ và hướng dẫn sử dụng và phân tích dữ liệu.

- *Phát triển chiến lược dữ liệu quốc gia* để đảm bảo tính nhất quán của các khuôn khổ quản trị dữ liệu quốc gia và khả năng tương thích của chúng với các chiến lược TTNT quốc gia.

*Các phương pháp kỹ thuật xử lý hạn chế về dữ liệu*

Một số thuật toán máy học như thuật toán nhận dạng hình ảnh vượt quá khả năng bình thường của con người. Tuy nhiên, để có khả năng này, các thuật toán phải được đào tạo cơ sở dữ liệu lớn gồm hàng triệu hình ảnh được dán nhãn. Nhu cầu dữ liệu đã khuyến khích nghiên cứu các kỹ thuật máy học cần ít dữ liệu để đào tạo hệ thống TTNT. Dưới đây là một số phương pháp để giải quyết tình trạng thiếu dữ liệu này.

- *Học tăng cường sâu* là kỹ thuật máy học kết hợp mạng nơron sâu với học tăng cường. Theo đó, học tăng cường sâu ủng hộ hành vi cụ thể dẫn đến kết quả mong đợi. Các "tác nhân" thông minh nhân tạo cạnh tranh thông qua các hành động trong môi trường phức tạp và được "thưởng" hoặc "phạt" tùy theo hành động đó có dẫn đến kết quả mong đợi hay không. Các tác nhân điều chỉnh hành động của chúng theo "phản hồi" này.

- *Chuyển giao việc học hoặc đào tạo trước* sử dụng lại các mô hình đã được đào tạo để thực hiện nhiều nhiệm vụ khác nhau. Ví dụ, một số lớp của mô hình được đào tạo để nhận dạng hình ảnh con mèo, có thể được sử dụng lại để phát hiện hình ảnh của chiếc váy màu xanh. Trong những trường hợp này, mẫu hình ảnh sẽ có độ lớn nhỏ hơn yêu cầu của các thuật toán máy học truyền thống.

- *Học dữ liệu tăng cường* hoặc "tổng hợp" dữ liệu có thể tạo ra dữ liệu một cách giả tạo thông qua mô phỏng hoặc nội suy dựa vào dữ liệu hiện có. Điều đó làm tăng hiệu quả của dữ liệu và cải thiện việc học. Phương pháp này đặc biệt có lợi trong những trường hợp  các ràng buộc về quyền riêng tư hạn chế việc sử dụng dữ liệu hoặc để mô phỏng các tình huống hiếm gặp trong thực tế.

- *Các mô hình học tập kết hợp* có thể mô hình hóa sự không chắc chắn bằng

cách kết hợp các loại mạng nơron sâu với các phương pháp tiếp cận theo xác suất hoặc Bayes. Qua đó có thể tăng hiệu suất và khả năng giải thích, đồng thời các dự đoán sai.

***Cạnh tranh***

Một số tác động đến cạnh tranh đặc biệt bắt nguồn từ TTNT như thúc đẩy việc tạo ra đầu vào mới. Bên cạnh đó, chính sách cạnh tranh giữa các chủ thể TTNT lớn được chú ý là vì các chủ thể đóng vai trò là nền tảng trực tuyến và người nắm giữ khối lượng lớn dữ liệu.

Vấn đề cụ thể liên quan đến TTNT là hiệu ứng mạng được định hướng bởi dữ liệu. Theo đó, tiện ích của mỗi người dùng từ việc sử dụng một số loại nền tảng nhất định sẽ tăng lên mỗi khi những người khác sử dụng nó. Ví dụ, thông qua một trong những nền tảng này, người dùng dạy các thuật toán cách phục vụ người dùng tốt hơn. Tuy nhiên, có người lại cho rằng trao đổi dữ liệu sẽ giảm sút về quy mô khi dữ liệu tăng vượt quá ngưỡng nhất định. Do đó, một số người băn khoăn liệu TTNT có gây lo ngại về cạnh tranh lâu dài hay không.

Lợi thế của kinh tế quy mô thể hiện ở giá trị kinh doanh của dữ liệu bổ sung. Nếu khả năng dẫn đầu về chất lượng dữ liệu giữa các đối thủ cạnh tranh cho phép một công ty thu hút nhiều khách hàng hơn, thì có thể tạo ra vòng phản hồi tích cực. Nhiều khách hàng nghĩa là sẽ có nhiều dữ liệu hơn, củng cố chu kỳ và cho phép gia tăng sự thống trị thị trường theo thời gian. Ngoài ra, lợi thế của kinh tế quy mô ở khía cạnh chuyên môn cũng cần có để xây dựng hệ thống TTNT hiệu quả.

Bên cạnh đó vẫn có lo ngại các thuật toán tạo sự thông đồng thông qua theo dõi các điều kiện thị trường, giá cả và phản ứng của đối thủ cạnh tranh với những thay đổi về giá. Qua đó, các công ty được cung cấp những công cụ cải tiến mới để điều phối các chiến lược, ấn định giá và thực thi thỏa thuận Cartel [1]. Tuy nhiên, theo suy đoán, các thuật toán học sâu tinh vi hơn thậm chí sẽ không yêu cầu các thỏa thuận thực tế giữa các đối thủ cạnh tranh để đạt được kết quả giống như thỏa thuận Cartel. Thay vào đó, kết quả đạt được mà không cần sự can thiệp của con người. Điều đó sẽ đặt ra những thách thức thực thi khó khăn. Luật cạnh tranh yêu cầu bằng chứng về các thỏa thuận hoặc "cuộc gặp gỡ giữa các bên" trước khi hành vi vi phạm thỏa thuận Cartel có thể được xác lập và trừng phạt.

***Sở hữu trí tuệ***

Các quy tắc SHTT thường thúc đẩy mức độ và tốc độ khám phá, phát minh và phổ biến công nghệ TTNT mới. Theo đó, các quy tắc SHTT tương tự như các quy tắc cho các công nghệ khác được bảo vệ bởi quyền SHTT. Dù quy tắc SHTT là "phần thưởng" dành cho cho các nhà phát minh, tác giả, nghệ sĩ và chủ sở hữu thương hiệu, nhưng chính sách SHTT cũng nên coi tiềm năng của TTNT là đầu vào ĐMST mạnh mẽ hơn.

Bảo vệ TTNT bằng quyền SHTT mà không phải bí mật thương mại có thể làm nảy sinh những vấn đề mới về cách khuyến khích các chuyên gia đổi mới tiết lộ những đổi mới sáng tạo TTNT, bao gồm các thuật toán và việc đào tạo chúng. Hội

*[1] Cartel là thỏa thuận giữa các công ty cạnh tranh để kiểm soát giá hoặc loại trừ các sản phẩm của một đối thủ cạnh tranh mới trong thị trường*

nghị của Văn phòng Nghị viện châu Âu đã thảo luận về ba loại hình cấp sáng chế TTNT. Loại đầu tiên là Core AI, thường liên quan đến các thuật toán vì các phương pháp toán học không được cấp sáng chế. Loại thứ hai là mô hình được đào tạo/máy học - xác nhận các biến thể và phạm vi. Cuối cùng, TTNT được cấp sáng chế như một công cụ trong lĩnh vực ứng dụng, được xác định thông qua các hiệu ứng kỹ thuật. Các tổ chức quốc tế khác và các nước OECD cũng đang tìm hiểu tác động của TTNT trong không gian SHTT.

Tính phổ biến của TTNT dẫn đến cân nhắc có nên điều chỉnh hệ thống SHTT trong thế giới mà hệ thống TTNT có thể tự tạo ra các phát minh. Một số hệ thống TTNT đã cho ra đời những phát minh có thể được cấp sáng chế, đặc biệt là trong lĩnh vực hóa học, dược phẩm và công nghệ sinh học. Trong những lĩnh vực này, nhiều phát minh liên quan đến việc kết hợp theo nhiều cách ngay từ ban đầu giữa các phân tử để tạo nên các hợp chất mới hoặc xác định đặc điểm mới của các phân tử hiện có. Ví dụ, KnIT, công cụ máy học do IBM phát triển, đã xác định thành công kinase - các enzym hoạt động như chất xúc tác để chuyển các nhóm photphat đến các chất nền cụ thể. Các kinase này trong số một tập hợp kinase quen thuộc, đã được kiểm tra bằng thực nghiệm. Phần mềm đã phát hiện ra những đặc điểm cụ thể của các phân tử đó. Các phát minh này đã được xin cấp sáng chế. Những vấn đề liên quan đến TTNT và SHTT, đang được các cơ quan chuyên môn của các nước OECD như Văn phòng Sáng chế châu Âu và Văn phòng Nhãn hiệu và Sáng chế Hoa Kỳ, cũng như Tổ chức SHTT thế giới xem xét. Vấn đề bảo vệ bản quyền dữ liệu do TTNT xử lý cũng được chú ý.

***Doanh nghiệp nhỏ và vừa***

Các chính sách và chương trình giúp doanh nghiệp nhỏ và vừa (DNNVV) định hướng quá chuyển đổi TTNT, là những ưu tiên lớn. Các công cụ tiềm năng hỗ trợ hệ sinh thái số cho DNNVV áp dụng TTNT, bao gồm:

• Nâng cao kỹ năng được xem là rất quan trọng vì cạnh tranh để tìm kiếm tài năng TTNT khan hiếm là mối quan tâm đặc biệt của DNNVV.

• Khuyến khích đầu tư có mục tiêu vào các ngành dọc đã được chọn lọc. Ví dụ, chính sách khuyến khích đầu tư ứng dụng TTNT trong ngành nông nghiệp của Pháp mang lại lợi ích cho tất cả các bên liên quan mà DNNVV không đủ khả năng đầu tư một mình.

• Giúp DNNVV truy cập dữ liệu thông qua xây dựng các nền tảng trao đổi dữ liệu.

• Hỗ trợ DNNVV cải thiện khả năng tiếp cận công nghệ TTNT thông qua chuyển giao công nghệ từ các viện nghiên cứu công, cũng như khả năng tiếp cận với năng lực tính toán và nền tảng đám mây.

• Cải thiện các cơ chế tài chính giúp DNNVV về TTNT mở rộng quy mô như thông qua quỹ đầu tư công mới và tăng tính linh hoạt và giới hạn tài chính của các kế hoạch đầu tư cho các công ty thâm dụng tri thức. Ủy ban châu Âu cũng đang tập trung hỗ trợ DNNVV ở châu Âu thông qua dự án AI4EU, nền tảng TTNT theo yêu cầu.

### *2.3.3. Môi trường chính sách cho đổi mới TTNT*

Chính phủ nhiều nước đã xây dựng môi trường có kiểm soát để thử nghiệm hệ thống TTNT, bao gồm khung thể chế thử nghiệm, trung tâm đổi mới sáng tạo và phòng thí nghiệm chính sách. Thử nghiệm chính sách diễn ra ở "*chế độ khởi động*". Trong trường hợp này, các thử nghiệm được triển khai, đánh giá và sửa đổi, sau đó mở rộng hoặc thu hẹp quy mô hoặc bị loại bỏ nhanh chóng.

Lựa chọn khác thúc đẩy đưa ra các quyết định nhanh và hiệu quả hơn là sử dụng các công cụ số để thiết kế chính sách, bao gồm chính sách ĐMST và theo dõi các mục tiêu chính sách. Ví dụ, chính phủ một số nước sử dụng "*mô hình dựa vào tác nhân*" để dự đoán tác động của các biến thể chính sách đến các loại hình doanh nghiệp khác nhau.

Nhiều nước khuyến khích các chủ thể TTNT phát triển cơ chế tự điều chỉnh như quy tắc ứng xử, tiêu chuẩn tự nguyện và thực hành tốt. Các cơ chế này định hướng các chủ thể TTNT thông qua vòng đời của TTNT, bao gồm theo dõi, báo cáo, đánh giá và giải quyết tác động có hại hoặc tình trạng sử dụng sai hệ thống TTNT.

Chính phủ các nước cũng có thể xây dựng và khuyến khích các cơ chế giám sát hệ thống TTNT của khu vực công và tư, bao gồm thẩm tra, đánh giá việc tuân thủ và tính phù hợp cùng với chương trình cấp chứng nhận. Các cơ chế này có thể được sử dụng khi xem xét các nhu cầu cụ thể của DNNVV và những khó khăn mà doanh nghiệp vấp phải.

### *2.2.4. Chuẩn bị chuyển đổi việc làm và xây dựng kỹ năng*

***Việc làm***

TTNT là lĩnh vực đang phát triển nhanh với việc bắt đầu công việc phân tích dựa vào bằng chứng, sẽ làm thay đổi bản chất công việc khi TTNT lan tỏa khắp các lĩnh vực. TTNT sẽ hỗ trợ con người trong một số nhiệm vụ, thay thế con người trong những nhiệm vụ khác và cũng tạo ra các loại việc làm mới. Nội dung dưới đây đề cập đến một số thay đổi dự kiến của thị trường lao động do tác động của TTNT, cũng như các cân nhắc chính sách đi kèm quá trình chuyển đổi sang nền kinh tế TTNT.

*TTNT sẽ cải thiện năng suất*

TTNT được kỳ vọng sẽ cải thiện năng suất theo hai cách. Đầu tiên, một số hoạt động trước đây do con người đảm nhiệm, sẽ được tự động hóa. Thứ hai, thông qua sự tự chủ của máy móc, các hệ thống sẽ hoạt động và thích ứng với môi trường không có hoặc có ít sự kiểm soát của con người. Nghiên cứu 12 nền kinh tế phát triển đưa ra ước tính, TTNT có thể tăng 40% năng suất lao động vào năm 2035 so với mức cơ bản dự kiến. Ví dụ, Watson (phần mềm trí tuệ nhân tạo) của IBM hỗ trợ chuyên viên tư vấn tại ngân hàng Crédit Mutuel của Pháp trả lời các câu hỏi của khách hàng với tốc độ nhanh hơn 60%. Phần mềm Chatbot của Alibaba đã xử lý hơn 95% các câu hỏi của khách hàng trong dịp bán hàng hạ giá năm 2017. Điều này cho phép các đại diện dịch vụ khách hàng xử lý nhiều vấn đề phức tạp hoặc mang tính cá nhân. Về lý thuyết, tăng năng suất lao động sẽ dẫn đến tăng tiền lương vì mỗi người lao động tạo ra giá trị gia tăng cao hơn.

Các nhóm có sự kết hợp giữa con người và TTNT giúp giảm sai sót và mở rộng cơ hội cho người lao động. Nhóm con người - TTNT đã được chứng minh là cho năng suất cao hơn nhóm chỉ có TTNT hoặc người lao động. Ví dụ, nhóm con người - TTNT trong nhà máy BMW đã tăng 85% năng suất so với các nhóm riêng lẻ. Ngoài sản xuất, rôbốt của Walmart còn quét hàng tồn kho để các cộng sự của cửa hàng tập trung trợ giúp khách hàng. Bác sỹ X quang khi kết hợp với mô hình TTNT để sàng lọc chụp X quang ngực cho bệnh nhân lao, đã đạt độ chính xác lên tới 100%, cao hơn khi chỉ TTNT hoặc bác sỹ tiến hành chụp.

TTNT cũng giúp các nhiệm vụ tự động trước đây diễn ra suôn sẻ và nhanh hơn. Kết quả là các công ty có thể sản xuất nhiều hơn với chi phí thấp hơn. Từ đó, nhu cầu hàng hóa sẽ tăng lên. Như vậy, nhu cầu về nguồn lao động (như trong vai trò sản xuất) cũng như các lĩnh vực cho hàng hóa trung gian tăng lên.

*TTNT sẽ thay đổi và gia tăng các nhiệm vụ có thể được tự động hóa*

Tự động hóa không phải mới, nhưng TTNT sẽ thay đổi và làm tăng số lượng các nhiệm vụ có thể được tự động hóa. Không như máy tính, công nghệ TTNT không được lập trình sẵn và dựa trên các quy tắc. Máy tính có xu hướng làm giảm số lượng việc làm thông thường đòi hỏi kỹ năng ở mức trung bình. Tuy nhiên, các ứng dụng mới có sự hỗ trợ của TTNT có thể thực hiện những nhiệm vụ tương đối phức tạp liên quan đến việc đưa ra dự báo. Những công việc này bao gồm phiên âm, dịch thuật, lái xe, chẩn đoán bệnh tật và giải đáp thắc mắc của khách hàng.

*Tác động của TTNT đến việc làm sẽ phụ thuộc vào tốc độ lan tỏa của nó trên các lĩnh vực khác nhau*

Trong những thập kỷ tới, tác động của TTNT đến việc làm sẽ phụ thuộc vào tốc độ phát triển và phổ biến của công nghệ TTNT trong các lĩnh vực khác nhau. Xe tự động được cho là sẽ hủy hoại công việc lái xe và dịch vụ giao hàng. Ví dụ, các công ty xe tải được thành lập như Volvo và Daimler đang cạnh tranh với các công ty khởi nghiệp như Kodiak và Einride trong việc phát triển và thử nghiệm xe tải không người lái. Theo Diễn đàn Giao thông quốc tế, xe tải không người lái sẽ xuất hiện thường xuyên trên nhiều tuyến đường trong 10 năm tới. Khoảng 50 - 70% trong số 6,4 triệu việc làm vận tải đường bộ chuyên nghiệp ở Hoa Kỳ và châu Âu có thể bị cắt giảm vào năm 2030. Tuy nhiên, việc làm mới sẽ được tạo ra song song để cung cấp các dịch vụ hỗ trợ cho số lượng xe tải không người lái ngày càng tăng. Xe tải không người lái có thể giảm 30% chi phí vận hành cho vận tải hàng hóa đường bộ, đặc biệt là do tiết kiệm chi phí lao động. Điều này có thể khiến các công ty vận tải đường bộ truyền thống ngừng hoạt động kinh doanh, dẫn đến sự sụt giảm việc làm thậm chí còn nhanh hơn.

Công nghệ TTNT tác động đến việc làm truyền thống đòi hỏi kỹ năng cao

Công nghệ TTNT cũng đang thực hiện các nhiệm vụ dự báo trước đây do lao động tay nghề cao thực hiện từ luật sư đến nhân viên y tế. Một "luật sư" rôbốt đã khiếu nại thành công các vé phạt trị giá hơn 12 triệu USD. Năm 2016, nền tảng trí tuệ nhân tạo Watson và DeepMind Health của IBM đã vượt trội hơn các bác sĩ trong việc chẩn đoán các bệnh ung thư hiếm gặp. TTNT đã được chứng minh có khả năng dự báo biến động của thị trường chứng khoán tốt hơn các chuyên gia tài chính.

*TTNT hỗ trợ con người và tạo việc làm mới*

TTNT hỗ trợ con người và tạo cơ hội việc làm cho người lao động. Các lĩnh vực đáng chú ý bao gồm những lĩnh vực hỗ trợ dự đoán và khai thác các kỹ năng của con người như tư duy phản biện, tính sáng tạo và sự đồng cảm.

• Các nhà khoa học dữ liệu và chuyên gia máy học*:* Cần có các chuyên gia để xây dựng và làm sạch dữ liệu cũng như lập trình và phát triển ứng dụng TTNT. Tuy nhiên, dù dữ liệu và máy học mang đến một số nhiệm vụ mới, nhưng không có khả năng tạo ra khối lượng lớn việc làm mới cho người lao động.

• Hành động*:* Một số hành động khi được thực hiện bởi con người, sẽ có giá trị hơn là máy móc như hành động của các vận động viên chuyên nghiệp, người chăm sóc trẻ em hoặc nhân viên kinh doanh. Nhiều người nghĩ rằng con người sẽ tập trung nhiều hơn vào công việc cải thiện cuộc sống của người khác như chăm sóc trẻ em, huấn luyện thể chất và chăm sóc người bệnh nan y.

• Phán đoán để xác định những gì cần dự báo*:* Khi TTNT được sử dụng để dự đoán, con người phải quyết định những gì cần dự báo và làm gì với các dự báo. Đặt ra các tình huống khó xử, diễn giải tình huống hoặc trích xuất ý nghĩa từ văn bản thì cần những người có các tố chất phán đoán và công bằng. Ví dụ, trong khoa học, TTNT có thể hỗ trợ những người có ý tưởng để xây dựng khung nghiên cứu và thiết lập bối cảnh cho các thí nghiệm cụ thể.

• *Phán đoán để quyết định cần làm gì với một dự đoán:* Một quyết định không thể được đưa ra chỉ với một dự đoán. Ví dụ, một quyết định bình thường như có nên mang ô khi ra ngoài đi dạo hay không, sẽ xem xét dự báo về khả năng mưa. Tuy nhiên, quyết định sẽ phụ thuộc phần lớn vào sở thích của người đó. Ví dụ này có thể được mở rộng cho nhiều quyết định quan trọng.

*Dự báo tác động thực của TTNT đến sự sụt giảm số lượng việc làm khác nhau*

Trong 5 năm qua, các ước tính về tác động tổng thể của tự động hóa đến mất việc làm đã được thực hiện. Ví dụ, 47% việc làm ở Hoa Kỳ được dự báo có nguy cơ thay đổi trong 10-15 năm tới. Viện Nghiên cứu toàn cầu McKinsey đã sử dụng phương pháp tiếp cận theo định hướng nhiệm vụ và phát hiện ra rằng năm 2017, khoảng 1/3 hoạt động trong 60% số việc làm có thể được tự động hóa. Tuy nhiên, các công việc bị ảnh hưởng bởi tự động hóa không chỉ do sự phát triển và triển khai áp dụng riêng TTNT mà còn do các công nghệ khác.

Ngoài ra, việc dự báo sự xuất hiện của việc làm tương lai trong các lĩnh vực mới là một thách thức. Một nghiên cứu ước tính TTNT sẽ tạo ra 2 triệu việc làm vào năm 2025, có thể từ các nghề mới xuất hiện và thông qua các kênh gián tiếp. Ví dụ, TTNT có khả năng giảm chi phí sản xuất hàng hóa và dịch vụ nhưng lại gia tăng về chất lượng. Điều này sẽ dẫn đến nhu cầu tăng và việc làm được tạo ra nhiều hơn.

Các ước tính gần đây của OECD tính đến sự không đồng nhất của các nhiệm vụ trong những việc làm được xác định trong phạm vi hẹp thông qua sử dụng dữ liệu của Chương trình quốc tế đánh giá năng lực người trưởng thành (PIAAC). Dựa vào các công nghệ hiện có, 14% việc làm tại các nước thành viên có nguy cơ tự động hóa cao; 32% người lao động khác sẽ có sự thay đổi lớn trong cách thực hiện công việc. Phân tích gần đây của OECD cho thấy sự sụt giảm việc làm trong các ngành

nghề được coi là "có khả năng tự động hóa cao" tại 82% khu vực của 16 quốc gia châu Âu. Đồng thời, phân tích xác định sự gia tăng mạnh hơn các công việc "tự động hóa thấp" tại 60% khu vực gây mất việc làm. Nghiên cứu này ủng hộ ý tưởng tự động hóa có thể thay đổi sự kết hợp giữa các việc làm, mà không làm giảm tổng số việc làm.

*TTNT sẽ thay đổi bản chất công việc*

TTNT được áp dụng rộng rãi, sẽ làm thay đổi bản chất của công việc. Thông qua tự động hóa các công việc thường ngày, TTNT làm cho công việc trở nên thú vị hơn và tạo sự cân bằng tốt hơn giữa công việc và cuộc sống. Tính sáng tạo và sự khéo léo của con người giúp khai thác hiệu quả các nguồn lực tính toán, dữ liệu và thuật toán để tạo ra các nhiệm vụ và định hướng mới cần có sự sáng tạo của con người.

TTNT thúc đẩy thay đổi cách thị trường vận hành bằng cách tăng hiệu quả. Ngày nay, các kỹ thuật TTNT cùng với dữ liệu lớn có tiềm năng giúp các công ty xác định vai trò của người lao động và kết nối người lao động với công việc. Ví dụ, IBM sử dụng TTNT để tối ưu hóa đào tạo người lao động, đề xuất các mô-đun đào tạo người lao động dựa vào hiệu quả lao động trước đây, mục tiêu nghề nghiệp và nhu cầu kỹ năng của IBM. Các công ty như KeenCorp và Vibe đã phát triển các kỹ thuật phân tích văn bản, giúp họ phân tích thông tin về người lao động nhằm đánh giá các chỉ số như tinh thần và năng suất lao động. Qua đó, TTNT giúp các công ty tối ưu hóa năng suất lao động.

*Các thông số về sự thay đổi của tổ chức cần được xây dựng*

Yêu cầu đối với các tiêu chuẩn công nghiệp mới hoặc sửa đổi và các thỏa thuận công nghệ giữa cấp quản lý và người lao động hướng tới những nơi làm việc đáng tin cậy, an toàn và hiệu quả ngày càng gia tăng. Ủy ban Kinh tế và Xã hội châu Âu (EESC) khuyến nghị "*các bên liên quan hợp tác với nhau trên nền tảng của hệ thống TTNT hỗ trợ và đồng sáng tạo tại nơi làm việc*". Nơi làm việc cũng cần sự linh hoạt, bảo vệ quyền tự chủ của người lao động và chất lượng công việc, bao gồm cả việc chia sẻ lợi nhuận. Gần đây giữa công đoàn công ty IG Metall của Đức và người sử dụng lao động đã có thỏa thuận tập thể về việc thay đổi thời gian làm việc. Thỏa thuận cho thấy tùy thuộc vào nhu cầu của tổ chức và cá nhân trong thế giới việc làm mới, người sử dụng lao động và công đoàn có thể đạt được thỏa thuận mà không cần thay đổi các biện pháp bảo vệ việc làm hợp pháp.

*Sử dụng TTNT để hỗ trợ các chức năng của thị trường lao động với các biện pháp bảo vệ*

TTNT hỗ trợ đào tạo và kết nối việc làm hiệu quả hơn. TTNT là cầu nối giữa những người tìm việc với các chương trình phát triển lực lượng lao động. Ở nhiều nước OECD, người sử dụng lao động và các dịch vụ việc làm công đã sử dụng các nền tảng trực tuyến để làm việc. Trong tương lai, TTNT và các công nghệ số khác có thể cải thiện các phương pháp tiếp cận sáng tạo và cá nhân hóa đối với quá trình tìm kiếm việc làm và tuyển dụng, đồng thời tăng hiệu quả kết nối cung cầu lao động. Nền tảng LinkedIn sử dụng TTNT giúp các nhà tuyển dụng tìm thấy những ứng viên phù hợp và kết nối các ứng viên với những công việc phù hợp. Nền tảng này dựa vào dữ liệu về hồ sơ và hoạt động của 470 triệu người dùng đã đăng ký.

Các công nghệ TTNT khai thác dữ liệu lớn, cũng có thể giúp thông tin cho chính phủ, người sử dụng lao động và người lao động về các điều kiện thị trường lao động địa phương. Thông tin này hỗ trợ xác định và dự báo nhu cầu về kỹ năng, các nguồn đào tạo trực tiếp và kết nối các cá nhân với công việc. Các dự án phát triển thông tin thị trường lao động đã và đang được tiến hành tại các nước như Phần Lan, Cộng hòa Séc và Latvia.

*Quản lý sử dụng dữ liệu lao động*

Dù TTNT cần có các bộ dữ liệu lớn mới hoạt động hiệu quả, nhưng có một số rủi ro tiềm ẩn khi đây là dữ liệu về người lao động, đặc biệt nếu hệ thống TTNT phân tích dữ liệu không rõ ràng. Việc lập kế hoạch về nhân sự và năng suất sẽ đòi hỏi phải khai thác ngày càng nhiều dữ liệu về người lao động và các thuật toán. Khi làm như vậy, các nhà hoạch định chính sách công và các bên liên quan phải nghiên cứu tác động của việc thu thập và xử lý dữ liệu đến triển vọng và điều khoản việc làm. Dữ liệu thu thập từ các ứng dụng, dấu vân tay, thiết bị đeo trên người và cảm biến trong thời gian thực, cho biết vị trí và nơi làm việc của người lao động.

Tuy nhiên, tại một số quốc gia đã có các thỏa thuận về dữ liệu người lao động và quyền ngắt kết nối. Công ty viễn thông Orange France Telecom và 5 trung tâm nghiệp đoàn là những đơn vị đầu tiên cam kết bảo vệ dữ liệu cho người lao động. Các biện pháp bảo vệ cụ thể bao gồm tính minh bạch trong việc sử dụng, đào tạo và đưa vào sử dụng thiết bị mới. Ngoài ra, trong các quy định đối với dữ liệu về người lao động có các điều khoản liên quan đến việc thành lập bộ phận quản lý dữ liệu trong các công ty, trách nhiệm giải trình, khả năng chuyển dữ liệu, quyền giải thích và xóa dữ liệu.

***Quản lý quá trình chuyển đổi TTNT***

Khả năng xảy ra gián đoạn và hỗn loạn trong thị trường lao động khi công nghệ vượt xa khả năng thích ứng của tổ chức. Sự lạc quan dài hạn không ngụ ý về sự chuyển đổi suôn sẻ sang một nền kinh tế với sự hỗ trợ ngày càng mạnh mẽ của TTNT: một số lĩnh vực có khả năng tăng trưởng, trong khi những lĩnh vực khác suy giảm. Việc làm hiện tại có thể biến mất, trong khi việc làm mới xuất hiện. Do đó, các vấn đề chính sách quan trọng về TTNT và việc làm đều liên quan đến việc quản lý quá trình chuyển đổi. Các chính sách quản lý chuyển đổi bao gồm mạng lưới an sinh xã hội, bảo hiểm y tế, đánh thuế lũy tiến đối với lao động và vốn cũng như giáo dục. Hơn nữa, các chính sách cạnh tranh và các chính sách khác có thể ảnh hưởng đến sự tập trung, sức mạnh thị trường và phân phối thu nhập cũng cần được quan tâm.

***Kỹ năng sử dụng TTNT***

*Khi việc làm thay đổi, các kỹ năng cần thiết của người lao động cũng phải thay đổi*

Khi việc làm thay đổi, các kỹ năng cần thiết của người lao động cũng thay đổi theo. Một số tác động có thể của TTNT đến các kỹ năng nhấn mạnh TTNT là lĩnh vực đang phát triển nhanh với việc khởi động công việc phân tích dựa vào bằng chứng. Chính sách giáo dục sẽ cần được điều chỉnh để mở rộng học tập suốt đời, đào tạo và phát triển kỹ năng. Cũng như các lĩnh vực công nghệ khác, TTNT sẽ tạo ra

nhu cầu trong ba lĩnh vực kỹ năng. Đầu tiên, cần có các kỹ năng chuyên môn để lập trình và phát triển ứng dụng TTNT. Đây có thể là các kỹ năng nghiên cứu cơ bản, kỹ thuật và ứng dụng TTNT, cũng như khoa học dữ liệu và tư duy tính toán. Thứ hai, cần có các kỹ năng chung để khai thác TTNT như thông qua các nhóm TTNT - con người để kiểm soát chất lượng. Thứ ba, TTNT sẽ cần có các kỹ năng bổ sung, bao gồm sử dụng các kỹ năng của con người như tư duy phản biện; tính sáng tạo, đổi mới và khởi nghiệp; và sự đồng cảm.

*Cần có các sáng kiến xây dựng và phát triển kỹ năng TTNT để giải quyết tình trạng thiếu kỹ năng*

Thiếu kỹ năng TTNT sẽ ngày càng phổ biến và trở nên rõ nét khi nhu cầu chuyên gia trong các lĩnh vực như máy học tăng nhanh. Các DNVVN, các trường đại học công lập và các trung tâm nghiên cứu đã cạnh tranh với các công ty thống lĩnh để tìm kiếm nhân tài. Các sáng kiến xây dựng và phát triển kỹ năng TTNT bắt đầu xuất hiện trong các khu vực công, tư và học thuật. Ví dụ, chính phủ Singapo đã xây dựng chương trình nghiên cứu 5 năm về quản trị TTNT và sử dụng dữ liệu tại Đại học Quản lý Singapo. Trung tâm Quản trị dữ liệu & Trí tuệ nhân tạo thuộc Đại học Quản lý Singapo tập trung nghiên cứu TTNT và ngành công nghiệp, xã hội và thương mại hóa. Về mặt học thuật, Đại học Công nghệ Massachusetts (MIT) đã cam kết dành 1 tỷ USD để thành lập Trường Cao đẳng Máy tính Schwarzman nhằm trang bị cho sinh viên và nhà nghiên cứu trong tất cả các ngành sử dụng kỹ thuật tính toán và TTNT đẩy mạnh phát triển lĩnh vực của họ và ngược lại.

Tình trạng thiếu kỹ năng TTNT cũng đã khiến một số quốc gia phải hợp lý hóa quy trình nhập cư đối với các chuyên gia tay nghề cao. Ví dụ, Vương quốc Anh đã tăng gấp đôi số thị thực Cấp 1 (Tài năng đặc biệt) lên 2.000 thị thực mỗi năm và sắp xếp hợp lý quy trình cho các sinh viên và nhà nghiên cứu hàng đầu làm việc tại đó. Tương tự, Canada đã xác định thời gian xử lý là hai tuần đối với các đơn xin thị thực từ lao động tay nghề cao và miễn thị thực cho các nhiệm vụ nghiên cứu ngắn hạn. Đây là một phần của Chiến lược kỹ năng toàn cầu năm 2017 nhằm thu hút người lao động và nhà nghiên cứu có kỹ năng cao từ nước ngoài.

*Kỹ năng chung để khai thác TTNT*

Tất cả các nước OECD đều đánh giá kỹ năng và dự báo nhu cầu kỹ năng trước mắt, trung hạn hoặc dài hạn. Phần Lan đã đề xuất Chương trình TTNT, bao gồm quà tặng là một tài khoản kỹ năng hoặc chương trình học tập suốt đời để kích thích nhu cầu giáo dục và đào tạo. Vương quốc Anh đang thúc đẩy phát triển lực lượng lao động đa dạng và đầu tư khoảng 530 triệu USD đào tạo kỹ năng TTNT. Nội dung trọng tâm là khoa học, công nghệ, kỹ thuật và toán học và giáo viên khoa học máy tính.

Các học viên hiện phải trở thành người có kỹ năng kép. Đây là những người có thể chuyên về một lĩnh vực như kinh tế, sinh học hoặc luật, nhưng cũng có kỹ năng TTNT như máy học. Theo xu hướng này, tháng 10/2018, MIT đã công bố sự thay đổi lớn về cơ cấu trong 50 năm với kế hoạch xây dựng một trường đào tạo về máy tính không thuộc ngành kỹ thuật và có sự đan xen tất cả các ngành học khác. Trường sẽ đào tạo những sinh viên có kỹ năng kép áp dụng TTNT và máy học vào giải quyết những thách thức trong lĩnh vực của họ. MIT dành 1 tỷ USD để thành lập ngôi

trường mới này.

*Kỹ năng bổ sung*

Dựa vào nghiên cứu hiện có, các kỹ năng này bao gồm khả năng phán đoán, phân tích của con người và giao tiếp giữa các cá nhân. Năm 2021, OECD sẽ đưa một mô-đun vào Chương trình Đánh giá học sinh quốc tế (PISA) để kiểm tra các kỹ năng tư duy sáng tạo và phản biện. Kết quả sẽ cung cấp đánh giá tiêu chuẩn về khả năng sáng tạo giữa các quốc gia để thông tin về chính sách và các hành động của đối tác xã hội.

## III. CHÍNH SÁCH VÀ SÁNG KIẾN TTNT

### 3.1. Nguyên tắc TTNT trong xã hội

Một số bên liên quan vẫn tích cực thảo luận cách đẩy mạnh triển khai ứng dụng TTNT phục vụ toàn xã hội. Ví dụ, tháng 4/2016, Viện Kỹ sư điện và điện tử (IEEE) đã khởi động Sáng kiến toàn cầu về đạo đức của hệ thống tự động và thông minh. Đến tháng 12/2017, Viện đã công bố phiên bản 2 các nguyên tắc thiết kế phù hợp về đạo đức. Sáng kiến Hợp tác TTNT mang lại lợi ích cho con người và xã hội, được công bố vào tháng 9/2016 với rất nhiều nguyên lý, đã bắt đầu được triển khai để xây dựng các nguyên tắc cho nhiều vấn đề cụ thể như sự an toàn. Các Nguyên tắc TTNT của Asilomar là một tập hợp các nghiên cứu, khía cạnh đạo đức và giá trị cho sự phát triển an toàn và có lợi cho xã hội của TTNT trước mắt và lâu dài. Sáng kiến TTNT thu hút các chuyên gia, học viên và công dân trên toàn cầu cùng cung cấp hiểu biết chung về các khái niệm như khả năng giải thích của TTNT.

***Bảng 1.*** Các bộ hướng dẫn TTNT chọn lọc do các bên liên quan đưa ra

| Tham khảo | Hướng dẫn TTNT do các bên liên quan đưa ra |
|---|---|
| ACM | ACM (2017), "*Quy tắc đạo đức và ứng xử nghề nghiệp ACM 2018: Dự thảo 3*", Hiệp hội Máy tính thuộc Ủy ban Đạo đức nghề nghiệp, https://ethics.acm.org/2018-code-draft-3/<br>USACM (2017), "*Tuyên bố về tính minh bạch và trách nhiệm giải trình của thuật toán*", Hiệp hội máy tính thuộc Hội đồng chính sách công Hoa Kỳ, www.acm.org/binaries/content/assets/public-policy/2017_usacm_statement_algorithms.pdf |
| An toàn | Amodei, D. và cộng sự. (2016), "*Các vấn đề cụ thể về an toàn TTNT*", ngày 25 tháng 7, https://arxiv.org/pdf/1606.06565.pdf |
| Asilomar | FLI (2017), "*Nguyên tắc TTNT của Asilomar*", Viện Nghiên cứu tương lai của sự sống, https://futureoflife.org/ai-principles/ |
| COMEST | COMEST (2017), "*Báo cáo của COMEST về đạo đức rô bốt*", Ủy ban thế giới về đạo đức của tri thức khoa học và công nghệ,<br>http://unesdoc.unesco.org/images/0025/002539/253952E.pdf |
| Economou | Economou, N. (2017) "*TTNT "có nguyên tắc"có thể cải thiện công lý"*, 3 tháng 10, tạp chí Aba,<br>www.abajournal.com/legalrebels/article/a_principled_artificial_intelligence_could_improve_justice |
| EGE | EGE (2018), "*Tuyên bố TTNT, rô bốt và hệ thống tự động*", Nhóm Đạo đức khoa học và công nghệ mới châu Âu, |

| | |
|---|---|
| | http://ec.europa.eu/research/ege/pdf/ege_ai_statement_2018.pdf |
| EPSRC | EPSRC (2010), "*Nguyên tắc của rô bốt*", Hội đồng Nghiên cứu Khoa học kỹ thuật và vật lý https://epsrc.ukri.org/research/ourportfolio/themes/engineering/activities/principlesofrobotics/ |
| FATML | FATML (2016), "*Nguyên tắc với thuật toán có thể giải trình và Tuyên bố về tác động xã hội đến thuật toán*", Công bằng, Trách nhiệm giải trình và Tính minh bạch của máy học, www.fatml.org/resources/principles-for-accountable-algorithms |
| FPF | FPF (2018), "K*hả năng giải thích: Hướng dẫn thực tiễn để Quản lý rủi ro trong mô hình học máy*", Diễn đàn tương lai của quyền riêng tư, https://fpf.org/wp-content/uploads/2018/06/Beyond-Explainability.pdf |
| Google | Google (2018), "*TTNT tại Google: Nguyên tắc của chúng tôi*", https://www.blog.google/technology/ai/ai-principles/ |
| IEEE | IEEE (2017), Sáng kiến toàn cầu về đạo đức của các hệ thống tự động và thông minh, "*Thiết kế phù hợp về đạo đức phiên bản 2*", Viện Kỹ sư điện và điện tử, http://standards.ieee.org/develop/indconn/ec/ead_v2.pdf |
| Intel | Intel (2017), "*TTNT - Cơ hội chính sách công*", https://blogs.intel.com/policy/files/2017/10/Intel-Artificial-Intelligence-Public- Policy-White-Paper-2017.pdf |
| ITI | ITI (2017), "*Nguyên tắc chính sách TTNT*", Hội đồng Công nghiệp công nghệ thông tin, www.itic.org/resources/AI-Policy-Principles- FullReport2.pdf |
| JSAI | JSAI (2017), "*Hướng dẫn đạo đức của Hiệp hội TTNT Nhật Bản*", Hiệp hội TTNT Nhật Bản, http://ai-elsi.org/wp-content/uploads/2017/05/JSAI-Ethical-Guidelines-1.pdf |
| MIC | MIC (2017), "*Dự thảo Hướng dẫn NC&PT TTNT cho các cuộc thảo luận quốc tế*", Bộ Nội vụ và Truyền thông Nhật Bản, www.soumu.go.jp/main_content/000507517.pdf |
| MIC | MIC (2018), "*Dự thảo Nguyên tắc sử dụng TTNT*", Bộ Nội vụ và Truyền thông Nhật Bản, www.soumu.go.jp/main_content/000581310.pdf |
| Montreal | UoM (2017), "*Tuyên bố Montreal về sự phát triển TTNT có trách nhiệm"*, Đại học Montreal, www.montrealdeclaration-responsibleai.com/ |
| Nadella | Nadella, S. (2016) "*Hợp tác của tương lai*", www.slate.com/articles/technology/future_tense/2016/06/microsoft_ceo_satya_nadella_humans_and_a_i_can_work_together_to_solve_society.html |
| PAI | PAI (2016), "TENETS", Hợp tác TTNT, www.partnershiponai.org/tenets/ |
| Polonski | Polonski, V. (2018) "*Vấn đề khó khăn của đạo đức TTNT - Ba nguyên tắc tích hợp đạo đức vào máy móc*", www.oecd-forum.org/users/80891-dr-vyacheslav-polonski/posts/30743-the-hard- problem-of-ai-ethics-three-guidelines-for-building-morality-into-machines |
| Taddeo and Floridi | Taddeo, M. và L. Floridi (2018), "*Làm thế nào để TTNT trở thành động lực tốt*", Science, 24 tháng 8, Vol. 61/6404, trang 751-752 |

| | |
|---|---|
| | http://science.sciencemag.org/content/361/6404/751 |
| Liên minh Tiếng nói công chúng | UGAI (2018), "*Nguyên tắc chung về TTNT*", Liên minh tiếng nói công chúng, https://thepublicvoice.org/ai- universal-guidelines/ |
| Tuyên bố Tokyo | Trung tâm Nghiên cứu trí tuệ Nhân tạo thế hệ mới (2017), "*Tuyên bố Tokyo - Hợp tác TTNT theo hướng có lợi*", www.ai.u-tokyo.ac.jp/tokyo-statement.html |
| Twomey | Twomey, P. (2018), "*Hướng tới Khung G20 cho TTNT tại nơi làm việc*", CIGI Papers, No 178, Trung tâm Đổi mới quản lý quốc tế, www.cigionline.org/sites/default/files/documents/Paper%20No.178.pdf |
| Liên minh công đoàn toàn cầu | UNI Global Union (2017), "*10 nguyên tắc hàng đầu về đạo đức TTNT*", www.thefutureworldofwork.org/media/35420/uni_ethical_ai.pdf |

Một số sáng kiến đã xây dựng những bộ nguyên tắc có giá trị định hướng phát triển TTNT (Bảng 1), trong đó một số tập trung vào các cộng đồng kỹ thuật thực hiện NC&PT hệ thống TTNT. Nhiều nguyên tắc đã được xây dựng dựa vào các quy trình có sự tham gia của nhiều bên liên quan. Tuy nhiên, các bộ nguyên tắc đó có thể được phân loại thành 5 nhóm: kỹ thuật, khu vực tư nhân, chính phủ, nghiên cứu và lao động. Cộng đồng kỹ thuật bao gồm Viện Nghiên cứu tương lai của sự sống; IEEE; Hiệp hội TTNT Nhật Bản; Liên minh Công bằng, Trách nhiệm giải trình và Minh bạch trong máy học; và Hiệp hội Máy tính. Khu vực tư nhân gồm có Liên minh hợp tác TTNT; Hội đồng Công nghệ thông tin công nghiệp; và Satya Nadella, giám đốc điều hành Microsoft. Về phía chính phủ có Bộ Nội vụ và Truyền thông Nhật Bản; Ủy ban thế giới về Đạo đức của khoa học và công nghệ; và Hội đồng Nghiên cứu khoa học vật lý và kỹ thuật. Cộng đồng nghiên cứu gồm Đại học Montréal và Nicolas Economou, giám đốc điều hành của công ty H5 và cố vấn đặc biệt về Sáng kiến TTNT của xã hội tương lai tại Trường Harvard Kennedy. Liên minh công đoàn toàn cầu là một ví dụ về cộng đồng lao động.

Các chủ đề chung được xác định dựa vào các sáng kiến. Các bên liên quan đã đưa ra các hướng dẫn về giá trị và quyền con người; không phân biệt đối xử; nhận thức và kiểm soát; truy cập dữ liệu; quyền riêng tư và kiểm soát; an toàn và bảo mật; kỹ năng; tính minh bạch và khả năng giải thích; trách nhiệm giải trình và trách nhiệm; đối thoại toàn xã hội; và đo lường.

Tháng 5/2018, Ủy ban Chính sách kinh tế số của OECD đã thành lập Nhóm chuyên gia TTNT của OECD (AIGO) để xây dựng các nguyên tắc về chính sách công và hợp tác quốc tế tạo niềm tin và thúc đẩy việc áp dụng TTNT. Động thái này đã dẫn đến sự ra đời của các Khuyến nghị TTNT của OECD vào ngày 22/5/2019, trong đó có sự tham gia của chính phủ 42 quốc gia.

### 3.2. Sáng kiến quốc gia

Nhiều quốc gia đã công bố các chiến lược và sáng kiến chính sách TTNT quốc gia, thường nhằm đảm bảo vị trí dẫn đầu trong lĩnh vực TTNT. Các chiến lược và sáng kiến TTNT đặt ra các mục tiêu và nhiệm vụ đòi hỏi sự phối hợp hành động của tất cả các bên liên quan. Chính phủ đóng vai trò tập hợp và hỗ trợ. Hộp 2 cung cấp các yếu tố thường thấy trong các chính sách và biện pháp nhằm tăng năng lực cạnh tranh TTNT của quốc gia. Ngoài ra, một số quốc gia đã thành lập hoặc giao trách

nhiệm cho một tổ chức công cụ thể về vấn đề dữ liệu và TTNT.

***Hộp 2.*** Cách các quốc gia phát triển lợi thế cạnh tranh về TTNT

Porter đã xác định bốn yếu tố quan trọng để đạt được lợi thế cạnh tranh quốc gia trong một ngành cụ thể: i) yếu tố điều kiện; ii) điều kiện nhu cầu; iii) các ngành công nghiệp hỗ trợ có liên quan; và iv) chiến lược, cấu trúc và cạnh tranh mạnh mẽ. Porter thừa nhận các công ty là chủ thể tạo lợi thế cạnh tranh trong các ngành. Tuy nhiên, ông nhấn mạnh vai trò chủ chốt của chính phủ trong việc hỗ trợ và tạo điều kiện cho bốn yếu tố quyết định quá trình phát triển công nghiệp quốc gia.

• *Yếu tố điều kiện*: Yếu tố này phụ thuộc vào vị trí địa lý, lao động tay nghề cao, trình độ học vấn và năng lực nghiên cứu. Các quốc gia đang tăng cường năng lực nghiên cứu TTNT thông qua các biện pháp khác nhau bao gồm: i) thành lập các viện nghiên cứu TTNT; ii) đào tạo bằng sau đại học và tiến sĩ liên quan đến TTNT tại các trường đại học và điều chỉnh các bằng cấp hiện có để đưa vào các khóa học TTNT, ví dụ: trong các ngành khoa học;

• *Điều kiện nhu cầu*: Một số quốc gia xác định các lĩnh vực chiến lược để phát triển TTNT, đặc biệt là giao thông vận tải, chăm sóc sức khỏe và dịch vụ công. Các nước này đang đưa ra các biện pháp kích thích nhu cầu tiêu dùng trong nước đối với các dịch vụ TTNT trong 3 lĩnh vực nêu trên. Trong các dịch vụ công, chính phủ một số quốc gia đang đảm bảo các hệ thống TTNT phải đáp ứng các tiêu chuẩn nhất định như độ chính xác hoặc tính mạnh mẽ thông qua các chính sách mua sắm công.

• *Các ngành công nghiệp hỗ trợ có liên quan*: Khả năng cạnh tranh của TTNT yêu cầu quyền truy cập vào cơ sở hạ tầng và dịch vụ kỹ thuật số, dữ liệu, sức mạnh tính toán và kết nối băng thông rộng. Một số quốc gia đang lập kế hoạch cho các cụm công nghệ tập trung vào TTNT và cấu trúc hỗ trợ DNVVN.

• *Chiến lược, cấu trúc và cạnh tranh mạnh mẽ*: Một số cách tiếp cận mà các quốc gia đang áp dụng để thúc đẩy đầu tư tư nhân và cạnh tranh TTNT bao gồm: i) chuẩn bị lộ trình phát triển TTNT để thúc đẩy đầu tư tư nhân; ii) khuyến khích các công ty TTNT quốc tế đầu tư trong nước như mở phòng thí nghiệm TTNT; và iii) thử nghiệm các cơ chế pháp lý cho ứng dụng TTNT để khuyến khích các công ty đổi mới.

Ngoài ra, để thực hiện hiệu quả các sáng kiến TTNT quốc gia, nhiều nước đang xem xét các cơ chế quản trị phù hợp đảm bảo phương pháp tiếp cận phối hợp toàn bộ chính phủ. Ví dụ: Pháp đã thiết lập chức năng điều phối TTNT trong Văn phòng Thủ tướng để thực hiện chiến lược AI quốc gia.

*Nguồn: Porter (1990 [4]), "Lợi thế cạnh tranh của các quốc gia", https://hbr.org/1990/03/the-competitive- advantage-of-nations.*

**Canada**

Canada đang tìm cách trở thành nước dẫn về TTNT, đặc biệt là với Chiến lược TTNT Pan-Canada được công bố vào tháng 3/2017. Chiến lược này được định hướng bởi Viện Nghiên cứu cao cấp phi lợi nhuận của Canada và được chính phủ hỗ trợ tài chính 100 triệu USD. Trong vòng 5 năm, các quỹ sẽ hỗ trợ các chương trình phát triển nguồn nhân lực của Canada, nghiên cứu TTNT ở Canada và chuyển các nghiên cứu TTNT sang ứng dụng trong khu vực công và tư. Các mục tiêu của Chiến lược TTNT Pan - Canada bao gồm:

1. Tăng số lượng các nhà nghiên cứu và sinh viên tốt nghiệp có tay nghề cao ở

Canada.

2. Thiết lập các nút kết nối về khoa học xuất sắc tại ba viện TTNT lớn của Canada: Edmonton (Viện Trí tuệ máy Alberta), Montreal (Viện Nghiên cứu thuật toán Montreal) và Toronto (Viện TTNT Vector).

3. Triển khai chương trình TTNT toàn cầu trong xã hội và lãnh đạo tư duy toàn cầu về tác động kinh tế, xã hội, đạo đức, chính sách và pháp lý của những tiến bộ TTNT.

4. Hỗ trợ cộng đồng nghiên cứu TTNT quốc gia.

Chính phủ liên bang thông qua Hội đồng Nghiên cứu quốc gia Canada (NRC) lập kế hoạch đầu tư nghiên cứu có tổng trị giá 40 triệu USD trong 7 năm để áp dụng TTNT cho các chủ đề chương trình chính bao gồm: phân tích dữ liệu, TTNT cho thiết kế, an ninh mạng, ngôn ngữ bản địa của Canada, hỗ trợ các siêu cụm liên bang và các trung tâm hợp tác với các trường đại học của Canada, cũng như hợp tác chiến lược với các đối tác quốc tế.

Ngoài tài trợ của liên bang, chính phủ Quebec đang phân bổ 80 triệu USD cho cộng đồng TTNT ở Montreal; Ontario đang cung cấp 40 triệu USD cho Viện TTNT Vector. Năm 2016, Quỹ Nghiên cứu xuất sắc Canada đã phân bổ 75 triệu USD cho ba trường đại học hàng đầu nghiên cứu về học sâu: Đại học Montréal, Đại học Bách khoa Montréal và Đại học HEC Montréal. Facebook và các công ty tư nhân năng động khác như ElementAI cũng đang hoạt động ở Canada.

Chính phủ Quebec đưa ra kế hoạch lập trạm quan sát thế giới về tác động xã hội của TTNT và công nghệ số. Tháng 3/2018, một hội thảo được tổ chức để xem xét nhiệm vụ và mô hình tiềm năng, phương thức quản lý, tài trợ và hợp tác quốc tế của trạm quan sát, cũng như các lĩnh vực và vấn đề trọng tâm. Chính phủ Quebec đã phân bổ 3,7 triệu USD để xây dựng trạm quan sát.

Canada cũng đang phối hợp với các đối tác quốc tế để thúc đẩy sáng kiến TTNT. Ví dụ, tháng 7/2018, chính phủ Canada và Pháp đã thông báo sẽ phối hợp thành lập Ủy ban TTNT quốc tế. Nhiệm vụ của Ủy ban sẽ là hỗ trợ và định hướng việc áp dụng TTNT có trách nhiệm lấy con người làm trung tâm và dựa vào quyền con người, sự hòa nhập, đa dạng, đổi mới và tăng trưởng kinh tế.

**Trung Quốc**

Vào tháng 5/2016, chính phủ Trung Quốc đã công bố kế hoạch TTNT quốc gia trong 3 năm do Ủy ban Cải cách và Phát triển quốc gia, Bộ Khoa học và Công nghệ, Bộ Công nghiệp và Công nghệ thông tin và Cục Quản lý Không gian mạng Trung Quốc phối hợp xây dựng. TTNT được lồng ghép vào sáng kiến Internet Plus công bố năm 2015 như chiến lược quốc gia nhằm thúc đẩy tăng trưởng kinh tế theo định hướng các công nghệ sáng tạo liên quan đến Internet trong giai đoạn 2016 - 2018. Sáng kiến này tập trung: i) nâng cao năng lực phần cứng về TTNT; ii) hệ sinh thái nền tảng mạnh mẽ; iii) Ứng dụng TTNT trong các lĩnh vực kinh tế - xã hội quan trọng; và iv) tác động của TTNT đến xã hội. Trong đó, Chính phủ Trung Quốc xây dựng thị trường trị giá 15 tỷ USD vào năm 2018 thông qua NC&PT ngành công nghiệp TTNT của Trung Quốc.

Giữa năm 2017, Hội đồng nhà nước Trung Quốc đã ban hành Hướng dẫn về Kế hoạch phát triển TTNT thế hệ mới, đưa ra quan điểm TTNT dài hạn của Trung Quốc với các mục tiêu của ngành công nghiệp cho từng thời kỳ. Các yếu tố này bao gồm: i) Tăng trưởng kinh tế dựa vào TTNT ở Trung Quốc vào năm 2020; ii) đột phá lớn về các lý thuyết cơ bản vào năm 2025 và đột phá trong việc xây dựng xã hội thông minh; và iii) để Trung Quốc trở thành trung tâm đổi mới TTNT toàn cầu vào năm 2030 và xây dựng ngành công nghiệp TTNT trị giá 150 tỷ USD. Kế hoạch đang được triển khai thực hiện trong toàn chính phủ và Trung Quốc đang nỗ lực dẫn đầu về TTNT với sự hỗ trợ của nhà nước và khai thác sự năng động của các công ty tư nhân. Hội đồng Nhà nước Trung Quốc đặt mục tiêu để "công nghệ thông tin thế hệ mới" trở thành ngành chiến lược chiếm 15% tổng sản phẩm quốc nội năm 2020.

Trong khung thời gian của Kế hoạch 5 năm lần thứ 13 (2016 - 2020), Trung Quốc đặt ra tham vọng trở thành nước dẫn đầu về khoa học và công nghệ với 16 siêu dự án "Đổi mới Khoa học và Công nghệ 2030" trong đó có siêu dự án "TTNT 2.0". Kế hoạch này đã tạo động lực hành động trong khu vực công. Kế hoạch yêu cầu các công ty đẩy nhanh NC&PT phần cứng và phần mềm TTNT, bao gồm tầm nhìn dựa vào TTNT, nhận dạng giọng nói và sinh trắc học, giao diện người-máy và điều khiển thông minh.

Ngày 18/1/2018, Trung Quốc đã thành lập một nhóm tiêu chuẩn hóa TTNT quốc gia và một nhóm chuyên gia cố vấn TTNT quốc gia. Đồng thời, Ủy ban Quản lý tiêu chuẩn hóa quốc gia đã công bố Sách trắng Tiêu chuẩn hóa TTNT, được soạn thảo với sự hỗ trợ của Viện Tiêu chuẩn hóa điện tử Trung Quốc (trực thuộc Bộ Công nghiệp và Công nghệ thông tin).

Các công ty tư nhân Trung Quốc chú ý đến TTNT trước khi chính phủ có động thái hỗ trợ gần đây. Các công ty Trung Quốc như Baidu, Alibaba và Tencent đã đầu tư lớn cho TTNT. Ngành công nghiệp Trung Quốc tập trung vào các ứng dụng và tích hợp dữ liệu, trong khi chính quyền trung ương chú trọng đến các thuật toán cơ bản, dữ liệu mở và khía cạnh khái niệm. Chính quyền thành phố tập trung sử dụng các ứng dụng và dữ liệu mở ở cấp thành phố.

**Pháp**

Tổng thống Pháp Emmanuel Macron đã công bố chiến lược TTNT của Pháp vào ngày 29/3/2018. Theo đó, vào năm 2022, 1,5 tỷ EUR tài trợ công sẽ được phân bổ cho TTNT để Pháp trở thành quốc gia đi đầu trong lĩnh vực nghiên cứu và đổi mới TTNT. Chiến lược kêu gọi đầu tư cho nghiên cứu công và giáo dục, xây dựng các trung tâm nghiên cứu đẳng cấp thế giới liên kết với ngành công nghiệp thông qua hợp tác công tư và thu hút các nhà nghiên cứu TTNT ưu tú của Pháp và của nước ngoài. Để phát triển hệ sinh thái TTNT ở Pháp, cách tiếp cận của chiến lược là đổi mới các ngành công nghiệp hiện có. Chiến lược đề xuất ưu tiên quyền truy cập vào dữ liệu bằng cách tạo ra các "điểm chung dữ liệu" giữa các chủ thể công và tư; điều chỉnh luật bản quyền để tạo điều kiện khai thác dữ liệu; và công khai dữ liệu của khu vực công như dữ liệu y tế cho các đối tác trong ngành.

Chiến lược TTNT cũng phác thảo kế hoạch ban đầu cho những gián đoạn do TTNT gây ra, có lập trường vững vàng về việc chuyển giao dữ liệu ra ngoài châu Âu. Chiến lược đề cập đến việc thành lập cơ quan dữ liệu trung tâm với một nhóm

khoảng 30 chuyên gia cố vấn ứng dụng TTNT trong toàn chính phủ. Ranh giới đạo đức và triết học được nêu ra trong chiến lược bao gồm tính minh bạch của thuật toán như một nguyên tắc cốt lõi. Ví dụ, các thuật toán được phát triển bởi chính phủ Pháp hoặc với sự tài trợ của công chúng, sẽ được cho là mở. Tôn trọng quyền riêng tư và các quyền khác của con người sẽ được "thiết kế riêng". Chiến lược cũng triển khai đào tạo nghề trong các ngành bị TTNT đe dọa. Chiến lược kêu gọi thử nghiệm chính sách trong thị trường lao động và đối thoại về cách chia sẻ giá trị gia tăng do TTNT tạo ra trong chuỗi giá trị.

**Nhật Bản**

Văn phòng Nội các Nhật Bản đã thành lập Hội đồng Chiến lược công nghệ TTNT vào tháng 4/2016 để thúc đẩy NC&PT công nghệ TTNT và các ứng dụng kinh doanh. Tháng 3/2017, Chiến lược Công nghệ TTNT do Hội đồng công bố, đã xác định các vấn đề quan trọng, bao gồm nhu cầu tăng đầu tư, tạo điều kiện sử dụng và truy cập dữ liệu cũng như tăng số lượng các nhà nghiên cứu và kỹ sư TTNT. Chiến lược này cũng xác định những lĩnh vực chiến lược mà TTNT có thể mang lại lợi ích to lớn: năng suất; sức khỏe, chăm sóc y tế và phúc lợi; tính di động; và bảo mật thông tin.

Chiến lược Đổi mới tích hợp của Nhật Bản được Văn phòng Nội các công bố vào tháng 6/2018, bao gồm một tập hợp các hành động chính sách TTNT. Chiến lược này liên quan đến các cuộc thảo luận giữa nhiều bên liên quan về các vấn đề đạo đức, luật pháp và xã hội của TTNT. Từ đó dẫn đến việc Văn phòng Nội các công bố Nguyên tắc xã hội lấy con người làm trung tâm vào tháng 4/2019.

Tại cuộc họp Bộ trưởng Công nghệ thông tin và truyền thông G7 ở Takamatsu vào tháng 4/2016, Nhật Bản đã đề xuất các nguyên tắc chung cho NC&PT TTNT. Một nhóm chuyên gia đã xây dựng Dự thảo Hướng dẫn NC&PT TTNT cho các cuộc thảo luận quốc tế, được Bộ Nội vụ và Truyền thông Nhật Bản công bố tháng 7/2017. Các hướng dẫn này chủ yếu nhằm cân bằng lợi ích và rủi ro của mạng TTNT, đồng thời đảm bảo sự trung tính của công nghệ và tránh gánh nặng quá mức cho các nhà phát triển. Các hướng dẫn bao gồm chín nguyên tắc mà các nhà nghiên cứu và phát triển hệ thống TTNT nên xem xét. Bảng 2 cung cấp tóm tắt của các hướng dẫn.

***Bảng 2.*** Nguyên tắc NC&PT được đưa ra trong Hướng dẫn NC&PT TTNT

| Nguyên tắc | Các nhà phát triển |
|---|---|
| I. Hệ thống hợp tác | Chú ý đến tính liên kết và khả năng tương tác của TTNT |
| II. Tính minh bạch và khả năng giải thích | Chú ý đến khả năng xác minh đầu vào/đầu ra của hệ thống TTNT và khả năng giải thích các quyết định |
| III. Khả năng kiểm soát | Chú ý đến khả năng kiểm soát của hệ thống TTNT |
| IV. An toàn | Đảm bảo hệ thống TTNT không gây hại đến tính mạng, cơ thể hoặc tài sản của người dùng hoặc bên thứ ba thông qua bộ truyền động hoặc các thiết bị khác. |
| V. Bảo mật | Chú ý đến tính bảo mật của hệ thống TTNT |
| VI. Quyền riêng tư | Cân nhắc về việc hệ thống TTNT sẽ không xâm phạm quyền |

| | |
|---|---|
| | riêng tư của người dùng hoặc bên thứ ba |
| VII. Đạo đức | Tôn trọng phẩm giá con người và quyền tự chủ của cá nhân trong NC&PT hệ thống TTNT |
| VIII. Hỗ trợ người dùng | Cân nhắc về việc hệ thống TTNT sẽ hỗ trợ người dùng và giúp họ có cơ hội lựa chọn cách ứng xử phù hợp. |
| IX. Trách nhiệm giải trình | Nỗ lực thực hiện trách nhiệm giải trình với các bên liên quan, bao gồm cả người dùng hệ thống TTNT. |

*Nguồn: Japan (2017[29]), Draft AI R&D Guidelines for International Discussions, www.soumu.go.jp/main_content/000507517.pdf.*

**Hàn Quốc**

Chính phủ Hàn Quốc đã công bố Chiến lược Phát triển công nghiệp thông tin thông minh vào tháng 3/2016. Năm 2020, chính phủ đã đầu tư công 940 triệu USD cho lĩnh vực TTNT và các công nghệ thông tin liên quan như IoT và điện toán đám mây. Chiến lược này nhằm xây dựng hệ sinh thái công nghiệp thông tin thông minh mới và khuyến khích đầu tư tư nhân 2,3 tỷ USD vào năm 2020. Theo chiến lược, chính phủ đề ra ba mục tiêu. Đầu tiên là kế hoạch khởi động các dự án hàng đầu phát triển TTNT, ví dụ trong các lĩnh vực công nghệ ngôn ngữ - hình ảnh - không gian - trí tuệ cảm xúc. Thứ hai là tìm cách tăng cường kỹ năng TTNT của lực lượng lao động. Thứ ba là thúc đẩy quyền truy cập và sử dụng dữ liệu của chính phủ, các công ty và viện nghiên cứu.

Vào tháng 12/2016, chính phủ Hàn Quốc đã công bố Kế hoạch tổng thể trung và dài hạn để chuẩn bị cho xã hội thông tin thông minh. Kế hoạch này bao gồm các chính sách quốc gia nhằm ứng phó với những thay đổi và thách thức của cuộc Cách mạng công nghiệp lần thứ 4. Để có được tầm nhìn về "xã hội thông minh lấy con người làm trung tâm", kế hoạch đặt mục tiêu xây dựng nền tảng CNTT thông minh đẳng cấp thế giới. Nền tảng CNTT có thể được áp dụng trong các ngành và được sử dụng để cải thiện các chính sách xã hội. Để thực hiện kế hoạch, chính phủ đang tạo ra môi trường thử nghiệm quy mô lớn để phát triển các dịch vụ và sản phẩm mới, bao gồm các dịch vụ công tốt hơn.

Vào tháng 5/2018, chính phủ Hàn Quốc đã công bố kế hoạch quốc gia đầu tư 2 tỷ USD vào năm 2022 để nâng cao năng lực NC&PT TTNT. Kế hoạch sẽ thành lập sáu viện nghiên cứu TTNT; phát triển tài năng TTNT thông qua 4.500 suất học bổng TTNT và các khóa đào tạo chuyên sâu ngắn hạn; và tăng tốc phát triển chip TTNT.

**Singapo**

Cơ quan phát triển truyền thông Infocomm đã công bố Khung hành động kinh tế số vào tháng 5/2018. Đây là khuôn khổ hành động biến Singapo thành nền kinh tế số hàng đầu, xác định TTNT là công nghệ đột phá để thúc đẩy chuyển đổi số của Singapo.

Ủy ban Bảo vệ dữ liệu cá nhân đã công bố khung mô hình quản trị TTNT mẫu vào tháng 1/2019 để thúc đẩy áp dụng TTNT có trách nhiệm ở Singapo. Khung mô hình này cung cấp hướng dẫn thực tế để chuyển đổi các nguyên tắc đạo đức thành các phương thức thực hành. Khung mô hình được xây dựng dựa vào kết quả của các

cuộc thảo luận quốc gia trong Hội nghị bàn tròn giữa các cơ quan quản lý và cơ quan công quyền. Các tổ chức có thể chủ động áp dụng khung mô hình. Đây cũng là cơ sở để phát triển các khuôn khổ quản trị TTNT theo từng lĩnh vực cụ thể.

Hội đồng đa bên tư vấn sử dụng TTNT và dữ liệu có đạo đức được thành lập vào tháng 6/2018. Hội đồng tư vấn cho chính phủ Singapo các vấn đề đạo đức, luật pháp, quy định và chính sách từ việc triển khai thương mại TTNT. Để đạt được mục tiêu đó, Hội đồng tư vấn hỗ trợ ngành công nghiệp áp dụng TTNT và triển khai các sản phẩm và dịch vụ TTNT theo cách có trách nhiệm. Ngoài ra, tháng 9/2018, Singapo đã xây dựng Chương trình nghiên cứu 5 năm về Quản trị TTNT và Sử dụng dữ liệu để đưa quốc gia này thành trung tâm tri thức hàng đầu có chuyên môn quốc tế về các quy định và chính sách TTNT.

**Vương quốc Anh**

Chiến lược Kỹ thuật số của Vương quốc Anh được công bố vào tháng 3/2017 công nhận TTNT là chìa khóa để phát triển nền kinh tế số quốc gia. Chiến lược này đã phân bổ tài trợ 22,3 triệu USD cho các trường đại học để phát triển công nghệ TTNT và robot. Chính phủ đã tăng đầu tư cho NC&PT lên 6,6 tỷ USD trong 4 năm tới, một phần thông qua Quỹ Thách thức chiến lược công nghiệp.

Tháng 10/2017, chính phủ đã công bố báo cáo đánh giá ngành công nghiệp TTNT của quốc gia. Theo báo cáo, Vương quốc Anh là trung tâm tri thức TTNT quốc tế, một phần là nhờ có các nhà khoa học máy tính tiên phong như Alan Turing. Chính phủ Anh ước tính TTNT có thể bổ sung 814 tỷ USD cho nền kinh tế nội địa. Các công cụ AI được sử dụng tại quốc gia này, bao gồm hướng dẫn sức khỏe cá nhân (Your.MD), chatbot (công cụ phần mềm giúp tương tác / nói chuyện tự động với khách hàng) được phát triển cho khách hàng của ngân hàng và nền tảng giúp trẻ em học tập và giáo viên cung cấp các chương trình giáo dục được cá nhân hóa. Báo cáo đã đưa ra 18 khuyến nghị như cải thiện quyền truy cập vào dữ liệu và chia sẻ dữ liệu thông qua quyền quản lý dữ liệu được ủy thác (Data Trust). Báo cáo cũng khuyến nghị cải thiện nguồn cung cấp kỹ năng TTNT thông qua các Chương trình đào tạo thạc sĩ TTNT do ngành tài trợ. Các ưu tiên khác được đề xuất bao gồm tối đa hóa nghiên cứu TTNT bằng cách điều phối nhu cầu về năng lực tính toán cho nghiên cứu TTNT giữa các tổ chức có liên quan; hỗ trợ tiếp nhận TTNT thông qua Hội đồng TTNT của Vương quốc Anh; và phát triển khuôn khổ để cải thiện tính minh bạch và trách nhiệm giải trình đối với các quyết định dựa vào TTNT.

Chính phủ Anh đã công bố chiến lược công nghiệp quốc gia vào tháng 11/2017. Chiến lược này xác định TTNT là một trong bốn "thách thức lớn" nhằm đưa Vương quốc Anh lên vị trí hàng đầu trong các ngành công nghiệp tương lai và đảm bảo khai thác những thay đổi lớn trên toàn cầu. Tháng 4/2018, Vương quốc Anh đã công bố Thỏa thuận trong lĩnh vực TTNT: gói đầu tư 1,2 tỷ USD được đưa ra dựa vào thế mạnh của quốc gia và tìm cách duy trì hệ sinh thái đẳng cấp thế giới. Thỏa thuận đề cập đến ba lĩnh vực trọng tâm chính: kỹ năng và tài năng; động lực cho ứng dụng; và dữ liệu và hạ tầng.

Chính phủ đã thành lập Văn phòng Trí tuệ thân tạo (OAI) tập trung thực hiện Thỏa thuận theo ngành và thúc đẩy áp dụng TTNT rộng rãi. Ngoài ra, chính phủ cũng thành lập Trung tâm Đạo đức dữ liệu và Đổi mới sáng tạo nhằm tăng cường bối

cảnh quản trị cho phép đổi mới sáng tạo và tạo niềm tin cho người dân. Trung tâm sẽ cung cấp cho chính phủ tư vấn độc lập của chuyên gia về các biện pháp cần thiết cho đổi mới an toàn, đạo đức và đột phá về các công nghệ dựa vào dữ liệu và TTNT.

**Hoa Kỳ**

Ngày 11/2/2019, Tổng thống Trump đã ký Sắc lệnh hành pháp 13859 duy trì vai trò đi đầu của Hoa Kỳ về TTNT và khởi động Sáng kiến TTNT. Sáng kiến TTNT định hướng hành động trong năm lĩnh vực chính: i) đầu tư NC&PT TTNT; ii) giải phóng nguồn lực TTNT; iii) hướng dẫn các quy định và tiêu chuẩn kỹ thuật của TTNT; iv) xây dựng lực lượng lao động TTNT; và v) thu hút quốc tế tham gia hỗ trợ nghiên cứu và đổi mới TTNT ở Hoa Kỳ và mở cửa thị trường cho ngành công nghiệp TTNT của quốc gia.

Sáng kiến TTNT là đỉnh cao của một chuỗi các hành động của Chính quyền nhằm thúc đẩy vai trò lãnh đạo của Hoa Kỳ trong lĩnh vực TTNT. Nhà Trắng đã tổ chức Hội nghị thượng đỉnh TTNT đầu tiên cho ngành công nghiệp Hoa Kỳ vào tháng 5/2018, quy tụ các bên liên quan trong công nghiệp, học giả và các nhà lãnh đạo chính phủ. Những người tham gia đã thảo luận về tầm quan trọng của việc dỡ bỏ rào cản đổi mới TTNT ở Hoa Kỳ và thúc đẩy hợp tác NC&PT TTNT giữa các đồng minh của Hoa Kỳ. Những người tham gia cũng đề cập đến sự cần thiết phải nâng cao nhận thức về TTNT để người dân hiểu rõ phương thức hoạt động của các công nghệ này và cách chúng mang lại lợi ích cho cuộc sống hàng ngày. Cũng trong thời điểm này, tờ thông tin có tên *Trí tuệ nhân tạo cho người dân Mỹ* đã liệt kê các chính sách và biện pháp TTNT do chính quyền đưa ra, bao gồm tăng cường tài chính công cho NC&PT TTNT và sửa đổi quy định để đẩy mạnh phát triển và sử dụng máy bay và ô tô không người lái. Các chính sách cũng ưu tiên giáo dục về khoa học, công nghệ, kỹ thuật và toán học (STEM) với trọng tâm là khoa học máy tính, cũng như tăng cường chia sẻ dữ liệu liên bang để nghiên cứu và ứng dụng TTNT.

Gói ngân sách NC&PT cho năm tài chính 2019 và 2020 do tổng thống ban hành, đã xác định TTNT và máy học là những ưu tiên chính. Các lĩnh vực cụ thể bao gồm nghiên cứu TTNT cơ bản tại Quỹ Khoa học quốc gia và NC&PT ứng dụng tại Bộ Giao thông. Các ưu tiên nghiên cứu cũng bao gồm phân tích sức khỏe nâng cao tại Viện Y tế quốc gia và hạ tầng điện toán TTNT tại Bộ Năng lượng. Nhìn chung, đầu tư của chính phủ liên bang cho NC&PT TTNT và các công nghệ liên quan đã tăng hơn 40% kể từ năm 2015.

Tháng 9/2018, Ủy ban Lựa chọn TTNT thuộc Hội đồng Khoa học và Công nghệ quốc gia bắt đầu sửa đổi Kế hoạch Chiến lược nghiên cứu và phát triển TTNT quốc gia. Kể từ khi kế hoạch được công bố vào năm 2016, việc sử dụng và triển khai thương mại TTNT đã có bước phát triển nhanh chóng. Để phục vụ cho quá trình sửa đổi, Ủy ban Lựa chọn TTNT đã lấy ý kiến đóng góp của người dân từ những người trực tiếp thực hiện NC&PT TTNT và những người bị ảnh hưởng bởi TTNT.

Chính quyền cũng đã ưu tiên đào tạo lực lượng lao động quốc gia trong tương lai. Tổng thống Trump đã ký Sắc lệnh hành pháp quy định thành lập các cơ sở học nghề và Nhóm đặc nhiệm cấp nội các về việc mở rộng thời gian học nghề. Bên cạnh đó, Bản ghi nhớ của Tổng thống cũng ưu tiên giáo dục STEM chất lượng cao, đặc biệt là giáo dục khoa học máy tính với cam kết tài trợ 200 triệu USD.

Quốc hội Hoa Kỳ đã tổ chức cuộc họp kín về TTNT giữa hai đảng vào tháng 5/2017, do các nghị sĩ John K. Delaney và Pete Olson đồng chủ trì. Cuộc họp kín quy tụ các chuyên gia từ học viện, chính phủ và khu vực tư nhân để thảo luận tác động của công nghệ TTNT. Quốc hội đang xem xét luật thành lập cả ủy ban cố vấn TTNT liên bang và các tiêu chuẩn an toàn liên bang cho xe tự lái.

### 3.3. Các sáng kiến liên chính phủ

#### G7 và G20

Tại Hội nghị Bộ trưởng CNTT-TT G7 năm 2016 ở Takamatsu (Nhật Bản), Bộ trưởng Bộ Nội vụ và Truyền thông Nhật Bản đã trình bày và thảo luận bộ Nguyên tắc NC&PT TTNT.

Hội nghị Bộ trưởng Công nghiệp và CNTT-TT G7 được tổ chức tại Turin vào tháng 9/2017 dưới sự chủ trì của Ý, đã ban hành Tuyên bố cấp Bộ trưởng, trong đó các nước G7 thừa nhận những lợi ích tiềm năng to lớn của TTNT đối với xã hội và nền kinh tế, đồng thời nhất trí cách tiếp cận TTNT lấy con người làm trung tâm.

Dưới thời Canada làm chủ tịch G7 năm 2018, các Bộ trưởng G7 đã nhóm họp tại Montreal vào tháng 3/2018 và xác định tầm nhìn về TTNT lấy con người làm trung tâm cũng như tập trung vào quan hệ liên kết giữa việc hỗ trợ tăng trưởng kinh tế từ đổi mới TTNT. Các Bộ trưởng cũng tìm cách tăng lòng tin vào việc áp dụng TTNT, đồng thời tăng tính toàn diện trong việc triển khai TTNT. Các thành viên G7 đã nhất trí hành động trong các lĩnh vực liên quan, bao gồm:

• Đầu tư NC&PT cơ bản ở giai đoạn đầu cho đổi mới TTNT, hỗ trợ khởi nghiệp TTNT và tạo cho lực lượng lao động mức độ sẵn sàng tự động hóa.

• Tiếp tục khuyến khích nghiên cứu, bao gồm giải quyết các thách thức xã hội, thúc đẩy tăng trưởng kinh tế và xem xét các cân nhắc về đạo đức của TTNT, cũng như các vấn đề rộng hơn như ra quyết định tự động.

• Hỗ trợ nâng cao nhận thức cộng đồng để truyền tải những lợi ích thực tế, tiềm năng và tác động trên diện rộng của TTNT.

• Tiếp tục thúc đẩy các phương pháp tiếp cận thích hợp về kỹ thuật, đạo đức và công nghệ trung lập.

• Hỗ trợ luồng thông tin miễn phí thông qua chia sẻ các phương pháp hay và kỹ thuật use case (một kỹ thuật được sử dụng trong kỹ thuật phần mềm và hệ thống để nắm bắt yêu cầu chức năng của hệ thống) về việc cung cấp quyền truy cập mở, mang tính tương tác và an toàn vào dữ liệu TTNT của chính phủ.

• Phổ biến tuyên bố của G7 trên toàn cầu để đẩy mạnh sự phát triển và hợp tác TTNT trên trường quốc tế.

Tại Charlevoix vào tháng 6/2018, G7 đã đưa ra thông cáo chung để xúc tiến mục tiêu TTNT lấy con người làm trung tâm và đẩy mạnh áp dụng TTNT vì mục đích thương mại. Bên cạnh đó, G7 cũng nhất trí đẩy mạnh các phương pháp tiếp cận về kỹ thuật, đạo đức và công nghệ trung lập phù hợp.

Các Bộ trưởng G7 đã quyết định triệu tập một hội nghị đa bên về TTNT do Canada tổ chức vào tháng 12/2018. Hội nghị đã thảo luận phương thức khai thác

tiềm năng chuyển đổi tích cực của TTNT để thúc đẩy tăng trưởng kinh tế toàn diện và bền vững.

G20 cũng rất chú trọng đến TTNT. Năm 2018, Hội nghị Bộ trưởng Kinh tế số G20 tại Salta đã khuyến khích "*các quốc gia tạo điều kiện cho các cá nhân và doanh nghiệp được hưởng lợi từ số hóa và các công nghệ mới nổi*" như 5G, IoT và TTNT. Đến năm 2019, Nhật Bản với vai trò làm chủ tịch G20, tiếp tục công việc G20 đặt ra năm 2018 để thúc đẩy phát triển các ưu tiên khác trong đó có TTNT.

**OECD**

*Các nguyên tắc của OECD về sự tin tưởng trong việc áp dụng TTNT*

Tháng 5/2018, Ủy ban Chính sách Kinh tế số của OECD đã thành lập Nhóm chuyên gia TTNT trong xã hội để hướng tới các nguyên tắc về chính sách công và hợp tác quốc tế nhằm tăng sự tin tưởng trong việc áp dụng TTNT. Những nguyên tắc này đã trở thành cơ sở cho Khuyến nghị của Hội đồng TTNT được 40 quốc gia tuân thủ. Cũng trên tinh thần đó, Chủ tịch Hội đồng Bộ trưởng năm 2018 kêu gọi "OECD theo đuổi các cuộc thảo luận đa bên về khả năng phát triển các nguyên tắc làm nền tảng phát triển và ứng dụng đạo đức của TTNT để phục vụ con người".

Nhóm chuyên gia TTNT bao gồm hơn 50 chuyên gia từ nhiều lĩnh vực và chuyên ngành khác nhau, bao gồm các chính phủ, doanh nghiệp, cộng đồng kỹ thuật, lao động và xã hội dân sự, cũng như Ủy ban châu Âu và UNESCO. Nhóm đã tổ chức bốn hội nghị: hai hội nghị tại OECD ở Paris vào ngày 24-25/9 và ngày 12/11/2018; một hội nghị tại Viện Công nghệ Massachusetts (MIT) ở Cambridge vào ngày 16-17/1/2019; và hai hội nghị cuối cùng tại Dubai vào ngày 8-9/2/2019, bên lề Hội nghị thượng đỉnh chính phủ thế giới. Nhóm chuyên gia đã xác định các nguyên tắc để các bên liên quan quản lý có trách nhiệm TTNT theo cách đáng tin cậy. Các nguyên tắc này bao gồm tôn trọng quyền con người, sự công bằng, tính minh bạch và khả năng giải thích, độ mạnh mẽ và an toàn cũng như trách nhiệm giải trình. Để thực hiện các nguyên tắc này, nhóm cũng đề xuất các khuyến nghị cụ thể cho nhiều chính sách quốc gia. Bước khởi đầu này thông báo về việc xây dựng Khuyến nghị OECD của Hội đồng TTNT vào nửa đầu năm 2019.

*Trạm quan sát (*Observatory*) chính sách TTNT của OECD*

OECD đã lên kế hoạch khởi động Trạm quan sát chính sách TTNT vào năm 2019 để xem xét những tiến bộ của TTNT hiện tại và tương lai cũng như các hàm ý chính sách TTNT. Mục đích là giúp thực hiện các nguyên tắc TTNT nói trên thông qua phối hợp với nhiều bên liên quan bên ngoài, bao gồm chính phủ, ngành công nghiệp, học viện, chuyên gia kỹ thuật và người dân. Trạm quan sát được kỳ vọng trở thành trung tâm đa ngành, dựa trên bằng chứng và đa bên liên quan để thu thập, thảo luận và hướng dẫn bằng chứng liên quan đến chính sách cho các chính phủ. Đồng

thời, Trạm quan sát sẽ cung cấp cho các đối tác bên ngoài thông tin về các hoạt động và phát hiện về TTNT có liên quan đến chính sách tại các quốc gia OECD.

***Ủy ban châu Âu và các tổ chức khác ở châu Âu***

Tháng 4/2018, Ủy ban châu Âu đã đưa ra Thông báo TTNT cho châu Âu đề cập đến ba ưu tiên chính. Thứ nhất là nâng cao năng lực công nghệ và công nghiệp của Liên minh châu Âu và thúc đẩy tiếp nhận TTNT trên toàn nền kinh tế. Thứ hai là chuẩn bị cho những thay đổi kinh tế xã hội do TTNT mang lại. Thứ ba là đảm bảo khuôn khổ pháp lý và đạo đức phù hợp. Ủy ban đã trình bày kế hoạch phối hợp phát triển TTNT ở châu Âu vào tháng 12/2018 chủ yếu nhằm tối đa hóa tác động của các khoản đầu tư và phối hợp xác định con đường phía trước. Kế hoạch dự kiến sẽ được thực hiện đến năm 2027, bao gồm khoảng 70 biện pháp riêng lẻ trong các lĩnh vực sau:

• H*ành động chiến lược và sự phối hợp*: khuyến khích các quốc gia thành viên xây dựng chiến lược TTNT quốc gia quy định các mức đầu tư và biện pháp thực hiện.

• *Tăng tối đa các khoản đầu tư thông qua hợp tác*: thúc đẩy đầu tư nghiên cứu và đổi mới TTNT chiến lược thông qua hợp tác công - tư về TTNT, Nhóm các nhà lãnh đạo và một quỹ cụ thể hỗ trợ các doanh nghiệp khởi nghiệp TTNT và các DNVVN đổi mới sáng tạo.

• *Từ phòng thí nghiệm đến thị trường*: củng cố các trung tâm nghiên cứu xuất sắc và trung tâm đổi mới số, đồng thời thiết lập các cơ sở thử nghiệm và cơ chế thử nghiệm pháp lý.

• *Kỹ năng và học tập suốt đời*: phát triển tài năng, kỹ năng và học tập suốt đời.

• *Dữ liệu*: kêu gọi xây dựng Không gian dữ liệu chung châu Âu tạo điều kiện truy cập dữ liệu được người dân quan tâm và các nền tảng dữ liệu công nghiệp cho TTNT, bao gồm cả dữ liệu y tế.

• *Các bước sử dụng công nghệ có trách nhiệm (*Ethics by design**)** *và khuôn khổ quy định*: nhấn mạnh nhu cầu TTNT có đạo đức và khuôn khổ quy định phù hợp với mục đích (bao gồm các khía cạnh về an toàn và trách nhiệm pháp lý). Khung đạo đức sẽ được xây dựng dựa trên Nguyên tắc đạo đức TTNT. Ủy ban châu Âu cũng cam kết duy trì "Ethics by design" thông qua chính sách mua sắm.

• *TTNT cho khu vực công*: phác thảo các giải pháp TTNT cho khu vực công như mua sắm chung và chuyển đổi.

• *Hợp tác quốc tế*: nhấn mạnh tầm quan trọng của tiếp cận quốc tế và kết hợp TTNT vào chính sách phát triển.

Tháng 6/2018, Ủy ban châu Âu đã thành lập AI HLEG như một phần của Chiến lược TTNT. AI HLEG, bao gồm các đại diện từ học viện, hiệp hội dân sự và ngành công nghiệp, được giao hai nhiệm vụ. Đầu tiên là soạn thảo Nguyên tắc đạo đức TTNT hướng dẫn cho các nhà phát triển, người triển khai và người dùng để đảm bảo "TTNT đáng tin cậy". Thứ hai là chuẩn bị các khuyến nghị đầu tư và chính sách TTNT ("Khuyến nghị") cho Ủy ban châu Âu và các quốc gia thành viên về hướng phát triển TTNT trung và dài hạn nhằm tăng năng lực cạnh tranh toàn cầu của châu Âu. Bên cạnh đó, Ủy ban đã lập diễn đàn Liên minh TTNT châu Âu đa bên để khuyến khích các cuộc thảo luận rộng rãi về chính sách TTNT ở châu Âu. Mọi người có thể đóng góp ý kiến thông qua nền tảng của AI HLEG và cung cấp thông tin cho quá trình hoạch định chính sách của EU.

Tháng 12/2018, AI HLEG đã công bố dự thảo đầu tiên về Hướng dẫn đạo đức để lấy ý kiến. Dự thảo đưa ra khuôn khổ để đạt được TTNT đáng tin cậy dựa vào các quyền cơ bản của EU. Theo đó, TTNT phải hợp pháp, có đạo đức và mạnh mẽ về mặt kỹ thuật xã hội. Các hướng dẫn đưa ra một loạt nguyên tắc đạo đức cho TTNT và xác định các yêu cầu chính để đảm bảo TTNT đáng tin cậy và các phương pháp để thực hiện các yêu cầu này. Cuối cùng, các hướng dẫn đề cập đến danh sách đánh giá không đầy đủ để chuyển từng yêu cầu thành các câu hỏi thực tế giúp các bên liên quan thực hiện các nguyên tắc.

Ngoài ra, Ủy ban châu Âu về Hiệu quả tư pháp đã thông qua Hiến chương Đạo đức châu Âu đầu tiên về việc sử dụng TTNT trong các hệ thống tư pháp vào tháng 12/2018. Văn bản này đưa ra 5 nguyên tắc định hướng phát triển các công cụ TTNT trong các cơ quan tư pháp của châu Âu. Năm 2019, Ủy ban về các vấn đề pháp lý và nhân quyền đã quyết định thành lập một tiểu ban về TTNT và nhân quyền.

Tháng 5/2017, Ủy ban Kinh tế và Xã hội châu Âu (EESC) đã thông qua ý kiến về tác động xã hội của TTNT, kêu gọi các bên liên quan của EU đảm bảo triển khai và sử dụng TTNT theo hướng có lợi cho xã hội và phúc lợi xã hội. Theo EESC, con người nên kiểm soát thời gian và cách thức TTNT được sử dụng trong cuộc sống hàng ngày, đồng thời xác định những 12 lĩnh vực mà TTNT gây lo ngại cho xã hội như đạo đức, an toàn, minh bạch, quyền riêng tư, tiêu chuẩn, lao động, giáo dục, truy cập, luật và quy định, quản trị và dân chủ.

**Liên hiệp quốc**

Tháng 9/2017, Viện Nghiên cứu công lý và tội phạm liên vùng của LHQ đã ký

Thỏa thuận nước chủ nhà để thành lập Trung tâm TTNT và Người máy trong hệ thống LHQ tại The Hague, Hà Lan.

Liên minh Viễn thông quốc tế đã phối hợp với hơn 25 cơ quan khác của LHQ tổ chức Hội nghị thượng đỉnh toàn cầu "TTNT vì điều tốt đẹp". Liên minh cũng hợp tác với các tổ chức như Quỹ XPRIZE và Hiệp hội Máy tính. Sau hội nghị thượng đỉnh đầu tiên vào tháng 6/2017, Liên minh Viễn thông quốc tế đã tổ chức hội nghị thượng đỉnh thứ hai tại Geneva vào tháng 5/2018.

UNESCO đã khởi động một cuộc đối thoại toàn cầu về đạo đức của TTNT do tính phức tạp và tác động của nó đến xã hội và nhân loại. UNESCO đã tổ chức hội nghị bàn tròn công khai với các chuyên gia vào tháng 9/2018, cũng như hội nghị toàn cầu mang tên "Nguyên tắc TTNT: Hướng tới phương pháp tiếp cận nhân văn? - TTNT với các giá trị của con người vì sự phát triển bền vững" vào tháng 3/2019. Mục đích của các hành động này là nâng cao nhận thức và thúc đẩy phản ánh các cơ hội và thách thức do TTNT và các công nghệ liên quan đặt ra. Vào tháng 11/2019, Đại hội đồng lần thứ 40 của UNESCO đã xem xét đưa ra khuyến nghị TTNT trong giai đoạn 2020-21, đã được Ban chấp hành của UNESCO thông qua vào tháng 4/2019.

**Tổ chức Tiêu chuẩn hóa quốc tế (ISO)**

ISO và Ủy ban Kỹ thuật Điện quốc tế (IEC) đã thành lập ủy ban kỹ thuật chung ISO/IEC JTC 1 vào năm 1987 có nhiệm vụ xây dựng các tiêu chuẩn CNTT cho các ứng dụng kinh doanh và tiêu dùng. Tháng 10/2017, tiểu ban 42 (SC 42) được thành lập trực thuộc JTC 1 để xây dựng các tiêu chuẩn TTNT. SC 42 cung cấp hướng dẫn cho các ủy ban ISO và IEC triển khai ứng dụng TTNT. Các hoạt động bao gồm cung cấp khuôn khổ chung, xác định các phương pháp tiếp cận và cấu trúc tính toán của hệ thống TTNT, đồng thời đánh giá các mối đe dọa và rủi ro liên quan.

### 3.4. Sáng kiến của các bên liên quan tư nhân

Các bên liên quan tư nhân đã phối hợp đưa ra sáng kiến đa dạng để thảo luận các vấn đề về TTNT. Tuy nhiên, nội dung dưới đây chủ yếu đề cập đến các sáng kiến bắt nguồn từ cộng đồng kỹ thuật, khu vực tư nhân, người lao động hoặc học viện.

***Cộng đồng kỹ thuật và học viện***

Tháng 4/2016, IEEE đã khởi động Sáng kiến toàn cầu về Đạo đức của các Hệ thống tự động và thông minh nhằm thúc đẩy thảo luận công khai về việc triển khai các công nghệ TTNT và xác định các giá trị ưu tiên và đạo đức. IEEE đã công bố bản phiên bản 2 các Nguyên tắc Thiết kế phù hợp về mặt đạo đức vào tháng 12/2017, kêu gọi cộng đồng đóng góp ý kiến. IEEE dự kiến xuất bản phiên bản cuối cùng của

hướng dẫn thiết kế vào năm 2019 (Bảng 3). Cùng với MIT Media Lab, IEEE đã thành lập Hội đồng Trí tuệ mở rộng vào tháng 6/2018 chịu trách nhiệm thiết kế các hệ thống thông minh có trách nhiệm, dành quyền kiểm soát dữ liệu cá nhân và tạo ra thước đo về sự thịnh vượng kinh tế.

***Bảng 3.*** Nguyên tắc chung trong thiết kế phù hợp về mặt đạo đức của IEEE

| **Nguyên tắc** | **Mục tiêu** |
|---|---|
| Quyền con người | Đảm bảo các hệ thống tự động và thông minh (AIS) không vi phạm quyền con người được quốc tế công nhận |
| Ưu tiên phúc lợi | Ưu tiên các chỉ số đo lường mức độ hạnh phúc trong việc thiết kế và sử dụng AIS vì các thước đo truyền thống về sự thịnh vượng không tính đến tác động đầy đủ của các công nghệ hệ thống TTNT đến sức khỏe con người |
| Trách nhiệm giải trình | Đảm bảo các nhà thiết kế và vận hành AIS theo hướng có trách nhiệm và chịu trách nhiệm giải trình |
| Tính minh bạch | Đảm bảo AIS vận hành theo phương thức minh bạch |
| Lạm dụng AIS và nhận thức về điều đó | Giảm thiểu rủi ro do lạm dụng AIS |

*Source*: IEEE (2017[51]), *Ethically Aligned Design (Version 2),* *http://standards.ieee.org/develop/indconn/ec/ead_v2.pdf.*

Nguyên tắc TTNT được đưa ra tại Hội nghị Asilomar, bao gồm 23 nguyên tắc cho sự phát triển an toàn và có lợi cho xã hội của TTNT trước mắt và lâu dài, là kết quả từ hội nghị của Viện Nghiên cứu sự sống tương lai vào tháng 1/2017. Hội nghị Asilomar đã lựa chọn các nguyên tắc cốt lõi từ các cuộc thảo luận, phản ánh và tài liệu do IEEE, học viện và các tổ chức phi lợi nhuận cung cấp.

Các vấn đề được phân thành ba lĩnh vực. Vấn đề nghiên cứu kêu gọi tài trợ nghiên cứu TTNT bao gồm những khó khăn trong khoa học máy tính; các nghiên cứu kinh tế, luật và xã hội học; "liên kết khoa học-chính sách"; và văn hóa hợp tác nghiên cứu kỹ thuật, độ tin cậy và tính minh bạch. Đạo đức và giá trị đòi hỏi phải thiết kế và vận hành hệ thống TTNT an toàn và bảo mật, minh bạch và có trách nhiệm, bảo vệ quyền tự do của cá nhân, quyền riêng tư, phẩm giá con người, quyền và đa dạng văn hóa, trao quyền trên phạm vi rộng và chia sẻ lợi ích. Các vấn đề lâu dài đặc biệt là tránh đưa ra những giả định chắc chắn giới hạn trên về khả năng của TTNT trong tương lai và lập kế hoạch thận trọng về sự phát triển theo dự báo của

TTNT tổng hợp (AGI). Bảng 4 cung cấp danh sách các Nguyên tắc TTNT của Asilomar.

OpenAI, công ty nghiên cứu TTNT phi lợi nhuận được thành lập cuối năm 2015, sử dụng 60 nhà nghiên cứu toàn thời gian với sứ mệnh "xây dựng AGI an toàn và đảm bảo lợi ích của AGI được phân phối rộng rãi và đồng đều nhất có thể".

Sáng kiến TTNT do tổ chức phi lợi nhuận Future Society đưa ra vào năm 2015, giúp xây dựng khung chính sách TTNT toàn cầu. Sáng kiến TTNT đặt nền tảng trực tuyến cho thảo luận và tranh luận dân sự đa ngành. Nền tảng này giúp hiểu được động lực, lợi ích và rủi ro của công nghệ TTNT để đưa ra các khuyến nghị chính sách.

Ngoài ra còn có nhiều sáng kiến học thuật tại tất cả các nước OECD và nhiều nền kinh tế đối tác. Ví dụ, Sáng kiến Nghiên cứu chính sách Internet của MIT giúp thu hẹp khoảng cách giữa các cộng đồng kỹ thuật và cộng đồng chính sách. Trung tâm Berkman Klein tại Đại học Harvard đã công bố Sáng kiến Đạo đức và Quản trị TTNT vào năm 2017. MIT Media Lab đang tập trung vào các thuật toán và tư pháp, xe tự động cũng như tính minh bạch và khả năng giải thích của TTNT.

***Bảng 4.*** Các nguyên tắc TTNT của Asilomar

| | **Vấn đề nghiên cứu** | **Đạo đức và giá trị** | **Vấn đề lâu dài** |
|---|---|---|---|
| Tiêu đề của các nguyên tắc | - Mục tiêu nghiên cứu<br>-Tài trợ nghiên cứu<br>- Liên kết khoa học - chính sách<br>- Văn hóa nghiên cứu<br>-Tránh chạy đua (race avoidance) | - An toàn<br>- Tính không minh bạch<br>- Minh bạch tư pháp<br>- Trách nhiệm<br>- Liên kết giá trị<br>- Quyền riêng tư cá nhân<br>- Tự do và quyền riêng tư<br>- Lợi ích chung<br>- Thịnh vượng chung<br>- Kiểm soát con người | - Đảm bảo năng lực<br>- Tầm quan trọng<br>- Rủi ro<br>- Tự cải tiến đệ quy (*Recursive self-improvement)*<br>- Lợi ích chung |

*Nguồn:* FLI (2017[53]), *Nguyên tắc TTNT của Asilomar, https://futureoflife.org/ai-principles/.*

***Các sáng kiến của khu vực tư nhân***

Tháng 9/2016, Amazon, DeepMindGoogle, Facebook, IBM và Microsoft đã

khởi động Liên minh Hợp tác TTNT mang lại lợi ích cho con người và xã hội (PAI). Mục đích của PAI là nghiên cứu và đưa ra các phương thức hiệu quả nhất dựa vào công nghệ TTNT, nâng cao hiểu biết của người dân về TTNT và đưa Liên minh trở thành nền tảng mở để thảo luận và thu hút sự chú ý đến TTNT và ảnh hưởng của nó đến con người và xã hội. Kể từ khi thành lập, Liên minh PAI bao gồm các bên liên quan đa ngành với hơn 80 thành viên từ các công ty công nghệ vì lợi nhuận đến đại diện của hiệp hội dân sự, các tổ chức học thuật và nghiên cứu và các công ty khởi nghiệp.

Hội đồng Công nghiệp công nghệ thông tin (ITI) là hiệp hội kinh doanh gồm các công ty công nghệ đặt tại Washington, DC với hơn 60 thành viên. Tháng 10/2017, ITI đã công bố Nguyên tắc chính sách TTNT (Bảng 5), đã xác định trách nhiệm của ngành trong một số lĩnh vực nhất định và kêu gọi chính phủ hỗ trợ nghiên cứu TTNT và hợp tác công - tư.

***Bảng 5.*** Các nguyên tắc chính sách TTNT của ITI

| **Trách nhiệm:**<br>**Thúc đẩy phát triển và sử dụng có trách nhiệm** | **Cơ hội cho chính phủ:**<br>**Đầu tư và hỗ trợ hệ sinh thái TTNT** | **Cơ hội hợp tác công - tư:**<br>**Đẩy mạnh giáo dục suốt đời và tăng tính đa dạng** |
|---|---|---|
| - Thiết kế và triển khai có trách nhiệm<br>- An toàn và khả năng kiểm soát<br>- Dữ liệu đại diện và thiết thực<br>- Khả năng diễn giải<br>- Trách nhiệm của hệ thống TTNT | - Đầu tư NC&PT TTNT<br>- Cách tiếp cận quy định linh hoạt<br>- Thúc đẩy đổi mới và bảo mật của Internet<br>- An ninh mạng và quyền riêng tư<br>- Các tiêu chuẩn toàn cầu và thông lệ tốt nhất | - Dân chủ hóa quyền tiếp cận và tạo sự bình đẳng về cơ hội<br>- Giáo dục khoa học, công nghệ, kỹ thuật và toán học<br>- Lực lượng lao động<br>- Hợp tác công - tư |

***Hiệp hội dân sự***

Liên minh Tiếng nói người dân do Trung tâm Thông tin quyền riêng tư điện tử thành lập, đã công bố *Nguyên tắc chung về TTNT* (UGAI) vào tháng 10/2018. Nguyên tắc này nhấn mạnh đến những thách thức ngày càng tăng của các hệ thống tính toán thông minh và đề xuất các khuyến nghị cụ thể để cải thiện và thông tin về thiết kế của hệ thống. Về bản chất, UGAI làm tăng tính minh bạch và trách nhiệm giải trình của các hệ thống TTNT và tìm cách đảm bảo cho mọi người nắm giữ quyền kiểm soát đối với các hệ thống mà họ xây dựng. Nguyên tắc UGAI đề cập đến

các quyền và nghĩa vụ khác nhau. Các quyền này bao gồm quyền minh bạch và quyền quyết định của con người; các nghĩa vụ đối với việc nhận dạng; công bằng; đánh giá và trách nhiệm giải trình; độ chính xác, độ tin cậy và tính hợp lệ; chất lượng dữ liệu; sự an toàn của người dân; an ninh mạng.

***Các tổ chức lao động***

Liên minh công đoàn toàn cầu đại diện cho hơn 20 triệu công nhân từ hơn 150 quốc gia trong các lĩnh vực kỹ năng và dịch vụ. Tương lai trao quyền cho người lao động và cung cấp công việc tốt là ưu tiên chính của Liên minh công đoàn toàn cầu. Liên minh đã xác định 10 nguyên tắc chính về Đạo đức TTNT nhằm đảm bảo các thỏa thuận tập thể, thỏa thuận khung toàn cầu và các liên minh đa quốc gia liên quan đến công đoàn, đại diện công đoàn và liên minh toàn cầu tôn trọng quyền của người lao động (Bảng 6).

***Bảng 6.*** Mười nguyên tắc hàng đầu cho trí tuệ nhân tạo có đạo đức (Liên minh công đoàn toàn cầu)

| | |
|---|---|
| 1. Hệ thống TTNT cần minh bạch | Người lao động có quyền yêu cầu sự minh bạch trong các quyết định và kết quả của hệ thống TTNT, cũng như các thuật toán cơ bản của chúng. Người lao động cũng cần được tư vấn về việc xây dựng và phát triển hệ thống TTNT. |
| 2. Hệ thống TTNT cần được trang bị "hộp đen" đạo đức | Hộp đen đạo đức không chỉ chứa dữ liệu liên quan để đảm bảo tính minh bạch và trách nhiệm giải trình của hệ thống, mà cả dữ liệu và thông tin rõ ràng về các cân nhắc đạo đức được tích hợp trong hệ thống. |
| 3. TTNT phải phục vụ con người và hành tinh | Các quy tắc đạo đức để phát triển, ứng dụng và sử dụng TTNT là cần thiết trong toàn bộ quá trình vận hành. Hệ thống TTNT vẫn tương thích và làm tăng các nguyên tắc về phẩm giá con người, tính toàn vẹn, tự do, quyền riêng tư và sự đa dạng về văn hóa và giới tính cũng như quyền cơ bản của con người. |
| 4. Áp dụng phương pháp tiếp cận do con người chỉ huy | Sự phát triển của TTNT phải có trách nhiệm, an toàn và hữu ích, trong đó máy móc duy trì hiện trạng pháp lý của các công cụ và pháp nhân giữ quyền kiểm soát và chịu trách nhiệm đối với những máy móc này mọi lúc. |
| 5. Đảm bảo TTNT không thiên vị, phi giới tính | Trong thiết kế và bảo trì TTNT cũng như các hệ thống nhân tạo, hệ thống cần được kiểm soát ở khía cạnh thành kiến tiêu cực hoặc gây hại cho con người. Mọi thành kiến dù là giới tính, chủng tộc, khuynh hướng tình dục hoặc tuổi tác đều được xác định và không được hệ |

| | |
|---|---|
| | thống tuyên truyền. |
| 6. Chia sẻ những lợi ích của hệ thống TTNT | Sự thịnh vượng kinh tế do TTNT tạo ra, nên được phân phối rộng rãi và bình đẳng để mang lại lợi ích cho toàn nhân loại. Do đó, các chính sách quốc gia cũng như toàn cầu nhằm thu hẹp khoảng cách kinh tế, công nghệ và xã hội kỹ thuật số là cần thiết. |
| 7. Đảm bảo quá trình chuyển đổi công bằng và hỗ trợ các quyền và tự do cơ bản: | Khi hệ thống TTNT phát triển và thực tế tăng cường được hình thành, người lao động và các công việc sẽ được thay thế. Điều quan trọng là các chính sách được xây đưa ra để đảm bảo quá trình chuyển đổi công bằng sang thực tế số, bao gồm các biện pháp cụ thể của chính phủ giúp người lao động bị di dời tìm việc làm mới. |
| 8. Thiết lập cơ chế quản trị toàn cầu | Thành lập các cơ quan quản lý TTNT và việc làm theo cách có đạo đức gồm nhiều bên liên quan ở cấp độ toàn cầu và khu vực. Các cơ quan này bao gồm các nhà thiết kế TTNT, nhà sản xuất, chủ sở hữu, nhà phát triển, nhà nghiên cứu, người sử dụng lao động, luật sư, các tổ chức xã hội dân sự và công đoàn. |
| 9. Cấm phân bổ trách nhiệm cho rô bốt | Rôbốt phải được thiết kế và vận hành ở mức có thể tuân thủ các luật hiện hành cũng như các quyền và tự do cơ bản, bao gồm cả quyền riêng tư. |
| 10. Cấm chạy đua vũ trang TTNT | Vũ khí tự động gây chết người, bao gồm cả chiến tranh mạng, nên bị cấm. Liên minh Công đoàn toàn cầu cho rằng cần có thỏa thuận toàn cầu về đạo đức TTNT giúp giải quyết và ngăn chặn những hậu quả tiêu cực ngoài ý muốn của TTNT, đồng thời nhấn mạnh lợi ích của nó đối với người lao động và xã hội. Con người và các tập đoàn là những tác nhân có trách nhiệm. |

# KẾT LUẬN

Hệ thống TTNT theo giải thích của Nhóm chuyên gia TTNT của OECD là "*hệ thống dựa vào máy móc để đưa ra các dự báo, khuyến nghị hoặc quyết định ảnh hưởng đến môi trường thực hoặc ảo nhằm đạt được một tập hợp các mục tiêu nhất định do con người đặt ra. Hệ thống sử dụng đầu vào dựa vào máy móc và/hoặc con người để nhận thức môi trường thực và/hoặc ảo; trừu tượng hóa các nhận thức đó thành mô hình (theo cách tự động như với máy học hoặc theo cách thủ công; và sử dụng suy luận mô hình để đưa ra các tùy chọn cho thông tin hoặc hành động. Hệ thống TTNT được thiết kế để hoạt động với nhiều mức độ tự chủ khác nhau*".

Các giai đoạn trong vòng đời của hệ thống TTNT bao gồm: i) lập kế hoạch và thiết kế, thu thập và xử lý dữ liệu, xây dựng và giải thích mô hình; ii) xác minh và xác nhận; iii) triển khai; và iv) vận hành và giám sát. Phân loại nghiên cứu TTNT phân biệt các ứng dụng TTNT như xử lý ngôn ngữ tự nhiên; các kỹ thuật đào tạo hệ thống TTNT như mạng nơ-ron; tối ưu hóa như học một lần; và nghiên cứu giải quyết các cân nhắc của xã hội như tính minh bạch.

TTNT được ứng dụng trong nhiều ngành từ giao thông, khoa học đến y học và có triển vọng phát triển khi trở thành công nghệ đa dụng. Thông qua các dự đoán, khuyến nghị hoặc quyết định chính xác và ít tốn kém, TTNT hứa hẹn sẽ tăng năng suất, cải thiện sức khỏe và giúp giải quyết những thách thức phức tạp. Để khai thác TTNT cần bổ sung đầu tư cho dữ liệu, kỹ năng và quy trình làm việc được số hóa, cũng như những thay đổi đối với quy trình tổ chức. Do đó, phương thức triển khai có sự khác biệt giữa đơn vị và ngành công nghiệp.

Ngoài những lợi ích, TTNT còn đặt ra các cân nhắc về chính sách công và cần nỗ lực để đảm bảo hệ thống TTNT lấy con người làm trung tâm, đáng tin cậy. TTNT, đặc biệt là một số loại máy học, làm dấy lên lo ngại mới về vấn đề đạo đức và công bằng. Các vấn đề chủ yếu là tôn trọng nhân quyền, các giá trị dân chủ và những nguy cơ của việc chuyển giao các thành kiến từ thế giới vật lý sang thế giới kỹ thuật số. Một số hệ thống TTNT phức tạp đến mức không thể giải thích được các quyết định mà hệ thống đưa ra. Việc thiết kế các hệ thống minh bạch sử dụng TTNT và chịu trách nhiệm về kết quả của chúng là rất quan trọng. Hệ thống TTNT phải hoạt động đúng cách, an toàn và bảo mật.

Các chính sách quốc gia đóng vai trò cần thiết để đẩy mạnh phát triển hệ thống TTNT đáng tin cậy, bao gồm cả chính sách khuyến khích đầu tư NC&PT TTNT có trách nhiệm. Ngoài công nghệ TTNT và khả năng tính toán, TTNT còn khai thác

khối lượng dữ liệu khổng lồ. Điều này làm tăng nhu cầu về môi trường kỹ thuật số cho phép truy cập vào dữ liệu, cùng với các biện pháp bảo vệ dữ liệu và quyền riêng tư. Hệ sinh thái được sự hỗ trợ của TTNT cũng có thể giúp DNVVN điều hướng quá trình chuyển đổi TTNT và đảm bảo môi trường cạnh tranh.

TTNT sẽ thay đổi bản chất của công việc khi nó thay thế và thay đổi các thành phần của lực lượng lao động. Các chính sách sẽ cần tạo thuận lợi cho quá trình chuyển đổi khi mọi người chuyển từ công việc này sang công việc khác và đảm bảo giáo dục, đào tạo và phát triển kỹ năng liên tục.

Nhiều quốc gia đã xây dựng chiến lược TTNT coi TTNT là động cơ tăng trưởng và phúc lợi, tìm cách đào tạo và tuyển dụng các nhà nghiên cứu thế hệ mới, đồng thời xem xét giải quyết hiệu quả những thách thức bắt nguồn từ TTNT. Các bên liên quan phi chính phủ - các tổ chức kinh doanh, kỹ thuật, học viện, xã hội dân sự và tổ chức công đoàn - và các tổ chức quốc tế bao gồm G7, G20, OECD, Ủy ban châu Âu và Liên hợp quốc cũng đang tích cực hành động. Tháng 5/2019, OECD đã thông qua Nguyên tắc TTNT, tiêu chuẩn quốc tế đầu tiên được sự đồng thuận của chính phủ các nước về việc quản lý TTNT đáng tin cậy theo cách có trách nhiệm.

***Biên soạn: Trung tâm Thông tin và Thống kê Khoa học và Công nghệ***

***Tài liệu tham khảo chính:***


1. China (2017), *Guideline on Next Generation AI Development Plan*, Government of China, State Council, http://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm.
2. CIFAR (2017), *Pan-Canadian Artificial Intelligence Strategy*, CIFAR, https://www.cifar.ca/ai/pan-canadian-artificial-intelligence-strategy.
3. Colclough, C. (2018), "Ethical Artificial Intelligence – 10 Essential Ingredients", *A.Ideas Series*, No. 24, The Forum Network, OECD, Paris, https://www.oecd-forum.org/channels/722- digitalisation/posts/29527-10-principles-for-ethical-artificial-intelligence.
4. FLI (2017), *Asilomar AI Principles*, Future of Life Institute (FLI), https://futureoflife.org/ai- principles/.
5. ITI (2017), *AI Policy Principles*, Information Technology Industry Council, https://www.itic.org/resources/AI-Policy-Principles-FullReport2.pdf.
6. OECD (2019), *Artificial Intelligence in Society*, OECD Publishing, Paris, *https://doi.org/10.1787/eedfee77-en*
7. OECD (2019), *Recommendation of the Council on Artificial Intelligence*, OECD, Paris.
8. Price, A. (2018), "First international standards committee for entire AI ecosystem", *IE e-tech*, Issue 03, https://iecetech.org/Technical-Committees/2018-03/First-International-Standards- committee-for-entire-AI-ecosystem.