

→ **Dẫn nhập ngắn về khoa học**

# **TRÍ TUỆ NHÂN TẠO**

Minh họa sinh động bằng tranh

**Henry Brighton & Howard Selina**

Phạm Xuân Trường *dịch*



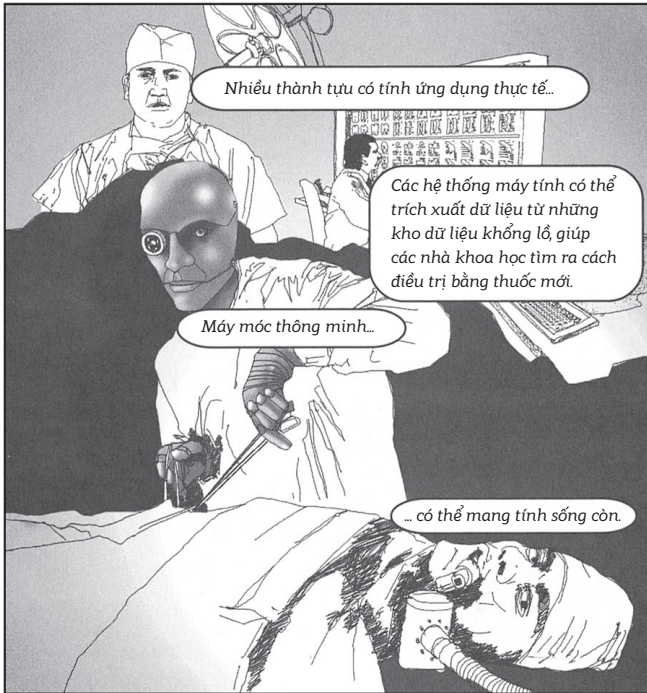
**NHÀ XUẤT BẢN  
DÂN TRÍ**

**ets**

Education  
Technology  
Science

## Trí tuệ nhân tạo

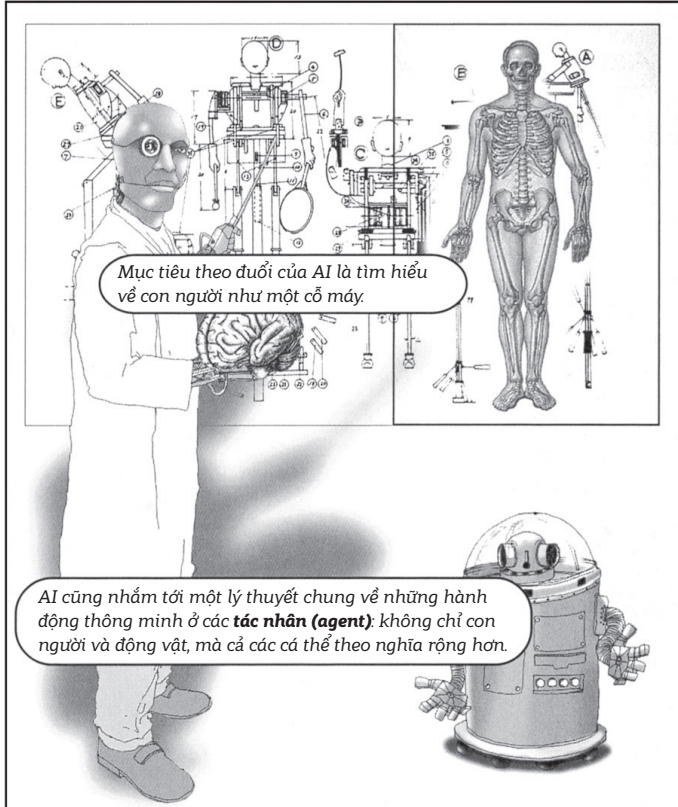
Trong nửa thế kỷ qua, giới khoa học đã tiến hành nhiều nghiên cứu chuyên sâu về một vấn đề liên quan đến việc chế tạo máy móc thông minh – vấn đề *Trí tuệ nhân tạo* (AI). Các nghiên cứu này đã tạo ra những chiếc máy tính chơi cờ có khả năng đánh bại những kỳ thủ giỏi nhất, cũng như các robot hình người, có thể thích nghi với những môi trường mới và tương tác với con người.



Các hệ thống máy tính được lắp đặt ở sân bay nhằm phát hiện hành lý chứa chất gây nổ. Vũ khí quân sự ngày càng phụ thuộc vào nghiên cứu về máy móc thông minh: tên lửa ngày nay có thể tìm mục tiêu với sự trợ giúp của hệ thống thị giác máy.

## Xác định vấn đề về AI

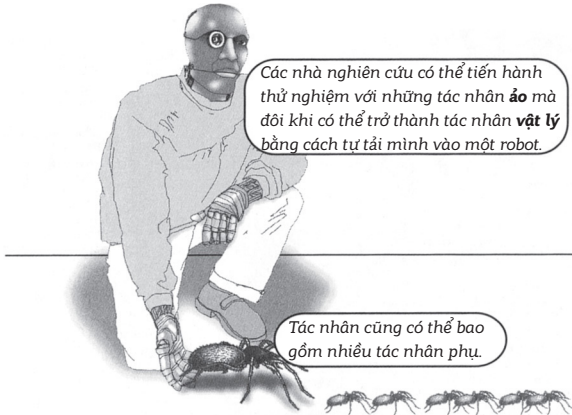
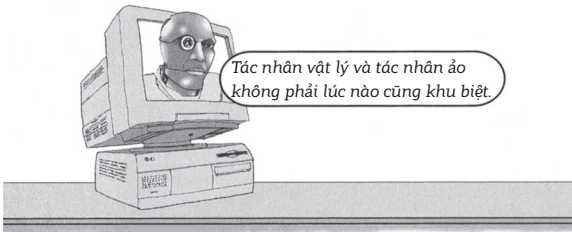
Nghiên cứu về AI đã mang lại thành công cho nhiều dự án kỹ thuật. Nhưng quan trọng hơn, AI đặt ra những câu hỏi vượt xa phạm vi của các ứng dụng kỹ thuật.



Khả năng của một tác nhân có thể vượt xa trí tưởng tượng của chúng ta hiện nay. Lĩnh vực táo bạo này đối đầu trực diện với những luận điểm về triết học đã tồn tại suốt hàng nghìn năm.

## Tác nhân là gì?

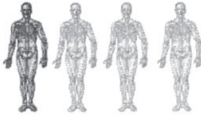
Tác nhân là thứ có khả năng thực hiện hành vi thông minh. Đó có thể là một robot hoặc một chương trình máy tính. *Tác nhân vật lý*, như robot, có thể được giải thích rõ ràng. Chúng là những thiết bị vật lý tương tác với những môi trường vật lý cụ thể. Tuy nhiên, phần lớn nghiên cứu về AI liên quan đến các tác nhân *ảo* hoặc *phần mềm* tồn tại dưới dạng các mô hình chiếm giữ một môi trường ảo bên trong máy tính.



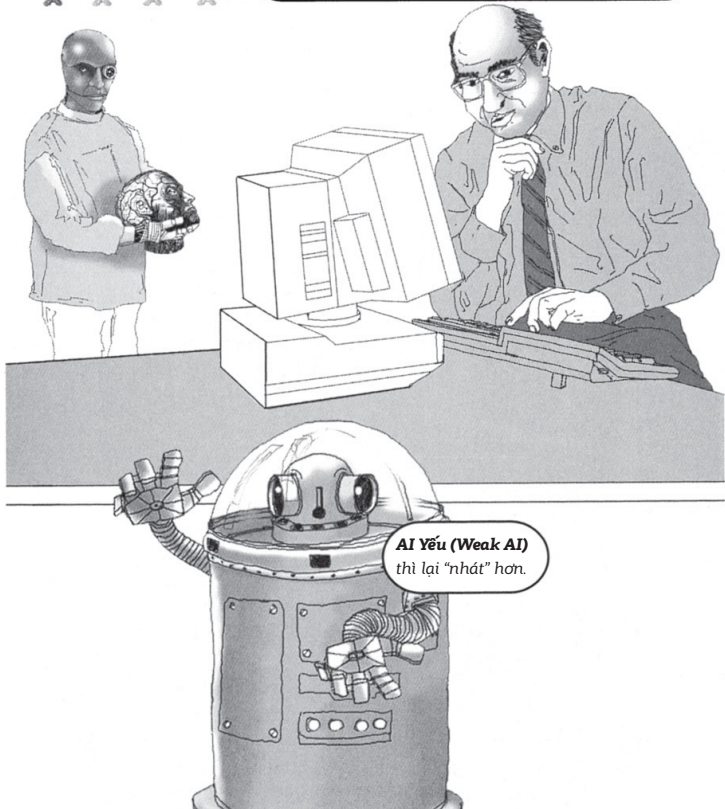
Một số hệ thống AI giải quyết vấn đề bằng cách sử dụng kỹ thuật mà ta vẫn thấy ở các đàn kiến. Như vậy, trong trường hợp này, tác nhân dường như là duy nhất có thể hoạt động dựa vào hành vi kết hợp của hàng trăm tác nhân phụ.

## AI: môn khoa học thực nghiệm

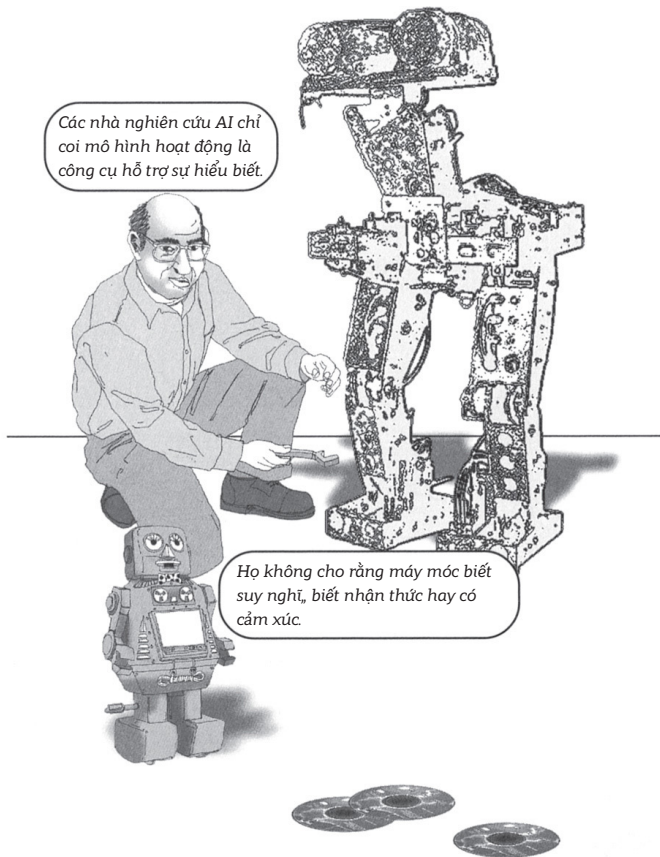
AI là một lĩnh vực đồ sộ. **Marvin Minsky** (1927-2016), một trong những người sáng lập ra ngành AI cho rằng: “AI là một trong những môn khoa học khó nhất từng được nghiên cứu.” AI là sự kết hợp của khoa học và kỹ thuật.



Ở hình thái cao nhất, **AI Mạnh (Strong AI)**, AI nhằm đến việc xây dựng một cỗ máy có thể suy nghĩ, nhận thức và mang cảm xúc. Theo quan điểm này, con người chỉ là những chiếc máy tính tinh vi.



Mục tiêu của AI Yếu là phát triển các lý thuyết về trí thông minh của con người và động vật, sau đó kiểm tra những lý thuyết này bằng cách xây dựng các mô hình hoạt động, thường ở dạng chương trình máy tính hoặc robot.



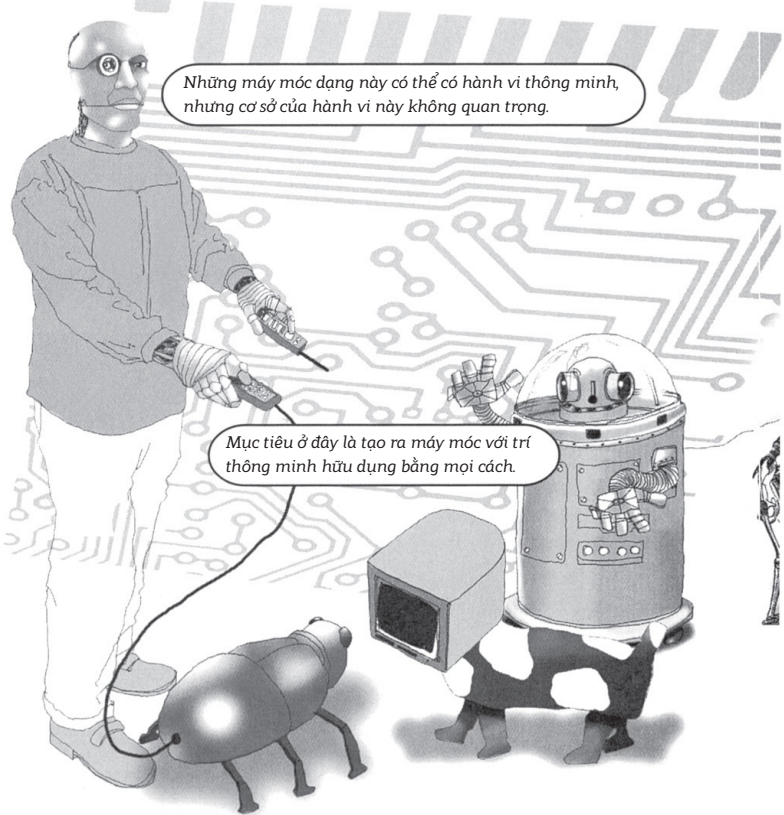
Các nhà nghiên cứu AI chỉ coi mô hình hoạt động là công cụ hỗ trợ sự hiểu biết.

Họ không cho rằng máy móc biết suy nghĩ, biết nhận thức hay có cảm xúc.

Như vậy, đối với AI Yếu, mô hình là một công cụ hữu ích để hiểu về ý thức. Trong khi đó, đối với AI Mạnh, mô hình *chính* là ý thức.

## Kỹ thuật AI ngoại lai (Alien-AI)

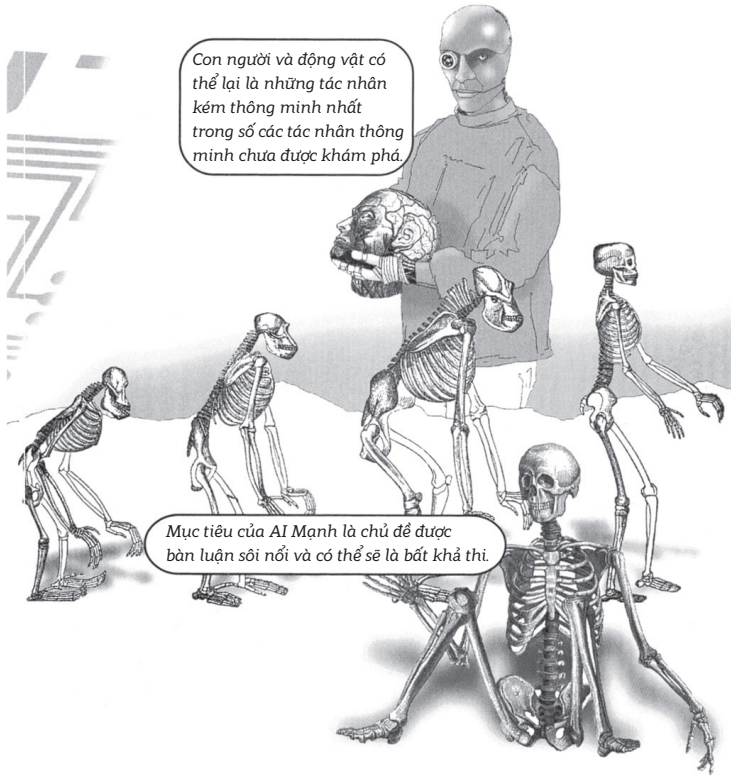
AI cũng hướng tới việc tạo ra những cỗ máy không nhất thiết dựa trên trí thông minh của con người hay động vật.



Do cơ chế hoạt động của những hệ thống như vậy không nhằm đến việc bắt chước cơ chế thông minh của con người, nên đôi khi phương pháp tiếp cận này được gọi là *AI ngoại lai*.

## Giải quyết vấn đề AI

Vì vậy, với một số người, giải quyết vấn đề AI đồng nghĩa với việc tìm cách xây dựng những cỗ máy có khả năng ngang tầm, hoặc cao hơn con người.

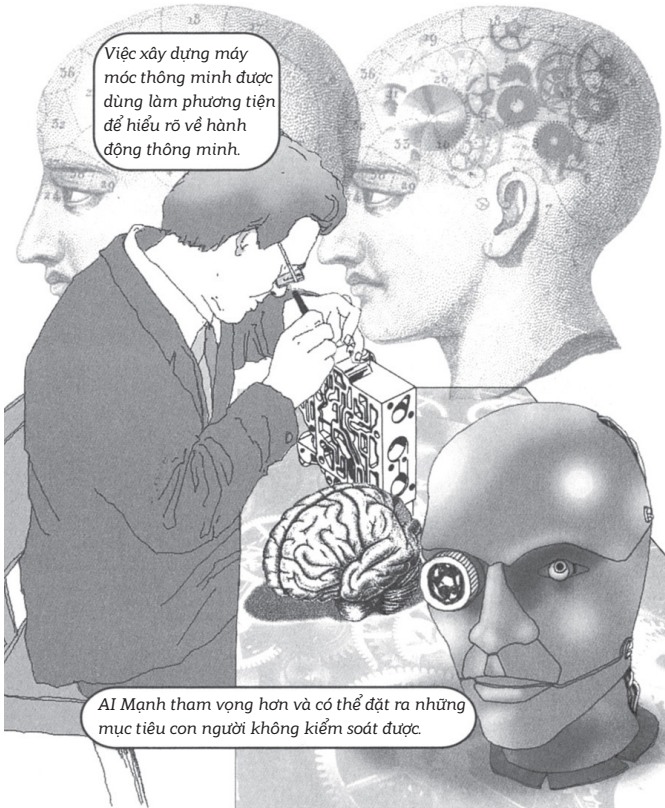


Nhưng đối với hầu hết những nhà nghiên cứu về AI, kết quả của cuộc tranh luận về AI Mạnh vẫn ít ảnh hưởng trực tiếp.



## Tham vọng trong giới hạn

Ở dạng thức yếu, AI có thể giúp chúng ta giải thích các cơ chế liên quan đến hành vi của con người và động vật.



Lập trường mạnh mẽ này có thể trái ngược với mục tiêu chế tạo máy móc thông minh, vốn là một mục tiêu phổ thông và thận trọng hơn, đồng thời cũng là cách tiếp cận đã có chỗ đứng vững chắc, được chứng minh thông qua các dự án kỹ thuật thành công.

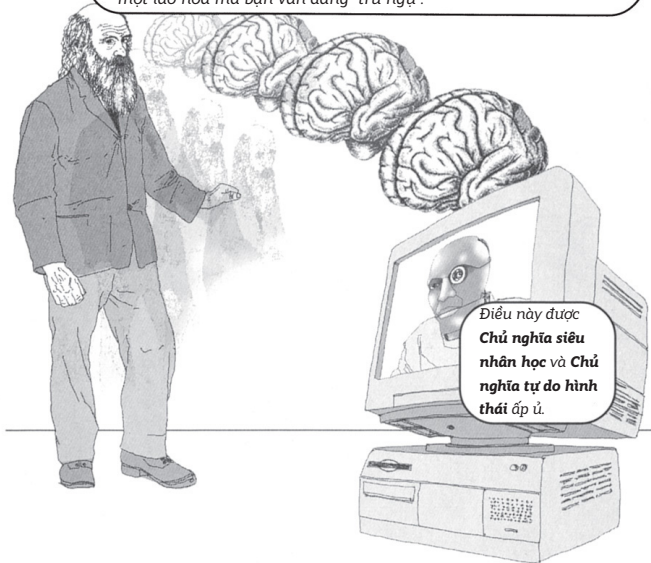
# Đưa AI tới giới hạn

## Sự bất tử và Chủ nghĩa siêu nhân học

*“Chúng ta không thể giữ chân AI, cũng như người nguyên thủy đã chẳng thể kiểm tỏa sức lan của ngôn ngữ.” – Doug Lenat và Edward Feigenbaum*

Nếu chúng ta giả định AI Mạnh là một khả năng có thật thì không ít câu hỏi cơ bản sẽ nảy sinh.

Hãy tưởng tượng bạn có thể rời bỏ cơ thể và đưa “linh hồn” mình vào một cỗ máy có “triển vọng” tốt hơn so với cơ thể sinh học ngày một lão hóa mà bạn vẫn đang “trú ngụ”.



Điều này được  
**Chủ nghĩa siêu  
nhân học** và **Chủ  
nghĩa tự do hình  
thái** áp ủ.

Vấn đề mà AI Mạnh nhắm tới cần làm sáng tỏ được khả năng này. Giả thuyết của AI Mạnh là: suy nghĩ, cũng như những đặc điểm về mặt trí óc khác, không liên kết chặt chẽ với cơ thể của chúng ta. Điều này có thể làm tăng khả năng tạo ra sự bất tử, khi “linh hồn” của một người có thể tồn tại trên một nền tảng khác mạnh mẽ hơn.