 Bản dịch cuốn sách "Artificial Intelligence For Dummies" sang tiếng Việt

# Giới thiệu

Bạn hầu như không thể tránh khỏi việc nghe nói về trí tuệ nhân tạo (AI) ngày nay. Bạn thấy AI trong phim ảnh, sách vở, các bản tin, và trên mạng. AI là một phần của robot, ô tô tự lái, drone, hệ thống y tế, trang web mua sắm trực tuyến và nhiều công nghệ khác ảnh hưởng đến cuộc sống hàng ngày của bạn. Một số người thậm chí đã tin tưởng AI đến mức ngủ thiếp đi trong khi ô tô tự lái đưa họ đến đích — điều này, dĩ nhiên, là bất hợp pháp.

Nhiều chuyên gia đang lấp đầy bạn với thông tin (và cả thông tin sai lệch) về AI. Một số coi AI là dễ thương và thân thiện; trong khi những người khác lại xem nó như một nguy cơ đe dọa nhân loại. Vấn đề ở đây là, khi bạn bị quá tải thông tin, bạn sẽ khó phân biệt điều gì là thật và điều gì chỉ là sản phẩm của trí tưởng tượng quá đà. Vậy bạn có thể tin tưởng AI đến mức nào? Phần lớn những lời thổi phồng về AI xuất phát từ kỳ vọng không thực tế của các nhà khoa học, doanh nhân và những người kinh doanh.

Artificial Intelligence For Dummies, 2nd Edition là cuốn sách bạn cần nếu bạn cảm thấy rằng mình thực sự không biết gì về công nghệ được cho là một phần không thể thiếu trong cuộc sống của mình. Cuốn sách này giúp bạn nhận ra những ứng dụng thực tế và cần thiết của AI trong cuộc sống hàng ngày, từ việc sử dụng bộ điều chỉnh nhiệt thông minh trong nhà đến những ứng dụng y tế tiên tiến.

Ngoài ra, cuốn sách còn chia sẻ những sự thật về những gì AI không thể làm. AI sẽ không bao giờ có thể thực hiện một số hoạt động nhất định cho đến khi công nghệ phát triển vượt bậc trong tương lai. Điều này giúp xua tan những ảo tưởng khiến bạn bị mù mờ về AI.

Về cuốn sách này Cuốn sách bắt đầu bằng cách giúp bạn hiểu rõ AI là gì, những gì nó cần để hoạt động, và tại sao nó từng thất bại trong quá khứ. Bạn sẽ khám phá cách AI liên quan đến các lĩnh vực như phân tích dữ liệu, học máy, và học sâu. Cuốn sách cũng cung cấp các đường dẫn đến thông tin bên ngoài để bạn có cái nhìn toàn diện hơn.

Ngoài ra, bạn còn được hướng dẫn cách hiểu các khái niệm chính về AI thông qua các biểu tượng và ghi chú, giúp việc học tập trở nên dễ dàng hơn.

# Mục lục

**Trang tiêu đề**  
**Lời giới thiệu**

**Phần 1: Giới thiệu về AI**

**Chương 1: Giới thiệu về AI**

* Định nghĩa thuật ngữ AI
* Tìm hiểu lịch sử AI
* Xem xét các ứng dụng của AI
* Tránh sự cường điệu hóa và đánh giá quá cao AI
* Mối liên kết giữa AI và công nghệ máy tính

**Chương 2: Định nghĩa vai trò của dữ liệu**

* Tìm kiếm dữ liệu như một tài nguyên phổ quát trong thời đại này
* Sử dụng dữ liệu thành công
* Chăm sóc và làm sạch dữ liệu
* Cân nhắc năm sự thật sai lệch trong dữ liệu
* Xác định giới hạn của việc thu thập dữ liệu
* Xem xét các vấn đề bảo mật dữ liệu

**Chương 3: Cân nhắc việc sử dụng thuật toán**

* Hiểu vai trò của thuật toán
* Khám phá cỗ máy học hỏi

**Chương 4: Đi đầu trong việc phát triển phần cứng chuyên dụng**

* Dựa vào phần cứng tiêu chuẩn
* Sử dụng GPU
* Làm việc với bộ xử lý học sâu (DLP)
* Tạo môi trường xử lý chuyên dụng
* Tăng cường khả năng của phần cứng
* Thêm cảm biến chuyên dụng
* Xây dựng phương pháp tương tác với môi trường

**Phần 2: Cân nhắc các ứng dụng AI trong xã hội**

**Chương 5: Nhìn nhận các ứng dụng AI trong máy tính**

* Giới thiệu các loại ứng dụng phổ biến
* Cách AI làm cho ứng dụng trở nên thân thiện hơn
* Tự động sửa lỗi
* Đưa ra gợi ý
* Cân nhắc các lỗi dựa trên AI

**Chương 6: Tự động hóa các quy trình thông thường**

* Phát triển giải pháp cho sự nhàm chán
* Làm việc trong môi trường công nghiệp
* Tạo ra môi trường an toàn

**Chương 7: Sử dụng AI để đáp ứng nhu cầu y tế**

* Áp dụng hệ thống giám sát bệnh nhân di động
* Làm cho con người trở nên hiệu quả hơn
* Giải quyết một loạt khả năng thể chất
* Hoàn thiện phân tích theo những cách mới
* Dựa vào công nghệ từ xa
* Tạo ra các kỹ thuật phẫu thuật mới
* Thực hiện nhiệm vụ bằng cách tự động hóa
* Kết hợp robot và chuyên gia y tế

**Chương 8: Dựa vào AI để cải thiện tương tác giữa con người**

* Phát triển các cách giao tiếp mới
* Trao đổi ý tưởng
* Cải thiện khả năng cảm nhận của con người

**Phần 3: Làm việc với các ứng dụng AI dựa trên phần mềm**

**Chương 9: Thực hiện phân tích dữ liệu cho AI**

* Định nghĩa phân tích dữ liệu
* Định nghĩa học máy
* Xem xét cách học hỏi từ dữ liệu

**Chương 10: Ứng dụng học máy trong AI**

* Đi theo nhiều con đường khác nhau để học hỏi
* Khám phá sự thật trong xác suất
* Phát triển các cây phân loại

**Chương 11: Cải thiện AI với học sâu**

* Tạo hình mạng nơ-ron tương tự não bộ con người
* Mô phỏng bộ não học hỏi
* Giới thiệu về học sâu
* Nhận diện các cạnh và hình dạng từ hình ảnh
* Học cách bắt chước nghệ thuật và cuộc sống

**Phần 4: Làm việc với AI trong các ứng dụng phần cứng**

**Chương 12: Phát triển robot**

* Định nghĩa vai trò của robot
* Lắp ráp một robot cơ bản

**Chương 13: Bay cùng drone**

* Ghi nhận trạng thái hiện đại
* Định nghĩa các ứng dụng của drone

**Chương 14: Sử dụng xe tự hành AI**

* Tìm hiểu lịch sử ngắn gọn
* Hiểu tương lai của khả năng di chuyển
* Sử dụng xe tự hành
* Vượt qua sự không chắc chắn trong nhận thức

**Phần 5: Cân nhắc tương lai của AI**

**Chương 15: Hiểu các ứng dụng không khả thi**

* Sử dụng AI ở nơi nó không hoạt động
* Cân nhắc tác động của "mùa đông AI"
* Tạo ra giải pháp cho các vấn đề

**Chương 16: Nhìn thấy AI trong không gian**

* Quan sát vũ trụ
* Khai thác tài nguyên không gian
* Khám phá các địa điểm mới
* Xây dựng cấu trúc trong không gian

**Chương 17: Tham gia vào các nỗ lực của con người**

* Giữ con người nổi bật
* Sống và làm việc trong không gian
* Xây dựng thành phố trong môi trường khắc nghiệt
* Làm cho con người hiệu quả hơn
* Giải quyết các vấn đề ở quy mô hành tinh

**Phần 6: Các nội dung thú vị**

**Chương 18: Mười nhóm nghề mà AI không thể thay thế**

* Thực hiện tương tác con người
* Tạo ra những điều mới
* Đưa ra quyết định trực giác

**Chương 19: Mười đóng góp đáng kể của AI cho xã hội**

* Cân nhắc các tương tác đặc trưng của con người
* Phát triển các giải pháp công nghiệp
* Tạo ra môi trường công nghệ mới
* Làm việc với AI trong không gian

**Chương 20: Mười cách AI đã thất bại**

* Hiểu biết
* Khám phá
* Đồng cảm

**Phụ lục: Mục lục tra cứu**  
**Thông tin về tác giả**  
**Trang quảng cáo**  
**Kết nối với Dummies**  
**Thỏa thuận cấp phép người dùng cuối**

# Table of Contents

Title Page

Introduction

Part 1: Introducing AI

Chapter 1: Introducing AI

Defining the Term AI

Understanding the History of AI

Considering AI Uses

Avoiding AI Hype and Overestimation

Connecting AI to the Underlying Computer

Chapter 2: Defining the Role of Data

Finding Data Ubiquitous in This Age

Using Data Successfully

Manicuring the Data

Considering the Five Mistruths in Data

Defining the Limits of Data Acquisition

Considering Data Security Issues

Chapter 3: Considering the Use of Algorithms

Understanding the Role of Algorithms

Discovering the Learning Machine

Chapter 4: Pioneering Specialized Hardware

Relying on Standard Hardware

Using GPUs

Working with Deep Learning Processors (DLPs)

Creating a Specialized Processing Environment

Increasing Hardware Capabilities

Adding Specialized Sensors

Devising Methods to Interact with the Environment

Part 2: Considering the Uses of AI in Society

Chapter 5: Seeing AI Uses in Computer Applications

Introducing Common Application Types

Seeing How AI Makes Applications Friendlier

Performing Corrections Automatically

Making Suggestions

Considering AI-based Errors

Chapter 6: Automating Common Processes

Developing Solutions for Boredom

Working in Industrial Settings

Creating a Safe Environment

Chapter 7: Using AI to Address Medical Needs

Implementing Portable Patient Monitoring

Making Humans More Capable

Addressing a Range of Physical Abilities

Completing Analysis in New Ways

Relying on Telepresence

Devising New Surgical Techniques

Performing Tasks Using Automation

Combining Robots and Medical Professionals

Chapter 8: Relying on AI to Improve Human Interaction

Developing New Ways to Communicate

Exchanging Ideas

Embellishing Human Sensory Perception

Part 3: Working with Software-Based AI Applications

Chapter 9: Performing Data Analysis for AI

Defining Data Analysis

Defining Machine Learning

Considering How to Learn from Data

Chapter 10: Employing Machine Learning in AI

Taking Many Different Roads to Learning

Exploring the Truth in Probabilities

Growing Trees that Can Classify

Chapter 11: Improving AI with Deep Learning

Shaping Neural Networks Similar to the Human Brain

Mimicking the Learning Brain

Introducing Deep Learning

Detecting Edges and Shapes from Images

Learning to Imitate Art and Life

Part 4: Working with AI in Hardware Applications

Chapter 12: Developing Robots

Defining Robot Roles

Assembling a Basic Robot

Chapter 13: Flying with Drones

Acknowledging the State of the Art

Defining Uses for Drones

Chapter 14: Utilizing the AI-Driven Car

Getting a Short History

Understanding the Future of Mobility

Getting into a Self-Driving Car

Overcoming Uncertainty of Perceptions

Part 5: Considering the Future of AI

Chapter 15: Understanding the Nonstarter Application

Using AI Where It Won’t Work

Considering the Effects of AI Winters

Creating Solutions in Search of a Problem

Chapter 16: Seeing AI in Space

Observing the Universe

Performing Space Mining

Exploring New Places

Building Structures in Space

Chapter 17: Engaging in Human Endeavors

Keeping Human Beings Popular

Living and Working in Space

Creating Cities in Hostile Environments

Making Humans More Efficient

Fixing Problems on a Planetary Scale

Part 6: The Part of Tens

Chapter 18: Ten Occupational Categories that AI Can’t

Replace

Performing Human Interaction

Creating New Things

Making Intuitive Decisions

Chapter 19: Ten Substantial Contributions of AI to Society

Considering Human-Specific Interactions

Developing Industrial Solutions

Creating New Technology Environments

Working with AI in Space

Chapter 20: Ten Ways in Which AI Has Failed

Understanding

Discovering

Empathizing

Index

About the Authors

Advertisement Page

Connect with Dummies

End User License Agreement

# Chương 1: Giới thiệu về Trí tuệ Nhân tạo (AI)

**TRONG CHƯƠNG NÀY**

* Định nghĩa AI và lịch sử của nó.
* Sử dụng AI cho các nhiệm vụ thực tế.
* Nhận diện những cường điệu về AI.
* Kết nối AI với công nghệ máy tính.

Trí tuệ nhân tạo (AI): Những khởi đầu sai lầm

AI đã trải qua nhiều **khởi đầu và tạm dừng** trong những năm qua, một phần vì:

* **Con người không thực sự hiểu:** AI là gì hoặc nó cần đạt được điều gì.
* **Ảnh hưởng của phim ảnh, truyền hình, và sách:** Những phương tiện này đã tạo ra hy vọng sai lệch về những gì AI có thể làm.

**Vấn đề chính:**  
Xu hướng **nhân cách hóa (anthropomorphize)** công nghệ, khiến mọi người kỳ vọng rằng AI phải làm được nhiều hơn khả năng thực tế của nó.

Để bắt đầu hiểu đúng về AI, cần:

1. **Định nghĩa chính xác:** AI thực sự là gì và không phải là gì.
2. **Hiểu mối liên hệ:** Cách AI kết nối với máy tính ngày nay.

## Định nghĩa thuật ngữ AI

Trước khi có thể sử dụng một thuật ngữ một cách có ý nghĩa và hữu ích, bạn phải có định nghĩa cho nó. Xét cho cùng, nếu không ai đồng ý về ý nghĩa của thuật ngữ, nó không có ý nghĩa gì cả; nó chỉ là một tập hợp các ký tự. Việc định nghĩa một thành ngữ (một thuật ngữ mà ý nghĩa của nó không rõ ràng từ ý nghĩa của các yếu tố cấu thành) đặc biệt quan trọng với các thuật ngữ kỹ thuật đã được đề cập nhiều lần trên các phương tiện truyền thông khác nhau.

**Nhận diện trí thông minh**

Mọi người định nghĩa trí thông minh theo nhiều cách khác nhau. Tuy nhiên, có thể nói rằng trí thông minh bao gồm một số hoạt động tinh thần nhất định, được cấu thành bởi các hoạt động sau:

* **Học hỏi:** Có khả năng thu thập và xử lý thông tin mới.
* **Lý luận:** Có khả năng thao tác thông tin theo nhiều cách khác nhau.
* **Hiểu biết:** Xem xét kết quả của việc thao tác thông tin.
* **Nắm bắt sự thật:** Xác định tính hợp lệ của thông tin đã được thao tác.
* **Nhận biết mối quan hệ:** Khám phá cách dữ liệu đã được xác thực tương tác với các dữ liệu khác.
* **Xem xét ý nghĩa:** Áp dụng sự thật vào các tình huống cụ thể theo cách phù hợp với mối quan hệ của chúng.
* **Phân biệt thực tế và niềm tin:** Xác định liệu dữ liệu có được hỗ trợ đầy đủ bởi các nguồn có thể chứng minh và được xác nhận một cách nhất quán hay không.

Danh sách này có thể mở rộng thêm rất nhiều, nhưng ngay cả danh sách này cũng dễ dàng bị giải thích khác nhau bởi bất kỳ ai chấp nhận nó là khả thi. Tuy nhiên, như bạn có thể thấy từ danh sách này, trí thông minh thường tuân theo một quy trình mà hệ thống máy tính có thể bắt chước như một phần của mô phỏng:

1. Xác định mục tiêu dựa trên nhu cầu hoặc mong muốn.
2. Đánh giá giá trị của bất kỳ thông tin nào đã biết để hỗ trợ mục tiêu.
3. Thu thập thêm thông tin có thể hỗ trợ mục tiêu. Trọng tâm ở đây là thông tin có thể hỗ trợ mục tiêu, thay vì thông tin mà bạn biết chắc chắn sẽ hỗ trợ mục tiêu.
4. Thao tác dữ liệu để nó đạt được một dạng nhất quán với thông tin hiện có.
5. Xác định các mối quan hệ và giá trị sự thật giữa thông tin hiện có và thông tin mới.
6. Xác định xem mục tiêu đã đạt được hay chưa.
7. Sửa đổi mục tiêu dựa trên dữ liệu mới và ảnh hưởng của nó đến xác suất thành công.
8. Lặp lại các bước từ 2 đến 7 khi cần thiết cho đến khi mục tiêu đạt được (được xác định là đúng) hoặc các khả năng để đạt được nó bị cạn kiệt (được xác định là sai).

### Khám phá bốn cách để định nghĩa AI

Như đã mô tả trong phần trước, khái niệm đầu tiên cần hiểu là AI không thực sự liên quan đến trí thông minh của con người. Đúng là một số AI được mô phỏng để bắt chước trí thông minh của con người, nhưng đó chỉ là mô phỏng. Khi nghĩ về AI, hãy chú ý đến sự tương tác giữa việc theo đuổi mục tiêu, xử lý dữ liệu để đạt được mục tiêu đó, và thu thập dữ liệu để hiểu rõ hơn về mục tiêu. AI dựa vào các thuật toán để đạt được kết quả, và kết quả đó có thể hoặc không liên quan đến mục tiêu hoặc phương pháp của con người. Với quan điểm này, AI có thể được phân loại theo bốn cách:

**1. Hành động như con người**

Khi một máy tính hành động giống con người, nó phản ánh rõ nhất **Bài kiểm tra Turing (Turing Test)**, trong đó máy tính thành công nếu không thể phân biệt được giữa máy tính và con người.

* **Bài kiểm tra Turing:** Để vượt qua bài kiểm tra này, AI phải sở hữu bốn công nghệ: xử lý ngôn ngữ tự nhiên, biểu diễn tri thức, lập luận tự động, và học máy.
* Một ví dụ tiêu biểu là chatbot **Mitsuku**, từng giành giải thưởng Loebner năm lần cho trí tuệ nhân tạo giống con người nhất. Bạn có thể tìm thấy Mitsuku tại:  
  <https://chat.kuki.ai/>  
  <http://www.square-bear.co.uk/mitsuku/home.htm>

**Các kỹ thuật hiện đại:** Thay vì hoàn toàn bắt chước con người, AI tập trung vào việc đạt được mục tiêu. Ví dụ, anh em nhà Wright không chế tạo máy bay bằng cách sao chép chính xác cách chim bay. Thay vào đó, họ lấy cảm hứng từ chim để phát triển khí động học, dẫn đến khả năng bay của con người.

**2. Suy nghĩ như con người**

Khi một máy tính suy nghĩ giống con người, nó thực hiện các nhiệm vụ đòi hỏi trí thông minh (trái ngược với các thủ tục lặp đi lặp lại) mà con người sử dụng để thành công, chẳng hạn như lái xe ô tô.

**Để xác định chương trình suy nghĩ giống con người:**  
Cần có phương pháp để xác định cách con người suy nghĩ, được định nghĩa bởi mô hình nhận thức (cognitive modeling). Mô hình này dựa trên ba kỹ thuật:

1. **Tự quan sát (Introspection):** Theo dõi và ghi lại các kỹ thuật được sử dụng để đạt được mục tiêu thông qua quá trình tư duy của chính mình.
2. **Kiểm tra tâm lý (Psychological Testing):** Quan sát hành vi của một người và thêm dữ liệu này vào cơ sở dữ liệu chứa hành vi của những người khác trong hoàn cảnh tương tự.
3. **Hình ảnh não bộ (Brain Imaging):** Theo dõi hoạt động não bộ thông qua các phương pháp cơ học như chụp cắt lớp vi tính (CAT), phát xạ positron (PET), chụp cộng hưởng từ (MRI), và từ trường não (MEG).

Sau khi tạo mô hình, bạn có thể viết một chương trình mô phỏng mô hình đó. Tuy nhiên, do sự biến đổi lớn trong quá trình tư duy của con người và khó khăn trong việc mô phỏng chính xác, kết quả chỉ mang tính thử nghiệm.

**3. Suy nghĩ hợp lý**

Khi một máy tính suy nghĩ hợp lý, nó dựa vào các tiêu chuẩn để tạo ra hướng dẫn mô tả hành vi con người điển hình trong những giới hạn nhất định.

* **Mục tiêu:** Giải quyết các vấn đề một cách logic khi có thể.
* **Phương pháp:** Bắt đầu bằng cách tạo một kỹ thuật cơ bản để giải quyết vấn đề, sau đó điều chỉnh để phù hợp với tình huống thực tế.

**4. Hành động hợp lý**

Khi một máy tính hành động hợp lý, nó dựa vào các hành động đã ghi nhận để tương tác với môi trường dựa trên điều kiện hiện tại, các yếu tố môi trường và dữ liệu hiện có.

* **Điểm tương đồng với suy nghĩ hợp lý:** Hành động hợp lý dựa trên giải pháp trong nguyên tắc, nhưng có thể không hữu ích trong thực tế.
* **Lợi ích:** Cung cấp một cơ sở để máy tính bắt đầu thương lượng và hoàn thành mục tiêu một cách hiệu quả và hợp lý.

### Quy trình con người và quy trình hợp lý

Quy trình của con người khác biệt với quy trình hợp lý ở kết quả của chúng. Một quy trình được coi là hợp lý nếu nó luôn làm điều đúng dựa trên thông tin hiện tại, với một thước đo hiệu suất lý tưởng. Nói ngắn gọn, quy trình hợp lý tuân thủ nghiêm ngặt các quy tắc, và giả định rằng các quy tắc này thực sự đúng. Trong khi đó, quy trình của con người bao gồm bản năng, trực giác, và các biến số khác không nhất thiết phải tuân thủ các quy tắc và có thể thậm chí không xem xét dữ liệu hiện có.

Ví dụ, cách hợp lý để lái xe là luôn tuân thủ luật giao thông. Tuy nhiên, giao thông không phải lúc nào cũng hợp lý. Nếu bạn tuân thủ luật một cách chính xác, bạn có thể bị mắc kẹt ở đâu đó vì những người lái xe khác không làm như vậy. Để thành công, một chiếc xe tự hành (SD car) phải hành động giống con người hơn là hành động hợp lý.

**Phân loại AI**

Các phân loại được sử dụng để định nghĩa AI cung cấp cách để xem xét các ứng dụng hoặc cách áp dụng AI. Một số hệ thống phân loại AI theo kiểu loại khá tùy tiện và không rõ ràng. Ví dụ, một số nhóm coi AI là **mạnh** (trí tuệ tổng quát có thể thích ứng với nhiều tình huống) hoặc **yếu** (trí tuệ cụ thể được thiết kế để thực hiện tốt một nhiệm vụ cụ thể).

**Vấn đề với AI mạnh:** Nó không thực hiện tốt bất kỳ nhiệm vụ nào.  
**Vấn đề với AI yếu:** Nó quá cụ thể để thực hiện các nhiệm vụ một cách độc lập.

Tuy nhiên, chỉ hai loại phân loại này là không đủ, ngay cả ở mức tổng quát. Bốn loại phân loại được đề xuất bởi **Arend Hintze** cung cấp cơ sở tốt hơn để hiểu AI:

1. **Máy phản ứng (Reactive machines):**  
   Các máy bạn thấy đánh bại con người trong cờ vua hoặc tham gia các chương trình trò chơi là ví dụ về máy phản ứng. Loại máy này không có trí nhớ hay kinh nghiệm để dựa vào khi ra quyết định. Thay vào đó, nó dựa hoàn toàn vào sức mạnh tính toán và các thuật toán thông minh để tái tạo từng quyết định mỗi lần. Đây là ví dụ của AI yếu được sử dụng cho một mục đích cụ thể.  
   **Chi tiết:** Phần “Considering the Chinese Room Argument” ở Chương 5 giải thích thêm về ý nghĩa của AI yếu.
2. **Bộ nhớ hạn chế (Limited memory):**  
   Một chiếc xe tự hành hoặc robot tự động không thể dành thời gian để đưa ra từng quyết định từ đầu. Các máy này dựa vào một lượng nhỏ bộ nhớ để cung cấp kiến thức từ kinh nghiệm về các tình huống khác nhau. Khi gặp tình huống tương tự, máy có thể dựa vào kinh nghiệm để giảm thời gian phản ứng và dành thêm tài nguyên cho các quyết định mới chưa từng gặp. Đây là ví dụ của cấp độ AI mạnh hiện tại.
3. **Lý thuyết về tâm trí (Theory of mind):**  
   Một máy có thể đánh giá cả mục tiêu cần thiết của mình và mục tiêu tiềm năng của các thực thể khác trong cùng môi trường. Dù hiện nay khả năng này khả thi ở một mức độ nhất định, nhưng chưa có dạng thương mại. Để xe tự hành trở nên thực sự tự động, mức độ AI này cần được phát triển hoàn toàn.
   * Ví dụ, một chiếc xe tự hành cần biết rằng nó phải đi từ điểm A đến điểm B, nhưng cũng phải dự đoán các mục tiêu xung đột tiềm tàng của những người lái xe xung quanh và phản ứng phù hợp.
   * Một ví dụ đơn giản khác là **bóng đá robot**, tham khảo tại:  
     <http://www.cs.cmu.edu/~robosoccer/main/>  
     <https://www.robocup.org/>
4. **Tự nhận thức (Self-awareness):**  
   Đây là loại AI mà bạn thường thấy trong các bộ phim. Tuy nhiên, nó đòi hỏi những công nghệ hiện tại chưa thể đạt được, vì loại máy này cần có ý thức về bản thân và tự nhận thức. Ngoài ra, thay vì chỉ dự đoán mục tiêu của người khác dựa trên môi trường và phản ứng, loại máy này có thể suy ra ý định của người khác dựa trên kinh nghiệm.

## Hiểu về Lịch sử của AI

Các phần trước của chương này giúp bạn hiểu trí tuệ từ góc nhìn của con người và thấy rằng các máy tính hiện đại còn kém xa khả năng mô phỏng trí thông minh đó, chứ chưa nói đến việc thực sự trở nên thông minh. Tuy nhiên, khát vọng tạo ra các cỗ máy thông minh (hoặc trong thời cổ đại là các tượng thần) đã có từ lâu đời. Mong muốn không còn cô đơn trong vũ trụ, được giao tiếp mà không phải đối mặt với sự bất nhất của con người, là một động lực mạnh mẽ. Tất nhiên, một cuốn sách không thể trình bày toàn bộ lịch sử loài người, vì vậy các phần sau cung cấp một cái nhìn tổng quan ngắn gọn, liên quan đến lịch sử của các nỗ lực AI hiện đại.

**Bắt đầu với logic biểu tượng tại Dartmouth**

Những máy tính đầu tiên chỉ là thiết bị tính toán. Chúng bắt chước khả năng của con người trong việc thao tác các ký hiệu để thực hiện các phép toán cơ bản, chẳng hạn như phép cộng. Sau đó, khả năng suy luận logic được bổ sung, cho phép thực hiện các phép lập luận toán học thông qua so sánh (ví dụ: xác định xem một giá trị có lớn hơn giá trị khác hay không). Tuy nhiên, con người vẫn cần phải:

1. Định nghĩa thuật toán được sử dụng để thực hiện tính toán.
2. Cung cấp dữ liệu cần thiết theo định dạng phù hợp.
3. Diễn giải kết quả.

Vào mùa hè năm 1956, nhiều nhà khoa học đã tham gia một hội thảo tại khuôn viên **Dartmouth College** để thực hiện một bước tiến xa hơn. Họ dự đoán rằng những cỗ máy có thể suy luận hiệu quả như con người sẽ được tạo ra chỉ trong một thế hệ. Nhưng họ đã sai. Chỉ đến ngày nay, chúng ta mới đạt được những cỗ máy có thể thực hiện suy luận toán học và logic hiệu quả như con người, và ngay cả điều này cũng chỉ là một phần nhỏ của trí tuệ con người.

**Những vấn đề tại hội thảo Dartmouth**

Vấn đề được nêu ra tại hội thảo Dartmouth và các nỗ lực cùng thời kỳ liên quan đến phần cứng — khả năng xử lý để thực hiện các tính toán đủ nhanh để tạo ra mô phỏng. Tuy nhiên, đây không phải là toàn bộ vấn đề.

* **Phần cứng:** Đúng là phần cứng đóng vai trò quan trọng, nhưng bạn không thể mô phỏng các quy trình mà bạn chưa hiểu.
* **Sự thiếu hiểu biết:** Vào thời điểm đó, chúng ta chưa hiểu đủ về cách con người suy nghĩ và xử lý thông tin để mô phỏng một cách hiệu quả.

Tuy nhiên, lý do AI ngày nay hiệu quả hơn là vì phần cứng cuối cùng đã đủ mạnh để hỗ trợ số lượng tính toán cần thiết.

### Các hệ thống chuyên gia

Các **hệ thống chuyên gia** (expert systems) lần đầu tiên xuất hiện vào những năm 1970 và tái xuất vào những năm 1980 như một nỗ lực nhằm giảm bớt yêu cầu tính toán mà AI đặt ra bằng cách sử dụng kiến thức của các chuyên gia. Một số biểu diễn của hệ thống chuyên gia đã được phát triển, bao gồm:

1. **Hệ thống dựa trên quy tắc (rule-based):**  
   Sử dụng các câu lệnh "nếu...thì" để đưa ra quyết định dựa trên các quy tắc kinh nghiệm.
2. **Hệ thống dựa trên khung (frame-based):**  
   Sử dụng các cơ sở dữ liệu được tổ chức thành các hệ thống phân cấp liên quan của thông tin chung gọi là "khung" (frames).
3. **Hệ thống dựa trên logic (logic-based):**  
   Dựa vào lý thuyết tập hợp (set theory) để thiết lập mối quan hệ giữa các yếu tố.

**Tầm quan trọng của hệ thống chuyên gia**

Sự xuất hiện của các hệ thống chuyên gia rất quan trọng vì chúng đại diện cho những triển khai **AI thực sự hữu ích và thành công đầu tiên**. Những hệ thống này không chỉ làm giảm yêu cầu về khả năng tính toán mà còn cho thấy tiềm năng thực tế của AI trong việc giải quyết các vấn đề phức tạp dựa trên kiến thức chuyên môn.

**Vấn đề với các hệ thống chuyên gia**

Một vấn đề lớn với các **hệ thống chuyên gia** là chúng có thể khó tạo ra và duy trì. Những người dùng ban đầu phải học các ngôn ngữ lập trình chuyên biệt như **List Processing (Lisp)** hoặc **Prolog**. Một số nhà cung cấp nhận thấy cơ hội đưa hệ thống chuyên gia đến với những lập trình viên ít kinh nghiệm hơn bằng cách sử dụng các sản phẩm như **VP-Expert** (tham khảo tại [The Illustrated VP-Expert trên Amazon.com](https://www.amazon.com)), dựa trên cách tiếp cận dựa trên quy tắc. Tuy nhiên, các sản phẩm này thường cung cấp chức năng rất hạn chế với cơ sở tri thức nhỏ.

**Sự biến mất của cụm từ "hệ thống chuyên gia"**

Vào những năm 1990, cụm từ **hệ thống chuyên gia** bắt đầu biến mất. Ý kiến cho rằng các hệ thống chuyên gia thất bại đã xuất hiện, nhưng thực tế là chúng thành công đến mức được tích hợp vào các ứng dụng mà chúng hỗ trợ.

Ví dụ, trước đây bạn cần mua ứng dụng kiểm tra ngữ pháp riêng biệt như **RightWriter** để sử dụng với trình xử lý văn bản. Tuy nhiên, ngày nay, các trình xử lý văn bản đã tích hợp công cụ kiểm tra ngữ pháp, vì chúng chứng minh được tính hữu ích của mình (dù không phải lúc nào cũng chính xác; tham khảo bài viết **“Hello, Mr. Chips PCS Learn English”** trên Washington Post để biết thêm chi tiết).

### Vượt qua các "mùa đông AI"

**Mùa đông AI** đề cập đến các giai đoạn mà nguồn tài trợ cho phát triển AI bị cắt giảm. Nhìn chung, AI đã đi theo một con đường mà các nhà ủng hộ thường thổi phồng khả năng của nó, thuyết phục những người không hiểu biết về công nghệ nhưng có nhiều tiền đầu tư. Khi AI không đáp ứng được kỳ vọng, một giai đoạn chỉ trích xảy ra, và cuối cùng là giảm nguồn tài trợ. Nhiều chu kỳ như vậy đã diễn ra trong lịch sử, và tất cả đều gây tổn hại cho sự tiến bộ thực sự của AI.

Hiện tại, AI đang trong một giai đoạn cường điệu mới nhờ công nghệ **học máy (machine learning)**, cho phép máy tính học từ dữ liệu. Thay vì dựa vào lập trình viên để thiết lập các hoạt động, học máy giúp máy tính rút ra các nhiệm vụ trực tiếp từ các ví dụ. Điều này giống như giáo dục một đứa trẻ bằng cách chỉ cho nó cách hành động thông qua ví dụ.

**Học máy và học sâu**

**Học máy** có một số rủi ro, bởi vì máy tính có thể học cách làm việc sai lầm thông qua sự giảng dạy bất cẩn. Hiện tại, **năm nhóm khoa học** đang làm việc trên các thuật toán học máy, mỗi nhóm tiếp cận từ một góc nhìn khác nhau (tham khảo thêm trong phần **“Avoiding AI Hype and Overestimation”** ở chương này để biết chi tiết).

Trong số đó, **học sâu (deep learning)** là công nghệ thành công nhất. Học sâu cố gắng bắt chước não người và khả thi nhờ:

* Máy tính mạnh mẽ.
* Các thuật toán thông minh hơn.
* Các tập dữ liệu lớn từ số hóa xã hội.
* Các khoản đầu tư lớn từ các công ty như Google, Facebook, Amazon và các doanh nghiệp khác tận dụng AI để phục vụ mục tiêu kinh doanh.

**AI và viễn cảnh tương lai**

Mọi người nói rằng mùa đông AI đã qua nhờ học sâu, và điều này đúng ở thời điểm hiện tại. Tuy nhiên, nếu quan sát cách mọi người nhìn nhận AI, có thể thấy rằng một giai đoạn chỉ trích khác sẽ xảy ra nếu các nhà ủng hộ không giảm bớt sự cường điệu. AI có thể làm được những điều tuyệt vời, nhưng chúng là những điều tuyệt vời ở cấp độ "thường nhật", chẳng hạn như thực hiện công việc lặp đi lặp lại để tìm ra vaccine Covid-19 (tham khảo **“How AI is being used for COVID-19 vaccine creation and distribution”** tại [TechRepublic.com](https://www.techrepublic.com)).

Phần tiếp theo sẽ mô tả cách AI hiện đang được sử dụng

## Xem xét các ứng dụng của AI

Ngày nay, bạn có thể tìm thấy AI được sử dụng trong rất nhiều ứng dụng. Vấn đề duy nhất là công nghệ này hoạt động quá hiệu quả đến mức bạn không nhận ra sự hiện diện của nó. Trên thực tế, bạn có thể sẽ ngạc nhiên khi biết rằng nhiều thiết bị gia đình hiện nay đã sử dụng AI.

Ví dụ:

* Một số bộ điều chỉnh nhiệt thông minh (smart thermostats) tự động tạo lịch trình dựa trên cách bạn điều chỉnh nhiệt độ thủ công.
* Các thiết bị sử dụng giọng nói để điều khiển học cách bạn nói để tương tác tốt hơn với bạn.

AI chắc chắn xuất hiện trong ô tô của bạn và đặc biệt là tại nơi làm việc. Thực tế, các ứng dụng AI có đến hàng triệu trường hợp — tất cả đều hoạt động một cách an toàn và thường không được chú ý ngay cả khi chúng khá đáng kể. Dưới đây là một số cách AI được sử dụng:

**1. Phát hiện gian lận**

Khi bạn nhận được cuộc gọi từ công ty thẻ tín dụng hỏi xem bạn có thực hiện một giao dịch cụ thể không, điều này không phải vì họ tò mò. Thay vào đó, AI được tích hợp trong mã của công ty thẻ tín dụng đã phát hiện một mẫu chi tiêu bất thường và thông báo cho bạn, vì có khả năng người khác đang sử dụng thẻ của bạn.

**2. Lập lịch tài nguyên (Resource scheduling)**

Nhiều tổ chức cần lên lịch sử dụng tài nguyên một cách hiệu quả. Ví dụ:

* Một bệnh viện có thể phải quyết định nơi đặt bệnh nhân dựa trên nhu cầu của bệnh nhân, sự sẵn có của chuyên gia lành nghề, và thời gian mà bác sĩ dự kiến bệnh nhân cần ở trong bệnh viện.

**3. Phân tích phức tạp (Complex analysis)**

Con người thường cần sự trợ giúp trong việc phân tích phức tạp vì có quá nhiều yếu tố cần xem xét.  
Ví dụ:

* Cùng một bộ triệu chứng có thể chỉ ra nhiều vấn đề khác nhau.
* Một bác sĩ hoặc chuyên gia khác có thể cần sự trợ giúp để đưa ra chẩn đoán kịp thời nhằm cứu sống bệnh nhân.

**4. Tự động hóa (Automation)**

Bất kỳ hình thức tự động hóa nào cũng có thể hưởng lợi từ việc bổ sung AI để xử lý các thay đổi hoặc sự kiện không mong đợi.

* Một vấn đề với một số loại tự động hóa hiện nay là khi xảy ra sự kiện bất ngờ, chẳng hạn như một vật thể ở sai vị trí, hệ thống tự động hóa có thể dừng hoạt động.
* Việc thêm AI vào hệ thống tự động hóa có thể giúp nó xử lý sự kiện không mong đợi và tiếp tục hoạt động như không có gì xảy ra.

**5. Dịch vụ khách hàng (Customer service)**

* Các đường dây hỗ trợ khách hàng mà bạn gọi ngày nay có thể không có con người đứng sau.
* Hệ thống tự động hóa đủ tốt để tuân theo các kịch bản và sử dụng các tài nguyên khác nhau để xử lý phần lớn câu hỏi của bạn.
* Với giọng nói ngữ điệu tự nhiên (cũng do AI cung cấp), bạn có thể thậm chí không nhận ra rằng mình đang nói chuyện với một máy tính.

**6. Hệ thống an toàn (Safety systems)**

Nhiều hệ thống an toàn trong các loại máy móc hiện nay dựa vào AI để kiểm soát phương tiện trong những tình huống khẩn cấp.  
Ví dụ:

* Nhiều hệ thống phanh tự động (ABS) sử dụng AI để dừng xe dựa trên tất cả các dữ liệu đầu vào mà xe có thể cung cấp, chẳng hạn như hướng trượt.
* Công nghệ **ABS** thực sự đã xuất hiện từ 40 năm trước.  
  Tham khảo thêm tại **“ABS (Anti-Lock Braking System) — A Brief History Of A 40-Year-Old Life-Saver”** trên [DriveSpark.com](https://www.drivespark.com).

**7. Hiệu quả máy móc (Machine efficiency)**

AI có thể kiểm soát máy móc để đạt được hiệu suất tối đa.

* AI quản lý việc sử dụng tài nguyên để hệ thống không vượt quá tốc độ hoặc các mục tiêu khác.
* Từng đơn vị năng lượng được sử dụng chính xác theo yêu cầu để cung cấp dịch vụ mong muốn.

## Tránh sự cường điệu và đánh giá quá cao về AI

Chương này đề cập nhiều đến việc cường điệu hóa AI. Thật không may, chương này chỉ mới chạm đến bề mặt của những lời cường điệu ngoài kia. Nếu bạn xem các bộ phim như **Her** hay **Ex Machina**, bạn có thể nghĩ rằng AI đã phát triển hơn thực tế rất nhiều. Vấn đề là AI thực sự đang ở giai đoạn sơ khai, và bất kỳ ứng dụng nào như trong phim đều là sản phẩm sáng tạo từ trí tưởng tượng phong phú. Các phần dưới đây sẽ giúp bạn hiểu cách sự cường điệu và đánh giá quá cao đang làm lệch lạc mục tiêu thực sự mà AI có thể đạt được hiện nay.

### Định nghĩa năm nhóm (tribes) và thuật toán tổng thể (master algorithm)

Bạn có thể đã nghe về khái niệm **điểm kỳ dị (singularity)**, một ý tưởng chịu trách nhiệm cho những tuyên bố tiềm năng được trình bày trên truyền thông và phim ảnh. Điểm kỳ dị về cơ bản là một **thuật toán tổng thể (master algorithm)** bao gồm tất cả năm nhóm học tập (tribes) được sử dụng trong học máy (machine learning). Để đạt được điều mà những nguồn này nói, máy móc phải có khả năng học tập như con người — như được chỉ ra qua **bảy loại trí thông minh** trong phần **“Discerning intelligence”** đã thảo luận trước đó trong chương này.

**Năm nhóm học tập:**

1. **Symbologists (Nhóm biểu tượng):**  
   Xuất phát từ logic và triết học, nhóm này dựa vào **suy luận nghịch (inverse deduction)** để giải quyết vấn đề.
2. **Connectionists (Nhóm kết nối):**  
   Xuất phát từ khoa học thần kinh, nhóm này sử dụng **lan truyền ngược (backpropagation)** để giải quyết vấn đề.
3. **Evolutionaries (Nhóm tiến hóa):**  
   Xuất phát từ sinh học tiến hóa, nhóm này dựa vào **lập trình di truyền (genetic programming)** để giải quyết vấn đề.
4. **Bayesians (Nhóm Bayes):**  
   Xuất phát từ thống kê, nhóm này sử dụng **suy luận xác suất (probabilistic inference)** để giải quyết vấn đề.
5. **Analogizers (Nhóm tương tự):**  
   Xuất phát từ tâm lý học, nhóm này dựa vào **máy kernel (kernel machines)** để giải quyết vấn đề.

**Mục tiêu cuối cùng của học máy**

Mục tiêu cuối cùng của học máy là kết hợp các công nghệ và chiến lược từ năm nhóm để tạo ra một **thuật toán tổng thể** có thể học bất cứ điều gì.

* Tuy nhiên, việc đạt được mục tiêu này vẫn còn rất xa.
* Các nhà khoa học như **Pedro Domingos** tại Đại học Washington hiện đang làm việc để hướng tới mục tiêu này.

Dù vậy, năm nhóm có thể không cung cấp đủ thông tin để giải quyết vấn đề về trí thông minh của con người. Vì vậy, ngay cả khi tạo ra các thuật toán tổng thể cho tất cả năm nhóm, chúng vẫn có thể không dẫn đến điểm kỳ dị. Điều này cho thấy rằng chúng ta vẫn còn chưa hiểu rõ về cách con người suy nghĩ hoặc tại sao họ suy nghĩ theo một cách nhất định. Bất kỳ tin đồn nào về việc AI sẽ thống trị thế giới hoặc vượt trội hơn con người đều **hoàn toàn sai sự thật**.

**Xem xét các nguồn gốc của sự cường điệu về AI**

Có rất nhiều nguồn gốc gây cường điệu về AI. Phần lớn đến từ truyền thông, được trình bày bởi những người không hiểu rõ về AI, có lẽ chỉ biết qua một cuốn tiểu thuyết khoa học viễn tưởng mà họ đã đọc.

* **Không chỉ phim ảnh và truyền hình gây ra vấn đề:** Các loại hình truyền thông khác cũng góp phần.
* **Ví dụ:** Báo cáo tin tức thường mô tả AI có thể làm được điều gì đó mà nó không thể làm, đơn giản vì phóng viên không hiểu công nghệ.
* Điều thú vị là nhiều dịch vụ tin tức hiện nay sử dụng AI để viết nháp các bài báo cho phóng viên.  
  **Tham khảo thêm tại:**  
  [“Did A Robot Write This? How AI Is Impacting Journalism”](https://www.forbes.com) trên **Forbes.com**.

**Thất bại của AI do cường điệu hóa**

Một số sản phẩm lẽ ra cần được kiểm tra kỹ lưỡng hơn trước khi tung ra thị trường.

* Bài viết **“2020 in Review: 10 AI Failures”** trên [SyncedReview.com](https://syncedreview.com) liệt kê mười sản phẩm được nhà phát triển thổi phồng nhưng thất bại hoàn toàn.

**Nguyên nhân thất bại:**

* Một số thất bại lớn và gây ảnh hưởng tiêu cực đến khả năng của AI nói chung.
* Một số thất bại do **người dùng can thiệp** vào thiết bị AI, dẫn đến lỗi không mong muốn.

Rõ ràng, quy trình kiểm tra cần xem xét khả năng người dùng cố tình gây rối với AI để gây ra lỗi. Nếu không, AI sẽ tiếp tục không đáp ứng kỳ vọng vì người dùng cố ý thao tác phần mềm để gây thất bại một cách hài hước.

### Hiểu về việc người dùng đánh giá quá cao AI

Do sự cường điệu (và đôi khi là sự lười biếng hoặc mệt mỏi), người dùng thường xuyên đánh giá quá cao khả năng của AI trong việc thực hiện các nhiệm vụ.

Ví dụ:

* Một chủ xe **Tesla** gần đây bị phát hiện đang ngủ trong xe khi xe di chuyển với tốc độ 90 mph (145 km/h) trên đường cao tốc.  
  **Tham khảo thêm tại:**  
  [“Tesla owner in Canada charged with ‘sleeping’ while driving over 90 mph”](https://www.theguardian.com/technology).

Mặc dù người dùng đã đánh giá quá cao khả năng lái xe của công nghệ này, dường như nó vẫn hoạt động đủ tốt (ít nhất là trong trường hợp của tài xế này) để tránh xảy ra sự cố nghiêm trọng.

**Ví dụ khác về đánh giá quá cao AI**

Bạn không cần phải lái xe với tốc độ 90 mph để gặp phải tình trạng đánh giá quá cao AI.

* **Robot hút bụi** cũng thường không đáp ứng được kỳ vọng, bởi vì người dùng nghĩ rằng chỉ cần cắm thiết bị vào là không bao giờ phải nghĩ đến việc hút bụi nữa.
* Thực tế, các bộ phim thường mô tả robot hút bụi hoạt động chính xác theo cách này.

Bài viết **“How to Solve the Most Annoying Robot Vacuum Cleaner Problems”** trên [RobotsInMyHome.com](https://www.robotsinmyhome.com) cung cấp các kỹ thuật khắc phục sự cố cho các loại robot hút bụi khác nhau, bởi vì robot vẫn cần sự can thiệp của con người.

**Điểm mấu chốt:**  
Hầu hết các robot vẫn cần sự can thiệp của con người ở một số thời điểm vì chúng thiếu kiến thức để hoạt động độc lập hoàn toàn.

## Kết nối AI với máy tính cơ bản

Để thấy AI hoạt động, bạn cần một số loại hệ thống máy tính, một ứng dụng chứa phần mềm cần thiết và một cơ sở tri thức.

1. **Hệ thống máy tính:**  
   Hệ thống máy tính có thể là bất kỳ thiết bị nào có chip bên trong; ví dụ:
   * Một **điện thoại thông minh** có thể hoạt động tốt như một máy tính để bàn trong một số ứng dụng.
   * Tuy nhiên, nếu bạn là **Amazon** và muốn đưa ra gợi ý mua sắm cá nhân hóa, một chiếc điện thoại thông minh là không đủ — bạn cần một hệ thống máy tính rất lớn để xử lý nhiệm vụ này.

**Kích thước của hệ thống máy tính tỷ lệ thuận với khối lượng công việc mà bạn kỳ vọng AI thực hiện.**

1. **Ứng dụng:**  
   Các ứng dụng AI có thể khác nhau về kích thước, độ phức tạp, và vị trí:
   * Ví dụ, nếu bạn là doanh nghiệp muốn phân tích dữ liệu khách hàng để tối ưu hóa cách tiếp cận bán hàng, bạn có thể sử dụng một **ứng dụng dựa trên máy chủ** để thực hiện nhiệm vụ này.
   * Nếu bạn là khách hàng và muốn tìm sản phẩm trên **Amazon** để mua kèm các sản phẩm hiện tại, ứng dụng đó có thể không nằm trên máy tính của bạn; bạn sẽ truy cập nó qua một **ứng dụng web** được lưu trữ trên máy chủ của Amazon.
2. **Cơ sở tri thức:**  
   Cơ sở tri thức có thể khác nhau về vị trí và kích thước:
   * **Dữ liệu càng phức tạp, bạn càng có thể khai thác được nhiều thông tin từ nó**, nhưng bạn cũng cần xử lý nhiều hơn.
   * **Không có "bữa trưa miễn phí" trong quản lý tri thức.**

**Vị trí và thời gian:**

* + Một kết nối mạng giúp bạn truy cập vào cơ sở tri thức lớn trực tuyến, nhưng sẽ mất thời gian do độ trễ của mạng.
  + Ngược lại, cơ sở dữ liệu cục bộ có tốc độ nhanh hơn nhưng thường thiếu chi tiết trong nhiều trường hợp.

# Chương 2: Định nghĩa vai trò của dữ liệu

**TRONG CHƯƠNG NÀY**

* Nhìn nhận dữ liệu như một tài nguyên phổ quát.
* Thu thập và xử lý dữ liệu.
* Tìm kiếm những sai lệch trong dữ liệu.
* Xác định giới hạn thu thập dữ liệu.
* Xem xét vấn đề bảo mật dữ liệu.

**Tổng quan về dữ liệu**

Dữ liệu không phải là khái niệm mới. Mọi ứng dụng thú vị từng được viết cho máy tính đều liên quan đến dữ liệu.

* Dữ liệu xuất hiện dưới nhiều hình thức khác nhau — một số được tổ chức, số khác thì không.
* Điều thay đổi là **khối lượng dữ liệu.** Ngày nay, chúng ta có thể truy cập lượng dữ liệu khổng lồ, chi tiết gần như mọi khía cạnh của cuộc sống, đôi khi đến mức mà ngay cả người liên quan cũng không nhận ra.

Ngoài ra, việc sử dụng phần cứng tiên tiến và cải tiến các thuật toán đã biến dữ liệu trở thành **nguồn tài nguyên phổ quát cho AI** ngày nay.

**Thu thập dữ liệu**

Để làm việc với dữ liệu, bạn phải thu thập nó.

* Ngày nay, các ứng dụng **thu thập dữ liệu thủ công** như trong quá khứ, và **tự động** bằng các phương pháp mới.
* Việc thu thập dữ liệu không chỉ sử dụng một hoặc hai kỹ thuật; nó trải dài từ **hoàn toàn thủ công đến hoàn toàn tự động**.

Hiện nay, sự chú trọng cũng nằm ở việc **thu thập dữ liệu một cách đạo đức** — chẳng hạn, không thu thập dữ liệu mà người dùng chưa đồng ý.

Chương này sẽ khám phá các vấn đề xoay quanh việc thu thập dữ liệu.

**Xử lý dữ liệu thô**

Dữ liệu thô thường không phù hợp để phân tích.

* Chương này giúp bạn hiểu **tầm quan trọng của việc xử lý và định hình dữ liệu** để đáp ứng các yêu cầu cụ thể.
* Bạn cũng sẽ khám phá nhu cầu xác định **giá trị sự thật (truth value)** của dữ liệu để đảm bảo rằng kết quả phân tích phù hợp với mục tiêu ban đầu của ứng dụng.

**Giới hạn trong việc thu thập dữ liệu**

Thật thú vị, bạn cũng phải đối mặt với các **giới hạn trong việc thu thập dữ liệu.**

* Hiện không có công nghệ nào cho phép lấy suy nghĩ từ tâm trí ai đó thông qua cách thức thần giao cách cảm.
* Tất nhiên, còn có những giới hạn khác — hầu hết có thể bạn đã biết nhưng có thể chưa cân nhắc đến.

**Bảo mật dữ liệu**

Việc thu thập dữ liệu theo cách không an toàn không mang lại hiệu quả.

* Dữ liệu phải không có thiên lệch, không bị hỏng, và đến từ một nguồn mà bạn biết rõ.

Chương này sẽ cung cấp thêm thông tin về **giới hạn thu thập** và **bảo mật dữ liệu.**

## Dữ liệu là một loại tài nguyên phổ quát

**Dữ liệu lớn (Big Data)** không chỉ là một thuật ngữ thời thượng mà các nhà cung cấp sử dụng để giới thiệu những cách mới lưu trữ và phân tích dữ liệu. Cuộc cách mạng dữ liệu lớn là một thực tế hàng ngày và là động lực chính của thời đại chúng ta.

Bạn có thể đã nghe về dữ liệu lớn trong nhiều ấn phẩm khoa học và kinh doanh chuyên ngành, thậm chí có thể tự hỏi thuật ngữ này thực sự có nghĩa là gì.

* **Về mặt kỹ thuật:** Dữ liệu lớn đề cập đến khối lượng dữ liệu máy tính lớn và phức tạp, đến mức mà các ứng dụng không thể xử lý dữ liệu chỉ bằng cách tăng dung lượng lưu trữ hoặc tăng sức mạnh tính toán.

**Ý nghĩa của Dữ liệu lớn**

**Dữ liệu lớn** ám chỉ một cuộc cách mạng trong việc lưu trữ và xử lý dữ liệu:

1. **Tăng khả năng lưu trữ:** Không chỉ lưu trữ nhiều hơn mà còn lưu trữ tốt hơn về chất lượng.
2. **Nâng cao khả năng thao tác:** Giúp thực hiện các tác vụ tốt hơn.

**Quan điểm của máy tính:**

* Máy tính lưu trữ dữ liệu lớn dưới các định dạng khác nhau (chẳng hạn như tệp cơ sở dữ liệu hoặc tệp .csv).
* Dù ở định dạng nào, dữ liệu với máy tính vẫn chỉ là một dòng các số 1 và 0 (ngôn ngữ cơ bản của máy tính).

**Quan điểm con người:**

* **Dữ liệu có cấu trúc (Structured Data):**
  + Có cấu trúc rõ ràng. Bạn biết chính xác dữ liệu chứa gì và vị trí của từng phần dữ liệu.
  + Ví dụ điển hình: Các bảng cơ sở dữ liệu, trong đó thông tin được sắp xếp thành các cột, mỗi cột chứa một loại thông tin cụ thể.
  + Dữ liệu có cấu trúc thường được tổ chức theo thiết kế: Bạn thu thập có chọn lọc và ghi lại nó ở đúng vị trí.
  + **Ví dụ:** Bạn có thể muốn đặt số lượng người mua một sản phẩm cụ thể trong một cột, trong một bảng, trong một cơ sở dữ liệu.

Tương tự như một thư viện: Nếu bạn biết chính xác thông tin mình cần, bạn có thể tìm thấy nó ngay lập tức.

* **Dữ liệu không có cấu trúc (Unstructured Data):**
  + Bao gồm hình ảnh, video và bản ghi âm.
  + **Ví dụ:** Một tệp văn bản không có định dạng sắp xếp rõ ràng nhưng được gắn thẻ với các đặc điểm như kích thước, ngày tháng hoặc loại nội dung.
  + Bạn thường không biết chính xác dữ liệu xuất hiện ở đâu trong tập dữ liệu không có cấu trúc. Thay vào đó, dữ liệu xuất hiện dưới dạng các chuỗi số 1 và 0 mà một ứng dụng phải diễn giải hoặc trực quan hóa.

Kho dữ liệu phong phú và tinh vi này không xuất hiện chỉ sau một đêm.  
Phải mất thời gian để phát triển công nghệ lưu trữ lượng dữ liệu khổng lồ này. Ngoài ra, cần có thời gian để phổ biến công nghệ tạo ra và phân phối dữ liệu, cụ thể là: **máy tính, cảm biến, điện thoại di động thông minh, và Internet với các dịch vụ trên World Wide Web.**

Các phần dưới đây sẽ giúp bạn hiểu điều gì khiến dữ liệu trở thành một tài nguyên phổ quát ngày nay.

### Hiểu ý nghĩa từ Moore's Law

Năm 1965, **Gordon Moore**, đồng sáng lập Intel và Fairchild Semiconductor, đã viết trong một bài báo có tựa đề **“Cramming More Components Onto Integrated Circuits”** trên IEEE.org rằng:

* Số lượng thành phần trong các mạch tích hợp (Integrated Circuits - IC) sẽ tăng gấp đôi mỗi năm trong thập kỷ tiếp theo.
* **Thời điểm đó, transistor chiếm ưu thế trong ngành điện tử.** Việc có thể tích hợp nhiều transistor hơn vào một mạch tích hợp đồng nghĩa với khả năng tạo ra các thiết bị điện tử mạnh mẽ và hữu ích hơn.
* Quá trình này được gọi là **tích hợp (integration)** và ám chỉ một quy trình **thu nhỏ hóa điện tử (electronics miniaturization)** — làm cho mạch điện nhỏ hơn nhiều nhưng vẫn giữ nguyên chức năng.

**Ví dụ:**

* Máy tính ngày nay không nhỏ hơn quá nhiều so với máy tính của một thập kỷ trước, nhưng chúng mạnh mẽ hơn đáng kể.
* Điều này cũng đúng với điện thoại di động: Dù có cùng kích thước (thậm chí nhỏ hơn), chúng có thể thực hiện được nhiều tác vụ hơn.

**Moore's Law và tác động lâu dài**

Những gì **Moore** tuyên bố trong bài báo đã đúng trong nhiều năm và trở thành quy luật được ngành công nghiệp bán dẫn gọi là **Moore's Law**.  
**Tham khảo thêm tại:**  
<http://www.mooreslaw.org/>.

* **10 năm đầu tiên:** Quy tắc tăng gấp đôi xảy ra đúng như dự đoán.
* **Năm 1975:** Moore đã sửa đổi dự đoán của mình, dự báo rằng số lượng thành phần sẽ tăng gấp đôi sau mỗi hai năm thay vì mỗi năm.

**Hình minh họa tác động của quy luật Moore**

A graph showing the number of transistors

Description automatically generated

Hình 2-1 minh họa tác động của sự tăng gấp đôi này qua thời gian.

**Tương lai của quy luật Moore**

Trong tương lai, **quy luật Moore** có thể không còn áp dụng được vì ngành công nghiệp sẽ chuyển sang các công nghệ mới, chẳng hạn như:

* Sử dụng **laser quang học (optical lasers)** để chế tạo các thành phần thay vì transistor.  
  **Tham khảo:** [“Optical Computing: Solving Problems at the Speed of Light”](https://www.bbvaopenmind.com) trên **BBVAOpenMind.com**.
* Dựa vào **máy tính lượng tử (quantum computers)**, như những máy tính đang được Google và IBM phát triển.  
  **Tham khảo:** [“Google confirms ‘quantum supremacy’ breakthrough”](https://www.theverge.com) trên **The Verge.com**.

### Ảnh hưởng của quy luật Moore đến dữ liệu

Kể từ năm 1965, việc số lượng thành phần tăng gấp đôi mỗi hai năm đã mang lại những tiến bộ lớn trong điện tử số, ảnh hưởng sâu rộng đến:

* **Thu thập, lưu trữ, thao tác, và quản lý dữ liệu.**

**Cách quy luật Moore tác động trực tiếp đến dữ liệu:**

1. **Thiết bị thông minh hơn:**  
   Thiết bị càng thông minh, sự phổ biến càng lớn.
   * Khi sự phổ biến tăng, giá giảm, tạo ra một vòng lặp vô tận thúc đẩy việc sử dụng máy tính mạnh mẽ và cảm biến nhỏ ở khắp nơi.
2. **Khả năng lưu trữ dữ liệu lớn hơn:**
   * Bộ nhớ máy tính và đĩa lưu trữ dữ liệu lớn hơn cho phép mở rộng khả năng sử dụng dữ liệu, chẳng hạn như:
     + Website
     + Ghi chép giao dịch
     + Đo lường
     + Hình ảnh kỹ thuật số
     + Và nhiều dạng dữ liệu khác.

**Sử dụng dữ liệu ở mọi nơi**

**Các nhà khoa học** cần máy tính mạnh mẽ hơn so với người dùng thông thường vì các thí nghiệm khoa học.

* Họ đã xử lý lượng dữ liệu ấn tượng từ nhiều năm trước, trước khi thuật ngữ **dữ liệu lớn** được ra đời.
* Khi đó, Internet chưa tạo ra lượng dữ liệu khổng lồ như ngày nay.

**Dữ liệu lớn không phải là một trào lưu nhất thời;** nó có nguồn gốc từ nhiều lĩnh vực khoa học, chẳng hạn như:

* **Thiên văn học (Astronomy):** Các nhiệm vụ không gian.
* **Vệ tinh (Satellite):** Giám sát và theo dõi.
* **Khí tượng học (Meteorology):** Dự báo thời tiết.
* **Vật lý (Physics):** Máy gia tốc hạt.
* **Di truyền học (Genomics):** Phân tích trình tự DNA.

**AI và dữ liệu khoa học**

Một ứng dụng AI có thể chuyên sâu vào lĩnh vực khoa học cụ thể, ví dụ:

* **Watson của IBM:** Có khả năng chẩn đoán y khoa ấn tượng nhờ học thông tin từ hàng triệu tài liệu khoa học về bệnh lý và y học.

Tuy nhiên, **AI thực tế** thường tập trung vào những khía cạnh "bình thường" hơn, chẳng hạn:

* Nhận diện vật thể.
* Di chuyển theo lộ trình.
* Hiểu và giao tiếp với con người.

**Dữ liệu từ Internet**

Ngày nay, Internet tạo ra và phân phối một lượng dữ liệu lớn.

* Sản lượng dữ liệu hàng ngày hiện tại ước tính vào khoảng **2,5 quintillion byte (2,5 tỷ tỷ byte)**.
* Phần lớn dữ liệu này thuộc dạng không cấu trúc, chẳng hạn như video và âm thanh.

Dữ liệu này liên quan đến:

* **Hoạt động, cảm xúc, trải nghiệm, và mối quan hệ** của con người.
* Thông qua dữ liệu này, AI có thể dễ dàng học cách suy luận và hành động giống con người hơn.

**Ví dụ về dữ liệu thú vị**

1. **Hình ảnh và biểu cảm khuôn mặt:**
   * Các kho dữ liệu lớn từ hình ảnh và video trên mạng xã hội như **Facebook**, **YouTube**, và **Google** cung cấp thông tin về giới tính, tuổi, cảm xúc, thậm chí là sở thích tình dục, khuynh hướng chính trị hoặc IQ.
   * **Tham khảo:** [“Face-reading AI will be able to detect your politics and IQ, professor says”](https://www.theguardian.com) trên **The Guardian.com**.
2. **Thông tin y tế và dữ liệu sinh trắc học:**
   * Đồng hồ thông minh đo dữ liệu cơ thể như nhiệt độ và nhịp tim trong cả bệnh và trạng thái khỏe mạnh.
   * Dữ liệu từ đồng hồ thông minh được xem là phương pháp phát hiện sớm các bệnh nghiêm trọng, chẳng hạn như Covid-19.
   * **Tham khảo:** [“Using Smartwatch Data to Detect COVID-19 Cases Early”](https://jamanetwork.com) trên **JAMA Network.com**.
3. **Dữ liệu về quan hệ xã hội:**
   * Nguồn dữ liệu từ mạng xã hội và công cụ tìm kiếm tiết lộ cách mọi người tương tác và điều gì thúc đẩy sự quan tâm của họ.
   * Một nghiên cứu từ Trung tâm Tâm lý học Cambridge cho biết:
     + Tương tác trên Facebook chứa đựng nhiều dữ liệu về các mối quan hệ thân mật.
   * **Tham khảo:** [“Your computer knows you better than your friends do, say researchers”](https://www.theguardian.com) trên **The Guardian.com**.
4. **Thông tin về cách chúng ta nói:**
   * Điện thoại di động ghi lại cách chúng ta nói chuyện, ví dụ:
     + Tính năng **OK Google** trên điện thoại Android thường xuyên ghi lại các câu hỏi và đôi khi nhiều hơn thế.
   * **Tham khảo:** [“Google’s been quietly recording your voice; here’s how to listen to—and delete—the archive”](https://www.qz.com) trên **Quartz.com**.

Hàng ngày, người dùng kết nối thêm nhiều thiết bị hơn với Internet

Ngày càng có nhiều **dữ liệu cá nhân mới** được lưu trữ từ các thiết bị kết nối Internet.  
Ví dụ:

* Các **trợ lý ảo** như **Amazon Echo** và các thiết bị nhà thông minh tích hợp khác đang điều chỉnh và hỗ trợ quản lý môi trường trong nhà.
* Đây chỉ là phần nổi của tảng băng, vì nhiều công cụ thông dụng khác trong đời sống hàng ngày (từ tủ lạnh đến bàn chải đánh răng) cũng đang trở nên kết nối và có khả năng xử lý, ghi lại, và truyền tải dữ liệu.

**Internet of Things (IoT)** đang trở thành hiện thực:

* **Năm 2015:** Có 3,6 tỷ thiết bị IoT được lắp đặt.
* **Dự báo năm 2025:** Số lượng này sẽ đạt **30,9 tỷ thiết bị**, tăng **858%**.

**Tham khảo thêm:**

* [“How Many IoT Devices Are There in 2021?”](https://techjury.net) tại **TechJury.net**.
* [“Internet of Things (IoT) and non-IoT active device connections worldwide from 2010 to 2025 (in billions)”](https://statista.com) tại **Statista.com**.

### Đưa thuật toán vào thực tế

Nhân loại hiện đang ở một giao lộ đáng kinh ngạc giữa:

1. **Lượng dữ liệu chưa từng có** được tạo ra bởi các phần cứng ngày càng nhỏ hơn và mạnh mẽ hơn.
2. **Khả năng xử lý và phân tích dữ liệu** bởi chính các máy tính đã giúp phát triển và lan truyền quá trình này.

Dữ liệu đã trở nên phổ biến đến mức **giá trị của nó không chỉ nằm trong thông tin mà nó chứa**, mà còn ở cách nó được sử dụng để tạo ra giá trị mới.

* Một số người gọi dữ liệu là **“dầu mỏ mới” (new oil)**.

**Giá trị mới của dữ liệu nằm ở đâu?**

* Cách các ứng dụng xử lý, lưu trữ và truy xuất dữ liệu.
* Cách dữ liệu thực sự được sử dụng thông qua các **thuật toán thông minh**.

**Thuật toán và AI đã thay đổi cách sử dụng dữ liệu**

Như đã đề cập trong chương trước, các thuật toán AI đã thử nghiệm nhiều cách tiếp cận khác nhau theo trình tự sau:

1. **Thuật toán đơn giản.**
2. **Lý luận biểu tượng dựa trên logic.**
3. **Hệ thống chuyên gia.**

**Những năm gần đây:** Thuật toán AI đã chuyển sang sử dụng **mạng nơ-ron (neural networks)**, và ở dạng trưởng thành nhất là **học sâu (deep learning).**

**Cách mạng dữ liệu:**

* Dữ liệu đã thay đổi từ việc chỉ là **nguyên liệu thô để cung cấp giải pháp** trở thành **yếu tố định hình thuật toán**, giúp thuật toán trở nên hữu ích hơn cho nhiệm vụ được giao.
* Dữ liệu không chỉ là "nhiên liệu" mà còn là "nghệ nhân" tạo ra chính giải pháp, như minh họa trong **Hình 2-2.**

A graph showing the value of deep learning

Description automatically generated

Hình **2-2**: Với các giải pháp AI hiện tại, **nhiều dữ liệu hơn đồng nghĩa với nhiều trí thông minh hơn**.

Hình minh họa này nhằm nhấn mạnh vai trò quan trọng của dữ liệu trong việc định hình và cải thiện trí thông minh nhân tạo. Khi lượng dữ liệu tăng lên, các thuật toán AI có thể học hỏi, thích nghi và đưa ra các kết quả thông minh hơn.

**Google và sức mạnh của dữ liệu**

* **Dữ liệu miễn phí:** Google khai thác dữ liệu từ nội dung trên các trang web, văn bản và sách có sẵn công khai.
* **Spider Software:** Google sử dụng phần mềm này để thu thập trang web, văn bản và hình ảnh.

**Thuật toán AI của Google giúp:**

* Hiểu và dự đoán nhu cầu của bạn ngay cả khi bạn không diễn đạt bằng các từ khóa rõ ràng.
* Hiểu ngôn ngữ tự nhiên, bao gồm ngôn ngữ nói hàng ngày không chính xác hoặc không rõ ràng.

**Tiến bộ với Hummingbird và RankBrain:**

1. **Hummingbird (2012):**
   * Cải thiện khả năng hiểu từ đồng nghĩa và khái niệm.
   * Tham khảo thêm:  
     [“FAQ: All About The New Google ‘Hummingbird’ Algorithm”](https://searchengineland.com).
2. **RankBrain (2015):**
   * Học từ hàng triệu truy vấn mỗi ngày và trả lời các truy vấn mơ hồ hoặc không rõ ràng, bao gồm tiếng lóng hoặc lỗi chính tả.
   * Không thay thế Hummingbird mà bổ sung và nâng cao hiệu quả.
   * Tham khảo thêm:  
     [“7 Things You May Not Know About Google’s RankBrain”](https://www.act-on.com).

**Sử dụng dữ liệu hiệu quả**

**Dữ liệu phong phú không đảm bảo một AI thành công.**

* Hiện nay, các thuật toán AI không thể tự trích xuất thông tin từ dữ liệu thô.
* **Thuật toán:** Phụ thuộc vào việc thu thập và xử lý dữ liệu trước khi phân tích.

**Nguồn gốc dữ liệu**

1. **Dữ liệu từ con người:**
   * **Ví dụ:** Trải nghiệm mua sắm: Người dùng chọn sản phẩm, thêm vào giỏ hàng, và cung cấp đánh giá. Mỗi trải nghiệm mua sắm trở thành một bài tập thu thập dữ liệu.
   * Dữ liệu thủ công: Tiếp tân nhập thông tin từ người gọi để đặt lịch hẹn.
2. **Dữ liệu từ cảm biến:**
   * **Ví dụ:**
     + Dữ liệu về thời tiết từ các cảm biến đo mưa, nhiệt độ, độ ẩm, mây.
     + Dữ liệu từ robot để sửa lỗi trong quá trình vận hành.
     + Một cảm biến kết hợp với AI nhỏ có thể xác định khi nào bữa tối được nấu chín hoàn hảo.

## Thu thập và xử lý dữ liệu

**Dữ liệu đáng tin cậy** nghe có vẻ dễ định nghĩa nhưng lại rất khó thực hiện.

**Định nghĩa:**  
Dữ liệu được coi là đáng tin cậy khi kết quả mà nó mang lại:

* **Dự đoán được:** Đáp ứng đúng kỳ vọng.
* **Nhất quán:** Kết quả không thay đổi theo thời gian hoặc điều kiện khác nhau.

Một nguồn dữ liệu đáng tin cậy cung cấp **dữ liệu bình thường**, không có bất ngờ. Kết quả không làm ai ngạc nhiên và thường dễ dự đoán.

**Mâu thuẫn trong mong đợi dữ liệu**

Dữ liệu có **tính hai mặt**:

1. **Dữ liệu dự đoán được:**  
   Chúng ta muốn dữ liệu đáng tin cậy, bình thường, và khẳng định những gì đã biết.
2. **Dữ liệu bất ngờ:**  
   Điều làm cho việc thu thập dữ liệu trở nên hữu ích chính là **những điểm bất ngờ**, điều này khiến việc phân tích và xem xét dữ liệu có ý nghĩa.

**Cân bằng khi thu thập dữ liệu**

Tuy nhiên, bạn không muốn dữ liệu quá khác thường đến mức gây khó khăn hoặc lo ngại khi xem xét.

* **Yêu cầu cân bằng:**  
  Dữ liệu cần nằm trong **giới hạn nhất định** (sẽ được mô tả chi tiết trong phần **“Manicuring the Data”**).
  + Dữ liệu cần đáp ứng **giá trị sự thật** (truth value), như sẽ được mô tả trong phần **“Considering the Five Mistruths in Data”**.

**Các yếu tố cần thiết:**

1. **Khoảng thời gian:** Dữ liệu phải đến theo các khoảng thời gian dự kiến.
2. **Đầy đủ trường thông tin:** Mọi trường dữ liệu trong bản ghi nhập phải hoàn chỉnh.

### Làm cho dữ liệu đầu vào của con người đáng tin cậy hơn

**Con người luôn mắc lỗi** — đó là một phần của bản chất con người. Mong đợi rằng con người không phạm lỗi là không thực tế.

* **Thực tế:** Nhiều thiết kế ứng dụng lại giả định rằng con người sẽ không mắc bất kỳ lỗi nào, rằng tất cả mọi người sẽ tuân thủ các quy tắc.
* **Thách thức:** Phần lớn người dùng thậm chí không đọc các quy tắc, vì họ lười biếng hoặc không có đủ thời gian dành cho những việc không mang lại lợi ích trực tiếp.

**Ví dụ về nhập liệu sai lầm**

Xem xét việc nhập **tên tiểu bang** vào một biểu mẫu. Nếu chỉ cung cấp một trường văn bản, người dùng có thể:

1. **Nhập tên đầy đủ:** Chẳng hạn, "Kansas."
2. **Gõ sai chính tả hoặc sai viết hoa:** Kết quả có thể là "Kanzuz," "Kansus," hoặc "kANSAS."
3. **Sử dụng các cách viết tắt khác nhau:**
   * Theo hướng dẫn của **Associated Press (AP):** "Kan."
   * Theo hướng dẫn của **Government Printing Office (GPO):** "Kans."
   * **Bưu điện Hoa Kỳ (USPS):** "KS."
   * **Lực lượng Bảo vệ Bờ biển Hoa Kỳ (U.S. Coast Guard):** "KA."
   * **Tổ chức Tiêu chuẩn Quốc tế (ISO):** "US-KS."

Điều này chỉ áp dụng với một mục đơn giản như nhập tên tiểu bang, tưởng chừng như rõ ràng. Nhưng như ví dụ trên, dữ liệu có thể trở nên phức tạp và không nhất quán.

**Giải pháp: Cung cấp tùy chọn chọn sẵn**

Đối với trường hợp này:

* **Phương án cải thiện:** Sử dụng một danh sách thả xuống (drop-down list) trong biểu mẫu để người dùng chọn tên tiểu bang theo định dạng yêu cầu.
* **Lợi ích:**
  1. Loại bỏ sự khác biệt trong cách viết tắt.
  2. Tránh lỗi chính tả.
  3. Giảm thiểu lỗi viết hoa không đúng cách.

Dù có kiểm tra chéo và mục nhập tĩnh, con người vẫn mắc lỗi

1. **Nhập số liệu không nhất quán:**
   * Khi cần nhập **2.00**, người dùng có thể nhập:
     + 2
     + 2.0
     + 2.
     + Hoặc các biến thể khác.
   * **Giải pháp:** Tự động phân tích và định dạng lại dữ liệu đầu vào. Điều này có thể thực hiện mà không cần người dùng can thiệp.
2. **Vấn đề phổ biến:**  
   Một số trang web yêu cầu nhập thông tin, như số thẻ tín dụng, với:
   * **Dấu gạch ngang (-).**
   * **Dấu cách ( ).**
   * **Không khoảng cách.**  
     Khi ứng dụng không tự động định dạng dữ liệu, điều này gây ra trải nghiệm người dùng khó chịu.

**Giới hạn của việc định dạng lại dữ liệu**

**Lỗi nhập số sai lệch:**

* Không thể sửa lỗi dữ liệu sai chỉ bằng cách định dạng lại.

**Giải pháp:**

1. **Kiểm tra giới hạn giá trị:**
   * Một khách hàng không thể mua -5 bánh xà phòng.
   * Nếu nhập 50,000 bánh xà phòng, điều này có thể là lỗi.
2. **Gợi ý rõ ràng:**
   * Hiển thị thông báo với phạm vi giá trị hợp lệ.
   * Nếu khách hàng thực sự muốn trả lại xà phòng, họ cần sử dụng quy trình **hoàn trả (return)** thay vì đặt hàng mới.

**Sử dụng thu thập dữ liệu tự động**

**Thu thập dữ liệu tự động** mang lại nhiều lợi ích:

* **Tăng tính nhất quán:** Dữ liệu đồng nhất hơn.
* **Cải thiện độ tin cậy:** Giảm thiểu lỗi do con người.
* **Giảm khả năng thiếu sót dữ liệu.**
* **Nâng cao độ chính xác:** Đặc biệt trong các tác vụ liên quan đến thời gian.
* **Giảm độ lệch:** Cho các đầu vào được định giờ.

**Hạn chế:**

1. **Phụ thuộc vào thiết kế con người:**
   * Hệ thống chỉ thu thập những dữ liệu mà con người thiết kế và cho phép.
   * Kết quả thu thập đôi khi không đầy đủ như mong đợi.
2. **Lỗi phần mềm và phần cứng:**
   * **Lỗi mềm (soft issues):** Hệ thống hoạt động nhưng không mang lại kết quả như ý.
   * Khi xảy ra lỗi mềm, hệ thống không thể nhận ra vấn đề như con người, dẫn đến dữ liệu bị hỏng hoặc kém chất lượng.

### Thu thập dữ liệu cá nhân một cách đạo đức

1. **Quan niệm sai lầm về dữ liệu công khai:**
   * Một số người cho rằng bất kỳ thứ gì trên Internet đều là **miền công cộng (public domain)**, bao gồm:
     + **Khuôn mặt.**
     + **Thông tin cá nhân.**
   * **Thực tế:** Bạn nên coi mọi thứ đều được bảo vệ bản quyền và không có sẵn để sử dụng một cách tự do.
2. **Nguyên tắc sử dụng hợp lý (Fair Use):**
   * Một số người hiểu rằng tài liệu được bảo vệ bản quyền nhưng vẫn dựa vào **nguyên tắc sử dụng hợp lý (fair use).**
   * **Fair use rất phức tạp:**
     + Ví dụ, vụ kiện **Author’s Guild v. Google** được quyết định có lợi cho Google.
     + **Tham khảo:** [“The Most Important Court Decision For Data Science and Machine Learning”](https://towardsdatascience.com) trên **Towards Data Science.com**.

Ngoài ra, vấn đề ‘fair use’ áp dụng với sách, chứ không áp dụng cho con người

Khi thảo luận về **fair use**, cần lưu ý rằng quyền riêng tư của cá nhân cũng là một vấn đề quan trọng.

* **Ví dụ:** Các công ty bắt đầu **thu thập hình ảnh của người đeo khẩu trang** từ Internet mà không có sự cho phép, dẫn đến tranh cãi lớn.  
  **Tham khảo:** [“Your face mask selfies could be training the next facial recognition tool”](https://www.cnet.com) trên **Cnet.com**.

**Các vụ kiện về quyền riêng tư**

1. **Facebook và vụ kiện dữ liệu khuôn mặt:**  
   Facebook đang bị kiện vì sử dụng sai dữ liệu người dùng.  
   **Tham khảo:** [“$35B face data lawsuit against Facebook will proceed”](https://www.techcrunch.com) trên **TechCrunch.com**.
2. **Công cụ bảo vệ quyền riêng tư khuôn mặt:**  
   Một ngành công nghiệp mới đã xuất hiện, phát triển công cụ làm cho khuôn mặt của một người trở nên khó sử dụng hơn với các công ty thu thập dữ liệu trái phép.  
   **Tham khảo:** [“This Tool Could Protect Your Photos From Facial Recognition”](https://www.nytimes.com) trên **New York Times**.

**Tóm lại:** Bất kể quan điểm của bạn về vấn đề sử dụng dữ liệu miễn phí, bạn cần đảm bảo **sử dụng dữ liệu một cách có đạo đức**, bất kể nguồn gốc của nó.

**Các cân nhắc khi thu thập dữ liệu cá nhân một cách đạo đức**

1. **Xin phép:**
   * Nếu nghiên cứu yêu cầu nhận dạng cá nhân trong một tập dữ liệu, việc **thu thập thông tin nhận dạng cá nhân (PII)** mà không xin phép là không đúng.
   * **Rủi ro:**
     + Dữ liệu có thể không đầy đủ hoặc không chính xác, khiến phân tích không đáng tin cậy.
     + Có thể gặp hậu quả pháp lý đắt đỏ và phức tạp.
   * **Giải pháp:**
     + Tốt nhất là **hỏi xin phép**.
     + Tham khảo nguồn thông tin, chẳng hạn như:  
       **“How to Obtain a Consumer’s Authorization before Gaining Access to Personally Identifiable Information (PII)”**.
2. **Kỹ thuật làm sạch dữ liệu (Sanitization):**
   * **Mục tiêu:** Xóa thông tin cá nhân như tên, địa chỉ, số điện thoại, ID, v.v., để không thể nhận dạng một cá nhân cụ thể từ tập dữ liệu.
   * **Ví dụ:**
     + Làm mờ khuôn mặt trong ảnh.
     + Xóa biển số xe trong hình ảnh.
   * **Lưu ý:** Khi tìm kiếm thông tin về **sanitization**, hãy thêm từ khóa "privacy" để tìm đúng bài viết liên quan.
3. **Tránh suy diễn dữ liệu (Data Inference):**
   * Một số người không muốn chia sẻ thông tin cá nhân như giới tính hoặc tuổi.
   * **Vấn đề:**
     + Suy diễn dựa trên tên có thể sai lệch, vì tên có thể khác giới tính trong các nền văn hóa khác nhau.
     + Sai lệch tuổi:
       - **Ví dụ:** Máy học có thể suy diễn sai tuổi của một người bạch tạng (albino), ảnh hưởng đến 1/3.000 người tùy khu vực.
       - **Tham khảo:** [“Information Bulletin – What is Albinism?”](https://www.albinism.org).
4. **Tránh tổng quát hóa (Generalizations):**
   * Áp dụng sai các kết quả từ thống kê và học máy vào cá nhân có thể dẫn đến phân biệt đối xử.
   * **Lưu ý:** Thống kê chỉ áp dụng cho **nhóm**, không phải **cá nhân.**

### Chăm sóc và xử lý dữ liệu

Một số người sử dụng thuật ngữ **"thao túng dữ liệu (data manipulation)"** khi nói về việc xử lý dữ liệu, khiến nó mang hàm ý rằng dữ liệu bị thay đổi một cách không trung thực hoặc mờ ám.

* **Thuật ngữ thay thế:** Có thể sử dụng thuật ngữ **"chăm sóc dữ liệu (data manicuring)"**, mang ý nghĩa làm cho dữ liệu trở nên có hình thức và giá trị hơn.

**Thực tế:** Dữ liệu thô hiếm khi đáp ứng yêu cầu xử lý và phân tích.

* Để khai thác giá trị từ dữ liệu, bạn cần phải xử lý và định hình nó để phù hợp với các nhu cầu cụ thể.

### Xử lý dữ liệu bị thiếu

1. **Tầm quan trọng của dữ liệu đầy đủ:**
   * Để trả lời một câu hỏi chính xác, bạn cần **đầy đủ các dữ kiện (facts).**
   * Nếu dữ liệu không đầy đủ, câu trả lời có thể chỉ là một **phỏng đoán**, với khả năng đúng hoặc sai như nhau.
2. **Hậu quả của dữ liệu thiếu:**
   * Khi đưa ra quyết định mà không có đầy đủ dữ kiện, điều này được gọi là **"kết luận vội vàng (jumping to a conclusion)."**
   * Trong phân tích dữ liệu, bạn có thể đã đưa ra nhiều kết luận vội vàng hơn bạn nghĩ, chỉ vì dữ liệu bị thiếu.
3. **Dữ liệu bị thiếu trong hồ sơ dữ liệu (data records):**
   * **Hồ sơ dữ liệu:** Một mục nhập trong tập dữ liệu, chứa các **trường (fields)** đại diện cho từng loại dữ liệu cụ thể.
   * Nếu một trường bị bỏ trống, bạn không có dữ liệu cần thiết để trả lời câu hỏi hoặc đưa ra quyết định chính xác.

### Xử lý dữ liệu thô

**1. Đối mặt với dữ liệu thiếu**

Khi phân tích dữ liệu, việc thiếu dữ liệu có thể dẫn đến kết quả sai lệch. Dữ liệu thiếu có thể chia thành hai loại chính:

1. **Dữ liệu cần thiết bị thiếu (Essential data missing):**
   * Nếu quá trình thu thập không bao gồm toàn bộ dữ liệu cần thiết để trả lời một câu hỏi cụ thể, kết quả sẽ không đáng tin cậy.
   * **Giải pháp:** Đôi khi tốt hơn là loại bỏ dữ liệu hỏng nặng thay vì sử dụng chúng.
2. **Dữ liệu bị thiếu một phần (Some data missing):**
   * **Dữ liệu thiếu ngẫu nhiên (Randomly missing data):**
     + Nguyên nhân: Lỗi con người hoặc cảm biến.
     + **Giải pháp:** Thay thế bằng giá trị trung vị hoặc trung bình. Mặc dù không hoàn toàn chính xác, nhưng đủ để đưa ra câu trả lời hợp lý.
   * **Dữ liệu thiếu tuần tự (Sequentially missing data):**
     + Nguyên nhân: Lỗi hệ thống tổng quát.
     + **Giải pháp:**
       - Khó khăn hoặc không thể sửa chữa vì thiếu dữ liệu xung quanh để dự đoán.
       - Nếu xác định được nguyên nhân, có thể tái tạo dữ liệu.

**2. Xem xét sự không đồng bộ của dữ liệu (Data misalignments)**

Dữ liệu trong một tập hợp có thể không khớp với dữ liệu trong tập hợp khác.

1. **Dạng lỗi không đồng bộ:**
   * **Dạng số:**
     + Ví dụ: Một tập dữ liệu sử dụng số thập phân (floating-point), trong khi tập khác sử dụng số nguyên (integer).
     + **Giải pháp:** Đồng bộ hóa kiểu dữ liệu trước khi kết hợp hai tập dữ liệu.
   * **Dạng ngày:**
     + Ví dụ: Ngày tháng có thể được định dạng khác nhau hoặc sử dụng các múi giờ khác nhau.
     + **Giải pháp:**
       - Đồng bộ hóa về cùng múi giờ.
       - Xử lý các trường hợp phức tạp, như sự khác biệt giữa múi giờ có **giờ mùa hè (DST)** và không có DST.
2. **Sự khác biệt trong trường dữ liệu:**
   * Ví dụ:
     + Một tập dữ liệu gộp họ và tên thành một trường duy nhất, trong khi tập khác tách biệt họ và tên.
     + **Giải pháp:**
       - Chuyển đổi tất cả các tập dữ liệu để sử dụng một trường duy nhất hoặc tách biệt các trường.
3. **Các giải pháp khả thi:**
   * Tính toán dữ liệu thiếu từ dữ liệu hiện có.
   * Tìm kiếm dữ liệu thiếu trong tập dữ liệu khác.
   * Kết hợp các tập dữ liệu để tạo thành một tập đầy đủ.
   * Thu thập thêm dữ liệu từ các nguồn khác.
   * Định nghĩa lại câu hỏi để không cần dữ liệu thiếu.

**3. Lọc dữ liệu hữu ích (Separating useful data)**

Nhiều tổ chức cho rằng **không bao giờ có quá nhiều dữ liệu**, nhưng **dữ liệu dư thừa** cũng là một vấn đề lớn như dữ liệu thiếu.

1. **Vấn đề của dữ liệu quá tải:**
   * **Tốn thời gian:** Cần nhiều thời gian hơn để tìm ra giải pháp khi phải xử lý lượng dữ liệu thừa thãi.
   * **Kết quả gây nhầm lẫn:** Quá nhiều dữ liệu có thể làm che khuất vấn đề thực sự.
2. **Giải pháp:**
   * Định nghĩa câu hỏi cần trả lời một cách rõ ràng và ngắn gọn.
   * Sử dụng thuật toán phù hợp hoặc tổ hợp thuật toán để xử lý dữ liệu hiệu quả hơn.

### Xử lý dữ liệu bị hỏng

**Dữ liệu bị hỏng** có thể đến từ nhiều nguyên nhân khác nhau:

1. **Đánh giá sản phẩm trực tuyến:**
   * Không chắc rằng:
     + Người thật tạo ra các đánh giá.
     + Không có ai bỏ phiếu nhiều lần.
     + Đánh giá không bị ảnh hưởng bởi ngày xấu (hoặc tốt) bất thường.
   * Đánh giá có thể quá thiên lệch đến mức không thể tin cậy.
2. **Nguồn gốc hỏng dữ liệu khác:**
   * Cảm biến bị hỏng, cung cấp dữ liệu sai lệch.
   * Tấn công bởi virus gây lỗi dữ liệu.
   * Lỗi trong cơ sở dữ liệu hoặc phần mềm.
   * Nhập liệu sai từ con người.
   * Thiên tai, như sét đánh, gây lỗi tạm thời.

**Giải pháp giảm thiểu dữ liệu bị hỏng**

1. **Lưu trữ trên đám mây:**
   * Giảm thiểu vấn đề do phần cứng hoặc thời tiết gây ra.
2. **Thủ tục và đào tạo:**
   * Đảm bảo nhân viên được đào tạo tốt và áp dụng quy trình chuẩn.
3. **Giám sát liên tục:**
   * Chủ động theo dõi để phát hiện và sửa lỗi sớm.
4. **Tường lửa và quản trị viên chủ động:**
   * Bảo vệ hệ thống trước các nguy cơ tấn công.

**Hủy botnet bằng phương pháp "sinkholing"**

**Botnet** là các nhóm máy tính được phối hợp hoạt động để thực hiện các tác vụ cụ thể, mà phần lớn là **hành vi độc hại.**

**Tác động của botnet đến dữ liệu:**

* **Cung cấp dữ liệu sai lệch:** Botnet có thể làm nhiễu dữ liệu của tập dữ liệu bằng cách đưa vào các thông tin sai.
* **Chiếm quyền kiểm soát tài khoản:** Botnet có thể xâm nhập tài khoản, thay đổi thông tin tài khoản theo những cách nhất định.
* **Hậu quả:** Botnet làm **hỏng hoặc thiên lệch dữ liệu,** khiến mọi phân tích dựa trên dữ liệu đó đều thất bại.

**Phương pháp xử lý: Sinkholing**

**Sinkholing** là một trong những phương pháp hiệu quả nhất để xử lý botnet.

* **Định nghĩa:** Sinkholing có nghĩa là **chuyển hướng botnet** đến một vị trí mà chúng không thể gây hại.

**Tham khảo thêm:**

Bài viết **“Hacker Lexicon: What Is Sinkholing?”** trên **Wired** cung cấp các kỹ thuật cụ thể để thực hiện phương pháp này.

## 5 Dạng Sai Lệch Trong Dữ Liệu

Con người thường nhận thức dữ liệu dưới dạng ý kiến, và trong một số trường hợp, dữ liệu bị bóp méo đến mức trở thành **sai lệch (mistruth)**.

* Máy tính không thể phân biệt dữ liệu đúng và sai — nó chỉ xử lý dữ liệu mà thôi.
* Một trong những thách thức lớn trong việc tạo ra AI có thể suy nghĩ như con người là:  
  **Con người có thể làm việc với sai lệch, nhưng máy tính thì không.**

**Cách xử lý sai lệch:**

* Phát hiện dữ liệu sai như các **ngoại lệ (outliers)** và loại bỏ chúng.
* Tuy nhiên, phương pháp này không luôn giải quyết được vấn đề, vì con người vẫn có thể sử dụng dữ liệu sai lệch để đưa ra kết luận.

**Ví dụ: Tai nạn giao thông**

Dưới đây là năm loại sai lệch thường thấy trong dữ liệu, minh họa qua ví dụ tai nạn giao thông:

### 1. Sai lệch do cố ý (Commission)

**Mô tả:**  
Sai lệch này xảy ra khi ai đó cố tình thay thế thông tin đúng bằng thông tin sai.

**Ví dụ:**

* Một người điền vào báo cáo tai nạn rằng họ bị **mặt trời làm chói mắt**, khiến họ không nhìn thấy người bị đâm.
* Thực tế, người này có thể đã mất tập trung vì điều gì đó khác hoặc đang nghĩ về bữa tối thay vì lái xe.

**Hệ quả:**

* Nếu không có ai chứng minh ngược lại, người này có thể thoát với một mức phạt nhẹ hơn.
* Dữ liệu bị ô nhiễm, và một công ty bảo hiểm có thể đưa ra mức phí bảo hiểm dựa trên thông tin sai lệch.

### 2. Sai lệch do bỏ sót (Omission)

**Sai lệch do bỏ sót** xảy ra khi mọi thông tin cung cấp đều đúng, nhưng một thông tin quan trọng bị bỏ qua, làm thay đổi toàn bộ nhận thức về sự việc.

**Ví dụ: Báo cáo tai nạn**

* Bạn khai báo:
  + Đường ướt.
  + Trời nhá nhem tối.
  + Bạn phản ứng chậm khi phanh.
  + Con hươu lao ra từ bụi cây.
* **Thực tế bị bỏ sót:** Bạn đang nhắn tin khi lái xe.

Nếu thông tin này được bổ sung:

* **Nguyên nhân tai nạn:** Chuyển từ "tai nạn bất khả kháng" thành "lái xe mất tập trung."
* **Hậu quả:**
  + Bạn có thể bị phạt.
  + Công ty bảo hiểm có thể điều chỉnh phí bảo hiểm theo mức cao hơn.

**Ảnh hưởng đến dữ liệu:**  
Dữ liệu thiếu chính xác sẽ dẫn đến sai lệch trong phân tích và ra quyết định.

### 3. Sai lệch do góc nhìn (Perspective)

**Sai lệch do góc nhìn** xảy ra khi các bên liên quan cùng tường thuật một sự kiện nhưng từ các góc nhìn khác nhau, dẫn đến sự khác biệt trong dữ liệu được ghi nhận.

**Ví dụ: Tai nạn giữa xe hơi và người đi bộ**

* **Tài xế:**
  + Thấy rõ bảng điều khiển xe và trạng thái xe tại thời điểm tai nạn.
  + Có thể cung cấp dữ liệu kỹ thuật, nhưng không biết về cảm xúc của người bị nạn.
* **Người bị nạn:**
  + Nhìn rõ biểu cảm trên khuôn mặt tài xế (có cố ý hay không).
  + Không biết về tình trạng xe hoặc hành động của tài xế.
* **Người chứng kiến:**
  + Có cái nhìn toàn cảnh:
    - Tài xế có cố gắng phanh hay né tránh không.
    - Các yếu tố ngoại cảnh khác.

**Kết quả:**

* **Báo cáo:** Sẽ phản ánh trung bình của các góc nhìn, thường được bổ sung bởi kinh nghiệm cá nhân của người lập báo cáo (chẳng hạn như cảnh sát).
* **Với AI:** Dữ liệu như vậy thường **không đủ chính xác**, vì AI cần thông tin đầy đủ và không thiên lệch để phân tích hiệu quả.

### 4.Thiên kiến (Bias)

Thiên kiến xuất hiện khi **người phân tích có khả năng nhìn thấy sự thật** nhưng vì quan điểm cá nhân hoặc niềm tin, họ không nhận thức được điều đó.

**Ví dụ:**  
Một người lái xe tập trung quá mức vào giữa đường và không nhận thấy con hươu ở rìa đường.

* Kết quả: Người lái không kịp phản ứng khi con hươu lao ra giữa đường.

**Phân loại thiên kiến:**

1. **Khó xác định:**
   * Một người không nhìn thấy con hươu vì:
     + Nó bị khuất sau bụi cây (**tai nạn thực sự**).
     + Họ không tập trung đúng cách (**thiếu chú ý**).
     + Họ bị phân tâm trong giây lát (**mất tập trung thoáng qua**).
2. **Xác định nguồn gốc thiên kiến:**
   * Điều này rất quan trọng khi thiết kế **thuật toán AI** để tránh các nguồn gốc thiên kiến.

### 5.Khung tham chiếu (Frame of Reference)

Khác với thiên kiến, khung tham chiếu thường xuất phát từ sự **hiểu nhầm** hơn là lỗi dữ liệu.

**Khung tham chiếu và hiểu nhầm:**

1. **Khi người nghe thiếu kinh nghiệm thực tế:**
   * Một sự kiện được mô tả nhưng không thể hiểu rõ vì người nghe không có kinh nghiệm thực tế liên quan.
   * **Ví dụ hài hước:** Đoạn kịch nổi tiếng **“Who’s On First”** của Abbott và Costello, dễ dàng tìm thấy trên **YouTube.com**.
2. **Khi trải nghiệm không thể truyền đạt đầy đủ:**
   * Một thủy thủ trải qua cơn bão trên biển — một cơn bão dữ dội, có thể nguy hiểm đến tính mạng.
   * Dù sử dụng video, phỏng vấn, hoặc mô phỏng, **trải nghiệm thực tế** không thể được truyền đạt hoàn chỉnh cho người chưa từng trải qua cơn bão.

**Khung tham chiếu và AI**

AI luôn gặp vấn đề với **khung tham chiếu** vì:

1. AI không thể tự tạo ra trải nghiệm.
2. **Ngân hàng dữ liệu (databank)** chỉ chứa các sự thật (facts) mà không bao gồm các kết luận hoặc cảm nhận dựa trên trải nghiệm.
3. **Trải nghiệm con người:** Dựa trên sự kết hợp giữa sự thật và các kết luận mà công nghệ hiện tại chưa thể sao chép.

## Xác định giới hạn của việc thu thập dữ liệu

Ngày nay, dữ liệu được thu thập bởi các tổ chức với tốc độ và khối lượng đáng kinh ngạc, thường không có mục tiêu hay ý định rõ ràng.

**Ví dụ về tốc độ thu thập dữ liệu:**  
Theo **“Data Never Sleeps”** trên [Domo.com](https://domo.com):

* Zoom tổ chức **208,333** người tham gia họp mỗi phút.
* Người dùng đăng tải **347,222** câu chuyện trên Instagram.
* Microsoft Teams kết nối **52,083** người dùng.
* Người dùng chia sẻ **41,666,667** tin nhắn trên WhatsApp.
* Mọi người thực hiện **1,388,889** cuộc gọi video/thoại.

**Cơn nghiện thu thập dữ liệu**

**Vấn đề chính:**

* Việc thu thập dữ liệu đã trở thành một **cơn nghiện** đối với các tổ chức trên toàn cầu, như thể tổ chức nào thu thập nhiều nhất sẽ chiến thắng một giải thưởng nào đó.
* Tuy nhiên, **thu thập dữ liệu đơn thuần không mang lại giá trị.**

**Minh họa qua "The Hitchhiker’s Guide to the Galaxy"**

Trong cuốn sách **"The Hitchhiker’s Guide to the Galaxy"** của Douglas Adams:

* Một chủng tộc siêu việt đã xây dựng một chiếc máy tính khổng lồ để tính toán ý nghĩa của "cuộc sống, vũ trụ, và mọi thứ."
* **Câu trả lời:** Số **42**, nhưng lại không giải quyết được bất kỳ điều gì.
* **Kết luận:**
  + Dữ liệu thu thập, phân loại, và phân tích không mang lại kết quả hữu ích nếu **câu hỏi được đặt ra không đúng.**
  + Bạn cần **biết câu hỏi** để làm cho dữ liệu và câu trả lời có ý nghĩa.

**Tạo ra câu hỏi đúng**

Thu thập dữ liệu không giới hạn không phải là mục tiêu, mà là việc xác định đúng câu hỏi cần giải quyết.

**Ví dụ về các câu hỏi liên quan đến kinh doanh:**

1. Mỗi ngày có bao nhiêu người đi qua trước cửa hàng?
2. Bao nhiêu người dừng lại để nhìn vào cửa sổ?
3. Họ nhìn trong bao lâu?
4. Thời điểm nào trong ngày họ nhìn nhiều nhất?
5. Những kiểu trưng bày nào mang lại kết quả tốt hơn?
6. Kiểu trưng bày nào khiến họ bước vào cửa hàng và mua sắm?

**Quy trình tạo câu hỏi hiệu quả:**

1. **Xây dựng danh sách câu hỏi:** Tập trung vào nhu cầu kinh doanh cụ thể.
2. **Xác minh từng câu hỏi:**
   * Câu hỏi có thực sự quan trọng không?
   * Có giải quyết nhu cầu thực sự không?
3. **Xác định thông tin cần thiết:** Để trả lời câu hỏi, cần loại dữ liệu nào?

## Xem xét các vấn đề bảo mật dữ liệu

Bảo mật dữ liệu không chỉ đơn thuần là ngăn chặn việc **đánh cắp dữ liệu** hoặc **bảo vệ quyền riêng tư**.

* **Mục tiêu:** Đảm bảo **tính toàn vẹn của dữ liệu** để dữ liệu có thể được sử dụng hiệu quả.
* **Thách thức:** Làm cho dữ liệu có thể truy cập được đồng nghĩa với việc chấp nhận rủi ro rằng ai đó có thể sử dụng dữ liệu không đúng mục đích.

**Hiểu về dữ liệu có chủ ý thiên lệch**

**Thiên lệch dữ liệu (bias)** xuất hiện trong hầu hết mọi bộ dữ liệu hiện nay:

* **Nguyên nhân:**
  + **Phương pháp thu thập dữ liệu** có thiên lệch.
  + **Phương pháp phân tích dữ liệu** có thiên lệch.
  + **Chính dữ liệu** cũng có thể thiên lệch.

Một số trường hợp dữ liệu bị cố tình thiên lệch:

1. **Chính trị:**
   * **Ví dụ:** Hai nhóm có quan điểm đối lập sử dụng cùng một bộ dữ liệu nhưng tạo ra kết quả hoàn toàn khác nhau để ủng hộ quan điểm của mình.
   * **Kỹ thuật:** Sử dụng **người trả lời giả** trong các cuộc khảo sát.
   * **Tham khảo:**  
     [“Assessing the Risks to Online Polls From Bogus Respondents”](https://www.pewresearch.org).
2. **Y tế:**
   * **Ví dụ:** Các thử nghiệm vaccine Covid-19 không bao gồm đủ **người cao tuổi** và **người thiểu số**, dẫn đến dữ liệu thiên lệch.
   * **Tham khảo:**  
     [“Older Adults, Minorities Underrepresented in COVID-19 Vaccine Trials”](https://www.aarp.org).
3. **Pháp lý:**
   * **Ví dụ:** Thuật toán COMPAS dùng để dự đoán **tái phạm tội** bị chỉ trích vì thiên lệch dữ liệu.
   * **Tham khảo:**  
     [“Injustice Ex Machina: Predictive Algorithms in Criminal Sentencing”](https://www.uclalawreview.org).
4. **Tuyển dụng:**
   * **Ví dụ:** Dữ liệu tuyển dụng có thể phản ánh và khuếch đại các thực hành thiên lệch trong tổ chức hoặc xã hội.
   * **Tham khảo:**  
     [“All the Ways Hiring Algorithms Can Introduce Bias”](https://www.hbr.org).
5. **Các lĩnh vực khác:**
   * **Fairness in Machine Learning:** Ý tưởng rằng kết quả phân tích phải phản ánh đúng thực tế xã hội.
   * **Tham khảo:**  
     [“A Tutorial on Fairness in Machine Learning”](https://towardsdatascience.com).

# Chương 3: Xem xét việc sử dụng thuật toán

**NỘI DUNG CHƯƠNG NÀY**

* Tìm hiểu vai trò của thuật toán trong AI.
* Chiến thắng trò chơi bằng tìm kiếm không gian trạng thái và thuật toán min-max.
* Phân tích cách hoạt động của hệ thống chuyên gia.
* Xem rằng học máy (machine learning) và học sâu (deep learning) là một phần của AI.

**Dữ liệu: Yếu tố thay đổi cục diện trong AI**

Những tiến bộ trong AI cho thấy rằng đối với một số vấn đề, **lượng dữ liệu phù hợp quan trọng hơn thuật toán tốt nhất.**

* **Ví dụ:**  
  Năm 2001, hai nhà nghiên cứu từ Microsoft, **Banko** và **Brill**, đã công bố một bài báo nổi bật:  
  **“Scaling to Very Very Large Corpora for Natural Language Disambiguation.”**
  + Kết luận: Nếu muốn máy tính xây dựng mô hình ngôn ngữ, bạn không cần thuật toán thông minh nhất.
  + Khi đưa hơn **một tỷ từ** trong ngữ cảnh vào bài toán, **bất kỳ thuật toán nào cũng bắt đầu hoạt động rất tốt.**

Chương này giúp bạn hiểu **mối quan hệ giữa thuật toán và dữ liệu** để thuật toán thực hiện công việc hữu ích.

**Tầm quan trọng của thuật toán và phân tích dữ liệu**

* **Dữ liệu phong phú:**  
  Dù có bao nhiêu dữ liệu, bạn vẫn cần một thuật toán để làm cho dữ liệu trở nên hữu ích.
* **Phân tích dữ liệu:**  
  Bạn phải thực hiện **phân tích dữ liệu** (một loạt các bước có thể định nghĩa được) để dữ liệu hoạt động hiệu quả với thuật toán đã chọn.
* **Không có đường tắt:**
  + Dù AI là **tự động hóa thông minh**, nhưng đôi khi tự động hóa phải nhường chỗ cho phân tích.
  + Máy học hoàn toàn tự động, không cần can thiệp của con người, vẫn là **tương lai xa.**

**Vai trò của các thành phần AI**

Phần thứ hai của chương này sẽ giúp bạn hiểu vai trò của:

* **Hệ thống chuyên gia (expert systems).**
* **Học máy (machine learning).**
* **Học sâu (deep learning).**
* Các ứng dụng như **AlphaGo**, trong việc đưa các khả năng tương lai đến gần hiện thực hơn.

## Vai trò của thuật toán trong AI

**Thuật toán (algorithm)** là một quy trình bao gồm các **chuỗi hoạt động**, thường được máy tính thực hiện để:

1. Tìm ra **giải pháp đúng** cho một vấn đề trong thời gian hữu hạn.
2. Thông báo rằng **không có giải pháp** tồn tại.

Mặc dù con người đã giải các thuật toán thủ công trong hàng nghìn năm, nhưng việc làm này có thể tốn rất nhiều thời gian và yêu cầu các phép tính phức tạp, đặc biệt với các vấn đề khó.

### AI và thuật toán

1. **Bản chất của thuật toán:**
   * Thuật toán phản ánh trí thông minh được mã hóa của con người đã tạo ra chúng.
   * Máy móc thực hiện các thuật toán này chỉ có thể phản ánh trí thông minh của con người, chứ không thể **thay thế hoặc mô phỏng sáng tạo của con người** trong việc tạo ra thuật toán mới.
2. **Nhận diện AI qua các công cụ:**
   * AI được công nhận khi một công cụ đưa ra phương pháp mới và tương tác theo cách giống con người.
   * **Ví dụ:** Trợ lý số như **Siri, Alexa, Cortana, và Google Assistant**.
3. **AI phổ biến nhưng ít được chú ý:**
   * Một số công cụ thông thường như **định tuyến GPS** hay các **bộ điều khiển thông minh (như thermostat)** không được nhận diện là AI do chúng quá quen thuộc và hoạt động ẩn sau hậu trường.
   * Điều này được gọi là **Hiệu ứng AI (AI Effect)**, được đặt tên bởi **Pamela McCorduck**, tác giả cuốn sách **Machines Who Think** (1979).
   * **Hiệu ứng AI:** Khi các chương trình máy tính thông minh hoạt động hiệu quả, chúng trở thành "diễn viên thầm lặng" và sự chú ý chuyển sang các vấn đề AI mới cần giải quyết.

### Định nghĩa thuật toán

Một thuật toán luôn trình bày một **chuỗi bước**, nhưng không nhất thiết phải thực hiện tất cả các bước này để giải quyết một vấn đề (một số bước có thể là tùy chọn).

**Ví dụ thuật toán pha trà:**

1. Đổ nước vào ấm.
2. Bật bếp để đun nước.
3. Khi nước sôi, đổ vào cốc.
4. Cho túi trà vào cốc và ngâm trà trong thời gian khuyến nghị.
5. Lấy túi trà ra.
6. (Tùy chọn) Thêm đường.
7. (Tùy chọn) Thêm sữa.
8. Uống trà.
9. (Tùy chọn) Đổ trà đi nếu không uống được.

### Phạm vi của thuật toán

Thuật toán có phạm vi rất lớn, có thể thực hiện:

* **Lưu trữ dữ liệu.**
* **Sắp xếp, cấu trúc dữ liệu.**
* Giải quyết vấn đề trong các lĩnh vực như:
  + Khoa học.
  + Y học.
  + Tài chính.
  + Sản xuất công nghiệp.
  + Truyền thông.

**Heuristics (thuật toán suy đoán):**

* Một phân lớp thuật toán tạo ra giải pháp **tốt** (nhưng không nhất thiết là hoàn hảo) khi thời gian quan trọng hơn độ chính xác.

### AI và các thuật toán phức tạp

Thuật toán AI khác biệt ở chỗ:

* Giải quyết những vấn đề phức tạp, thường là **NP-complete problems** (các vấn đề không xác định được lời giải trong thời gian đa thức).
* Những vấn đề này yêu cầu kết hợp giữa **lý luận và trực giác** của con người.

**Ví dụ thuật toán AI:**

1. **Lập lịch và phân bổ tài nguyên khan hiếm.**
2. **Tìm đường đi trong không gian phức tạp.**
3. **Nhận diện mẫu (hình ảnh, âm thanh):**
   * Nhận diện khuôn mặt, tiếng nói.
4. **Xử lý ngôn ngữ tự nhiên:**
   * Hiểu văn bản.
   * Dịch ngôn ngữ.
5. **Chơi và thắng trong các trò chơi cạnh tranh.**

## Lập kế hoạch và thuật toán phân nhánh Min-Max

**Lập kế hoạch** giúp xác định trình tự hành động cần thực hiện để đạt được một mục tiêu nhất định. Việc lập kế hoạch là một vấn đề cổ điển trong AI, với các ví dụ trong:

* **Sản xuất công nghiệp.**
* **Phân bổ tài nguyên.**
* **Di chuyển robot trong một không gian.**

**Cách thức hoạt động**

1. **Bắt đầu từ trạng thái hiện tại:**
   * AI xác định tất cả các hành động có thể thực hiện từ trạng thái này.
2. **Mở rộng trạng thái:**
   * Mở rộng trạng thái hiện tại thành nhiều trạng thái tương lai.
   * Sau đó, AI tiếp tục mở rộng các trạng thái tương lai thành các trạng thái tương lai khác.
3. **Dừng khi không thể mở rộng thêm:**
   * Khi không thể mở rộng thêm, AI tạo ra một **không gian trạng thái (state space)**, bao gồm tất cả những gì có thể xảy ra trong tương lai.

**Sử dụng không gian trạng thái**

* **Dự đoán mọi khả năng:** AI không chỉ dự đoán những gì có thể xảy ra mà còn xác định xác suất xảy ra của từng trạng thái.
* **Hỗ trợ ra quyết định:** Không gian trạng thái được sử dụng để khám phá các quyết định giúp AI đạt được mục tiêu theo cách tối ưu nhất.

Quá trình này được gọi là **tìm kiếm trong không gian trạng thái (state-space search).**

### Cấu trúc dữ liệu và thuật toán

Làm việc với không gian trạng thái yêu cầu:

1. **Cấu trúc dữ liệu:**
   * Cây (trees).
   * Đồ thị (graphs).
2. **Thuật toán tìm kiếm:**
   * **Tìm kiếm theo chiều rộng (Breadth-First Search - BFS).**
   * **Tìm kiếm theo chiều sâu (Depth-First Search - DFS).**

**Cây và nút**

Cây trong AI được xây dựng theo cách tương tự như cây trong tự nhiên:

* **Mỗi phần tử được thêm vào cây là một nút (node).**
* **Các nút kết nối với nhau thông qua liên kết (links).**
* **Tổng hợp các nút và liên kết tạo thành một cấu trúc giống cây.**

Hình minh họa **3-1** mô tả cấu trúc cây này.

A diagram of a tree

Description automatically generated

**Mở rộng cây bằng các nút đồ thị**

Một **đồ thị (graph)** là một dạng mở rộng của cây (tree). Tương tự như cây, đồ thị có các **nút (nodes)** kết nối với nhau để tạo thành mối quan hệ. Tuy nhiên, khác với **cây nhị phân (binary trees),** một nút trong đồ thị có thể có nhiều hơn một hoặc hai kết nối.

**Đặc điểm chính của đồ thị:**

1. **Nhiều kết nối:** Các nút trong đồ thị thường có một số lượng lớn kết nối.
2. **Kết nối đa chiều:** Các nút có thể kết nối theo bất kỳ hướng nào, không chỉ theo hướng từ cha đến con như trong cây.

### Cấu trúc và ứng dụng của đồ thị

1. **Cấu trúc đồ thị:**
   * Đồ thị bao gồm các **nút (nodes hoặc vertexes)** được kết nối bằng các **cạnh (edges hoặc arcs).**
   * Ví dụ: Trong bản đồ, mỗi vị trí trên bản đồ được xem là một **nút,** và các con đường là các **cạnh.**

**Khác biệt so với cây:**  
Trong cây, mỗi nhánh sẽ kết thúc tại một **nút lá (leaf node),** trong khi đồ thị không nhất thiết phải có điểm kết thúc.

**Ví dụ minh họa:**  
Tham khảo **Hình 3-2** để xem một đồ thị được biểu diễn.

1. **Ứng dụng thực tế:**
   * Đồ thị rất hữu ích trong việc biểu diễn các trạng thái không gian vật lý.
   * **Ví dụ:** Hệ thống định vị GPS sử dụng đồ thị để biểu diễn các địa điểm và đường phố.

**Tính năng nâng cao của đồ thị**

1. **Định hướng (Directionality):**
   * Một đồ thị có thể chứa khái niệm định hướng.
   * **Khác với cây,** trong đó các mối quan hệ cha/con luôn định hướng, đồ thị cho phép các nút kết nối với nhau với một hướng cụ thể.

**Ví dụ:**  
Trong các con đường thành phố:

* + Hầu hết các con đường là hai chiều (bidirectional).
  + Một số là đường một chiều (one-way streets) chỉ cho phép di chuyển theo một hướng.

1. **Trọng số (Weights):**
   * Một đồ thị có thể gán **trọng số** cho các kết nối cụ thể.
   * Trọng số này có thể đại diện cho:
     + **Khoảng cách** giữa hai điểm.
     + **Thời gian** cần thiết để di chuyển qua tuyến đường.
     + **Nhiên liệu** tiêu hao trên lộ trình.
     + Hoặc các thông tin khác.

A diagram of a flowchart

Description automatically generated

### Duyệt qua đồ thị

Duyệt qua đồ thị là quá trình tìm kiếm (hoặc thăm) từng đỉnh (vertex/nút) trong một thứ tự cụ thể.

* **Thăm một đỉnh:** Bao gồm việc đọc và/hoặc cập nhật nó.
* **Phát hiện đỉnh chưa được thăm:** Khi duyệt qua đồ thị, bạn phát hiện các đỉnh mới. Đỉnh có thể:
  + **Được phát hiện:** Vì bạn vừa thăm nó.
  + **Được xử lý:** Vì thuật toán đã thử tất cả các cạnh xuất phát từ nó.

**Thứ tự duyệt đồ thị** xác định loại tìm kiếm được thực hiện. Có hai kiểu tìm kiếm chính:

1. **Tìm kiếm không thông tin (Uninformed/Blind Search):**  
   AI khám phá không gian trạng thái chỉ dựa trên cấu trúc đồ thị mà nó phát hiện trong quá trình duyệt. Hai thuật toán tìm kiếm không thông tin phổ biến là:
   * **Tìm kiếm theo chiều rộng (Breadth-First Search - BFS):**
     + Bắt đầu từ gốc đồ thị, khám phá tất cả các nút liên kết với gốc.
     + Sau đó, chuyển sang cấp độ tiếp theo, lần lượt khám phá từng cấp độ cho đến hết.
     + Ví dụ: Trong đồ thị, BFS khám phá từ A đến B và C trước khi chuyển sang D.
   * **Tìm kiếm theo chiều sâu (Depth-First Search - DFS):**
     + Bắt đầu từ gốc đồ thị, khám phá mọi nút theo một đường đi từ gốc đến cuối.
     + Quay lui để khám phá các đường đi chưa được thăm cho đến khi quay lại gốc. Nếu còn đường đi khác từ gốc, thuật toán tiếp tục quá trình tìm kiếm theo cách tương tự.
     + DFS khám phá từng đường đi hoàn chỉnh trước khi chuyển sang đường khác.

**Chơi trò chơi đối kháng**

**Trò chơi đối kháng:** Là các trò chơi trong đó một người thắng và người kia thua, hoặc bất kỳ tình huống nào mà các người chơi có mục tiêu xung đột nhau.

* **Ví dụ:** Trò chơi cờ caro (tic-tac-toe).

**Cờ caro:**

* Bảng 9 ô, người đầu tiên đặt được 3 dấu (X hoặc O) liên tiếp (ngang, dọc, chéo) là người thắng.
* Khi xây dựng cây trạng thái cho trò chơi:
  + **Mỗi cấp độ trong cây:** Đại diện cho một lượt chơi.
  + **Nút cuối cùng:** Đại diện cho trạng thái cuối cùng của bảng và quyết định thắng, hòa, hoặc thua.

**Chiến lược Min-Max**

**Min-Max Approximation:**

* Được giới thiệu bởi Ronald Rivest, MIT năm 1987.
* Tham khảo thêm bài báo: [Ronald Rivest - Min-Max Approximation](https://people.csail.mit.edu/rivest/pubs/Riv87c.pdf).

**Cách hoạt động:**

1. **Nút cuối cùng:**
   * Điểm cao nhất cho trạng thái thắng.
   * Điểm thấp hơn cho hòa.
   * Điểm thấp hoặc âm cho trạng thái thua.
2. **Lan truyền điểm:**
   * AI lan truyền điểm từ các nút cuối lên các nhánh và nút gốc bằng cách cộng dồn.
3. **Lượt của AI:**
   * AI chọn nhánh có giá trị cao nhất (tăng cơ hội thắng).
4. **Lượt của đối thủ:**
   * Giá trị thấp nhất được chọn (giảm cơ hội thắng của AI).

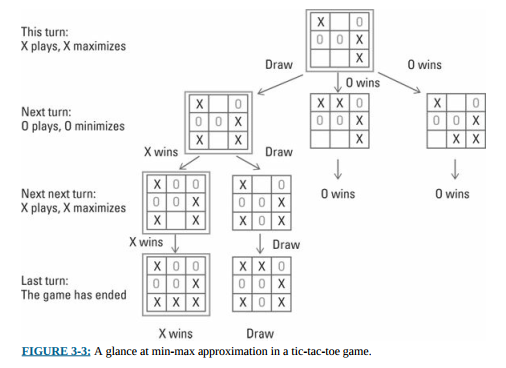
**Áp dụng chiến lược**

**Trò chơi hiện đại:**

* **AlphaGo của Google DeepMind:**
  + Sử dụng cách tiếp cận tương tự chiến lược Min-Max để chơi cờ vây.
  + Được xây dựng trên các thuật toán học sâu (Deep Learning).

**Ví dụ minh họa:**

* Hình 3-3 cung cấp minh họa trực quan về chiến lược Min-Max trong trò chơi đối kháng.



### Sử dụng tìm kiếm cục bộ và heuristics

**Giới hạn của tìm kiếm trạng thái toàn diện**

Không có máy móc nào, dù mạnh đến đâu, có thể liệt kê hết các khả năng phát sinh từ một tình huống phức tạp.

* **Ví dụ:**
  + Cờ vua và cờ vây có các quy tắc cố định nhưng số lượng vị trí bàn cờ rất lớn.
  + Thậm chí, cờ checkers — trò chơi đơn giản hơn — có **500 tỷ tỷ** (500,000,000,000,000,000,000) vị trí bàn cờ khả thi.

Theo tính toán của các nhà toán học tại Đại học Hawaii, để tính toàn bộ các nước đi có thể mất tới **18 năm** với máy tính mạnh. Vì vậy, việc giới hạn số nước đi cần đánh giá là rất cần thiết.

**Tối ưu hóa với tìm kiếm cục bộ và heuristics**

Tìm kiếm cục bộ và heuristics giúp giảm số lượng đánh giá ban đầu bằng cách:

1. Sử dụng **ràng buộc (constraints)** để giới hạn nước đi cần kiểm tra, như trong **cắt tỉa alpha** (bỏ qua những nước đi không đóng góp vào thành công của tìm kiếm).
2. Áp dụng **thuật toán tìm kiếm cục bộ (local search)**:
   * Xuất phát từ một giải pháp hiện tại (dù chưa hoàn hảo) và cải thiện dần dần.
   * Tập trung đánh giá các giải pháp gần kề (neighborhood solutions).
   * Có thể dựa vào ngẫu nhiên hoặc heuristics để tìm hướng đi đúng.

**Quy trình cơ bản của tìm kiếm cục bộ**

1. **Bắt đầu từ một tình huống hiện tại:**
   * Đó có thể là trạng thái hiện tại, một giải pháp ngẫu nhiên, hoặc giải pháp đã biết.
2. **Tìm kiếm các giải pháp mới gần kề:**
   * Tạo danh sách các ứng viên từ các giải pháp gần kề.
3. **Chọn giải pháp tối ưu:**
   * Dựa vào một heuristic để quyết định giải pháp nào tốt nhất từ danh sách ứng viên.
4. **Tiếp tục tìm kiếm:**
   * Lặp lại bước 2 và 3 cho đến khi không thể cải thiện giải pháp thêm nữa.

**Ưu và nhược điểm của tìm kiếm cục bộ:**

* **Ưu điểm:** Đơn giản, dễ thiết kế.
* **Nhược điểm:** Có thể không tìm được giải pháp trong thời gian hợp lý hoặc không đạt chất lượng mong muốn.

**Một số phương pháp heuristic phổ biến**

1. **Hill Climbing:**
   * Mô phỏng việc quả bóng lăn xuống dốc hoặc leo lên đỉnh núi.
   * Heuristic chỉ ra hướng đi có độ dốc cao nhất (leo lên) hoặc thấp nhất (xuống dốc).
   * **Nhược điểm:** Có thể bị mắc kẹt ở điểm cực đại cục bộ hoặc vùng phẳng không có tiến triển (plateau).
2. **Twiddle (Coordinate Descent):**
   * Giống hill climbing nhưng kiểm tra tất cả các hướng.
   * Tập trung vào hướng hiệu quả nhất, giảm dần bước di chuyển khi việc cải thiện trở nên khó khăn.
3. **Simulated Annealing:**
   * Dựa trên kỹ thuật luyện kim làm nóng và làm nguội kim loại để cải thiện cấu trúc.
   * Cho phép chọn các giải pháp kém hơn trong ngắn hạn để đạt mục tiêu dài hạn tốt hơn.
   * Điều chỉnh dần để tìm giải pháp tối ưu.
4. **Taboo Search:**
   * Ghi nhớ các vùng đã khám phá để tránh lặp lại.
   * Khi đạt giải pháp tiềm năng, thuật toán có thể quay lại để kiểm tra các hướng chưa thử.

**Ứng dụng tìm kiếm cục bộ và heuristics**

1. **Điều hướng robot:**
   * Robot phải tìm đường đi ngắn nhất đến mục tiêu, tránh chướng ngại vật và giảm thiểu chuyển động không cần thiết.
   * Dựa vào **cảm biến lidar hoặc sonar** để hiểu môi trường xung quanh.
2. **Tìm đường (Pathfinding):**
   * Áp dụng cho robot, trò chơi điện tử, hoặc các mục tiêu ảo.
   * **Mục tiêu:** Tìm đường ngắn nhất từ vị trí hiện tại đến đích, tránh chướng ngại vật.

**Vai trò của heuristic:**

* Giúp robot định hướng, ví dụ: Biết rằng mục tiêu nằm về hướng bắc.
* Tăng hiệu quả mà không cần xác định tất cả các phương án.

## Khám phá cách hoạt động của hệ thống chuyên gia

Các ví dụ thuật toán trước đây liên quan đến AI bởi chúng là những giải pháp thông minh giải quyết các vấn đề **lặp lại, có giới hạn nhưng phức tạp**, đòi hỏi trí thông minh.

* **Vai trò của kiến trúc sư:** Cần có người thiết kế nghiên cứu vấn đề và chọn thuật toán phù hợp để giải quyết.
* **Hạn chế:** Khi vấn đề thay đổi, biến đổi, hoặc hiển thị các đặc điểm không quen thuộc, thuật toán có thể gặp khó khăn trong việc thực thi.

**Ví dụ về AI giải quyết vấn đề cố định**

* Một thuật toán AI có thể được lập trình để **giải Sudoku** (trò chơi yêu cầu sắp xếp các số lên bảng theo một số quy tắc nhất định).
* **Peter Norvig**, Giám đốc nghiên cứu tại Google, đã viết một bài tiểu luận thú vị về việc sử dụng **DFS (Depth-First Search)**, giới hạn số lượng tính toán, áp dụng các ràng buộc và ưu tiên các nhánh nhỏ trước để làm cho giải pháp Sudoku khả thi.
  + **Tham khảo:** [Norvig's Sudoku Essay](http://norvig.com/sudoku.html).

**Tuy nhiên:** Không phải tất cả các vấn đề đều có thể giải quyết giống như Sudoku.

* **Vấn đề thực tế** không tồn tại trong các môi trường thông tin hoàn hảo và hành động được xác định rõ.
* **Ví dụ:**
  + **Phát hiện gian lận:** Số lượng gian lận có thể xảy ra là rất lớn, và kẻ gian thường che giấu thông tin.
  + **Chẩn đoán bệnh:** Nhiều bệnh có triệu chứng tương tự, và thông tin từ các xét nghiệm có thể không đầy đủ.

**Khó khăn của các vấn đề thực tế**

1. **Tập hợp lớn các quy tắc và khả năng:**
   * Số lượng gian lận hoặc bệnh tật là rất lớn.
2. **Thông tin bị thiếu:**
   * Gian lận thường che giấu thông tin.
   * Bác sĩ có thể phải chẩn đoán dựa trên thông tin không đầy đủ.
3. **Quy tắc không bất biến:**
   * Kẻ gian tìm ra cách gian lận mới.
   * Các bệnh mới xuất hiện hoặc được phát hiện.

Để giải quyết các vấn đề này, cần có một **cách tiếp cận linh hoạt** và tích lũy kiến thức hữu ích nhằm đối mặt với những thách thức mới. Nói cách khác, cần có **quá trình học tập liên tục**, giống như cách con người học để thích nghi với môi trường thay đổi.

### Sử dụng hệ thống chuyên gia (Expert Systems)

**Hệ thống chuyên gia** là một trong những nỗ lực đầu tiên để thoát khỏi lĩnh vực thuật toán được mã hóa cứng, tạo ra các cách thông minh và linh hoạt hơn để giải quyết các vấn đề thực tế.

1. **Cốt lõi của hệ thống chuyên gia:**
   * Dựa vào **các quy tắc** để đưa ra quyết định.
   * Phù hợp với thời kỳ mà việc lưu trữ và xử lý lượng lớn dữ liệu trong bộ nhớ máy tính còn tốn kém.
2. **Hạn chế của thập niên 1970:**
   * Bộ nhớ máy tính hạn chế.
   * **Ví dụ:** Nhà khoa học AI Ross Quillian phải xây dựng mô hình ngôn ngữ hoạt động chỉ dựa trên từ vựng gồm **20 từ** vì bộ nhớ máy tính chỉ chứa được bấy nhiêu.
3. **Giải pháp thời đó:**
   * Chỉ tập trung vào thông tin quan trọng của vấn đề.
   * Lấy thông tin từ **con người có chuyên môn**, những người biết rõ nhất về vấn đề.

### MYCIN: Một hệ thống chuyên gia khởi đầu

**MYCIN** là một trong những hệ thống chuyên gia đầu tiên được phát triển để chẩn đoán các bệnh liên quan đến đông máu hoặc nhiễm trùng do vi khuẩn, chẳng hạn như:

* **Bacteremia:** Nhiễm khuẩn máu.
* **Meningitis:** Viêm màng não.

**Cách hoạt động của MYCIN**

* **Chức năng chính:**  
  MYCIN khuyến nghị **liều lượng kháng sinh phù hợp** bằng cách sử dụng hơn **500 quy tắc.**
* **Tương tác với bác sĩ:**  
  Khi không đủ thông tin (ví dụ, thiếu kết quả xét nghiệm), MYCIN bắt đầu một cuộc đối thoại tham vấn bằng cách đặt các câu hỏi liên quan để đạt được **chẩn đoán và liệu pháp tự tin hơn.**

**Lịch sử và hiệu quả của MYCIN**

* **Người phát triển:**  
  Edward Shortliffe đã phát triển MYCIN như một phần luận án tiến sĩ tại Đại học Stanford.
* **Thời gian hoàn thiện:** Hơn 5 năm.
* **Hiệu quả:**
  + MYCIN có độ chính xác chẩn đoán cao hơn bất kỳ bác sĩ trẻ nào.
  + Hiệu suất tương đương với bác sĩ giàu kinh nghiệm.
* **Nguồn gốc:**  
  MYCIN được phát triển từ cùng phòng thí nghiệm đã tạo ra **DENDRAL**, hệ thống chuyên gia đầu tiên, chuyên về **hóa học hữu cơ.**

**DENDRAL:** Một ứng dụng thách thức, trong đó các thuật toán mạnh mẽ (brute-force) tỏ ra không khả thi, trong khi các **heuristics dựa trên kinh nghiệm** lại hiệu quả.

**Các vấn đề của MYCIN**

1. **Trách nhiệm pháp lý:**
   * Nếu hệ thống đưa ra chẩn đoán sai, ai sẽ chịu trách nhiệm?
2. **Khả năng sử dụng:**
   * Bác sĩ phải kết nối với MYCIN qua một **thiết bị đầu cuối từ xa (remote terminal)** đến máy chủ tại Stanford.
   * Điều này rất khó khăn và chậm chạp vào thời điểm Internet vẫn đang ở giai đoạn sơ khai.

**Di sản của MYCIN**

Dù gặp vấn đề, MYCIN chứng minh được:

* **Hiệu quả và tính hữu ích trong hỗ trợ quyết định của con người.**
* **Mở đường** cho nhiều hệ thống chuyên gia khác phát triển trong những năm 1970 và 1980.

**Thành phần của hệ thống chuyên gia**

Các hệ thống chuyên gia thời đó thường bao gồm **hai thành phần chính:**

1. **Cơ sở tri thức (Knowledge base):**
   * Lưu trữ kiến thức dưới dạng **các quy tắc if-then.**
   * **If:** Một hoặc nhiều điều kiện.
   * **Then:** Các kết luận.
2. **Động cơ suy luận (Inference engine):**
   * Sử dụng **logic Boolean** hoặc **logic bậc nhất phức tạp** (first-order logic) để thao tác các quy tắc, xử lý sự kiện, lớp, và các lớp con.

**Inference Engine và Quy trình Hoạt động trong Hệ thống Chuyên gia**

**Inference Engine** là một tập hợp các chỉ dẫn, cho phép hệ thống thao tác các điều kiện dựa trên tập hợp logic Boolean, bao gồm các toán tử như **AND, OR, NOT**.

* **Hoạt động:**
  + Kết hợp các điều kiện biểu tượng **TRUE** hoặc **FALSE** thành các lập luận phức tạp.
  + Khi điều kiện **TRUE**, một quy tắc được kích hoạt (hoặc "fired").
  + Khi điều kiện **FALSE**, quy tắc không áp dụng.

Hệ thống hoạt động dựa trên một loạt các câu lệnh **nếu-thì** (if-then), được sắp xếp theo lớp. Việc thu thập thông tin ban đầu giúp loại bỏ các kết luận không đúng và hướng hệ thống đến các câu hỏi hoặc thông tin có thể dẫn đến câu trả lời.

**Các Phương pháp Hoạt động Chính của Hệ thống Chuyên gia**

1. **Forward Chaining (Xâu chuỗi tiến):**
   * Dựa trên **dữ liệu hiện có**, hệ thống kích hoạt một loạt quy tắc và loại bỏ các quy tắc không liên quan.
   * **Tập trung:** Vào các quy tắc có thể dẫn đến kết luận cuối cùng.
   * **Đặc điểm:** Hướng dữ liệu.
2. **Backward Chaining (Xâu chuỗi ngược):**
   * Hệ thống **đánh giá từng kết luận khả thi** và cố gắng chứng minh nó dựa trên bằng chứng có sẵn.
   * **Tập trung:** Hướng mục tiêu; xác định câu hỏi cần đặt ra và loại bỏ các mục tiêu không liên quan.
   * **Ứng dụng:**
     + **MYCIN**, một hệ thống được sử dụng trong chẩn đoán y khoa, sử dụng phương pháp này.
     + Tiến hành từ giả thuyết ngược trở lại bằng chứng, một chiến lược phổ biến trong y học.
3. **Conflict Resolution (Giải quyết xung đột):**
   * Khi hệ thống đưa ra nhiều kết luận cùng lúc, nó ưu tiên kết luận dựa trên:
     + **Tính tác động.**
     + **Nguy cơ.**
     + **Yếu tố khác.**
   * **Ví dụ:**
     + **MYCIN** sử dụng **hệ số chắc chắn** (certainty factor) để ước tính xác suất chẩn đoán chính xác.
     + Hệ thống có thể tham khảo ý kiến người dùng để giải quyết xung đột dựa trên đánh giá của họ.

**Ưu Điểm của Hệ Thống Chuyên Gia**

1. **Biểu diễn kiến thức dễ đọc:**
   * Quy trình ra quyết định minh bạch.
   * Hệ thống có thể trả lại các quy tắc mà nó sử dụng để đạt đến kết luận, cho phép người dùng kiểm tra và sửa lỗi đầu vào nếu cần.
2. **Lập trình dễ dàng:**
   * Sử dụng các ngôn ngữ như **Lisp**, **Prolog**, hoặc **ALGOL**.
3. **Khả năng cải thiện theo thời gian:**
   * Người dùng có thể thêm quy tắc mới hoặc cập nhật quy tắc hiện có.
   * Có thể áp dụng **logic mờ (fuzzy logic)** để xử lý các điều kiện không chắc chắn.

**Logic Mờ (Fuzzy Logic) và Ví Dụ**

**Logic mờ:** Một loại logic đa giá trị, trong đó giá trị có thể nằm giữa **0 (hoàn toàn sai)** và **1 (hoàn toàn đúng)**.

* **Ứng dụng:**
  + Ví dụ: Quy tắc kích hoạt khi phòng "nóng" không phải ở một nhiệt độ chính xác, mà ở khoảng nhiệt độ xung quanh ngưỡng đó.
  + Tham khảo thêm: [“Artificial Intelligence — Fuzzy Logic Systems”](https://www.tutorialspoint.com) trên **TutorialsPoint.com**.

**Sự Suy Thoái của Hệ Thống Chuyên Gia**

Hệ thống chuyên gia bắt đầu suy thoái vào cuối những năm 1980 vì các lý do sau:

1. **Hạn chế về logic và biểu tượng:**
   * Không thể biểu đạt đầy đủ các quy tắc đằng sau một quyết định phức tạp, dẫn đến việc phải quay lại mã hóa thủ công bằng các thuật toán cổ điển.
2. **Độ phức tạp tăng cao:**
   * Hệ thống trở nên quá phức tạp, mất đi tính khả thi về kinh tế và kỹ thuật.
3. **Dữ liệu ngày càng phổ biến:**
   * Không còn cần thiết phải phỏng vấn và thu thập kiến thức chuyên gia hiếm hoi một cách cẩn thận khi có thể **khai thác kiến thức từ dữ liệu.**

**Ứng Dụng Hiện Đại của Hệ Thống Chuyên Gia**

Dù suy thoái, hệ thống chuyên gia vẫn tồn tại trong một số lĩnh vực:

1. **Chấm điểm tín dụng:**
   * Cung cấp quyết định rõ ràng và minh bạch.
2. **Phát hiện gian lận:**
   * Yêu cầu hệ thống giải thích được các quyết định đưa ra.
3. **Ứng dụng trong xe tự lái:**
   * Một số ứng dụng xe tự lái vẫn sử dụng hệ thống chuyên gia vì các phương pháp AI khác có thể chậm hơn.
   * Tham khảo thêm: [“Expert Systems and AI Self-Driving Cars: Crucial Innovative Techniques”](https://www.aitrends.com).

## Khám phá về Machine Learning (máy tự học)

Các giải pháp có khả năng **học trực tiếp từ dữ liệu mà không cần xử lý trước để biểu diễn thành biểu tượng** đã xuất hiện từ vài thập kỷ trước các hệ thống chuyên gia (expert systems).

* Một số phương pháp có tính chất thống kê, số khác bắt chước tự nhiên theo nhiều cách khác nhau, và một số khác nữa cố gắng tạo ra logic biểu tượng tự động dưới dạng các quy tắc từ dữ liệu thô.
* Tất cả những giải pháp này xuất phát từ các trường phái khác nhau và hiện nay được gọi chung là **machine learning (sự học của máy).**

**Machine Learning khác biệt như thế nào?**

**Học máy** thuộc thế giới thuật toán, nhưng khác với các thuật toán truyền thống đã được thảo luận, nó:

* **Không được thiết kế dưới dạng các bước xác định trước** để giải quyết một vấn đề cụ thể.
* Thay vào đó, học máy giải quyết những vấn đề mà con người không biết cách chi tiết hóa thành các bước nhưng con người lại giải quyết một cách tự nhiên.

**Ví dụ:**

* Nhận diện khuôn mặt trong hình ảnh.
* Nhận diện một số từ trong một cuộc trò chuyện.

### Machine Learning trong cuốn sách này

Học máy được nhắc đến trong hầu hết các chương, nhưng **Chương 9 đến Chương 11** tập trung giải thích cách hoạt động của các thuật toán học máy chính, đặc biệt là **học sâu (deep learning)**.

* Học sâu là công nghệ đứng sau **làn sóng ứng dụng AI mới**, xuất hiện trên các tiêu đề tin tức gần như hàng ngày.

**Chạm đến những đỉnh cao mới**

Vai trò của học máy trong làn sóng thuật toán AI mới là **thay thế một phần, bổ sung một phần** cho các thuật toán hiện có.

**Hoạt động của học máy:**

* Làm việc với các hoạt động đòi hỏi **trí tuệ từ góc nhìn con người**, nhưng không dễ dàng được hình thức hóa thành các bước chính xác.

**Ví dụ:**

* Trong cờ vây (**Go**), các chuyên gia có thể hiểu nhanh chóng các cơ hội và mối đe dọa trên bàn cờ và trực giác tìm ra nước đi đúng.

**Cờ vây - Thách thức lớn với AI**

**So sánh:**

* **Cờ vua:** Trung bình có 35 nước đi cần đánh giá trên một bàn cờ, với một ván cờ thường kéo dài hơn 80 nước đi.
* **Cờ vây:** Trung bình có 140 nước đi cần đánh giá, và một ván thường kéo dài hơn 240 nước đi.

**Hiện nay, không có sức mạnh tính toán nào trên thế giới có thể tạo ra không gian trạng thái đầy đủ cho một ván cờ vây.**

**AlphaGo - Sự thống trị của AI**

**Google DeepMind** đã phát triển chương trình **AlphaGo**, đánh bại nhiều người chơi cờ vây hàng đầu.  
**Tìm hiểu thêm tại:**

* [DeepMind’s AlphaGo](https://deepmind.com/research/alphago/).
* [DeepMind Gaming AI Dominance](https://www.kdnuggets.com/2020/05/deepmind-gaming-ai-dominance.html).

**Cách tiếp cận của AlphaGo:**

1. **Phương pháp tìm kiếm thông minh:**
   * Sử dụng các kiểm tra ngẫu nhiên (random tests) để đánh giá một nước đi có kết quả tích cực hay tiêu cực.
   * Áp dụng tìm kiếm sâu ưu tiên (DFS) nhiều lần để xác định kết quả.
2. **Thuật toán học sâu:**
   * **Mạng chính sách (policy network):** Phân tích hình ảnh bàn cờ để xác định nước đi tốt nhất trong tình huống đó.
   * **Mạng giá trị (value network):** Ước tính khả năng chiến thắng của AI nếu sử dụng nước đi đó.
3. **Học qua quan sát và tự chơi:**
   * **Học từ các ván đấu của chuyên gia cờ vây.**
   * **Học qua tự chơi:** Phiên bản **AlphaGo Zero** tự học từ đầu, không cần ví dụ của con người.
   * **Tham khảo:** [AlphaGo Zero Learning from Scratch](https://deepmind.com/blog/alphago-zero-learning-scratch/).

### Học tăng cường (Reinforcement Learning)

Khả năng học qua việc tự chơi của AlphaGo được gọi là **học tăng cường (reinforcement learning).**

* Đây là một bước tiến quan trọng trong việc phát triển các thuật toán học sâu tự học mà không cần dữ liệu gốc từ con người.

# Chương 4: Tiên phong phần cứng chuyên dụng

**TRONG CHƯƠNG NÀY**

* Sử dụng phần cứng tiêu chuẩn.
* Sử dụng phần cứng chuyên dụng.
* Cải thiện phần cứng.
* Tương tác với môi trường.

**Bối cảnh lịch sử và thách thức về phần cứng AI**

Trong **Chương 1**, bạn đã tìm hiểu rằng một trong những lý do khiến các nỗ lực AI ban đầu thất bại là do **thiếu phần cứng phù hợp.**

* Phần cứng lúc đó không thể thực hiện các tác vụ đủ nhanh, ngay cả đối với những nhu cầu cơ bản, chứ chưa nói đến những điều phức tạp như mô phỏng tư duy con người.

**Ví dụ lịch sử:**

* Bộ phim **The Imitation Game** mô tả cách **Alan Turing** phá mã Enigma.
* Ông đã khéo léo tìm kiếm một cụm từ cụ thể là **"Heil Hitler"** trong mỗi thông điệp.
* Nếu không có lỗ hổng đó trong cách các nhà điều hành sử dụng máy Enigma, thiết bị mà Turing sử dụng sẽ không đủ nhanh để giải quyết vấn đề.

**Điều đáng chú ý:**

* Các tài liệu lịch sử được giải mật cho thấy vấn đề mà Turing phải đối mặt còn nghiêm trọng hơn những gì bộ phim thể hiện.  
  **Tham khảo:**
  + [“Cracking the Uncrackable”](https://www.scienceabc.com) trên **ScienceABC.com**.
  + [Mô tả phim The Imitation Game](https://www.amazon.com) trên **Amazon.com**.

**Cải tiến phần cứng hiện nay**

May mắn thay, ngày nay, phần cứng tiêu chuẩn có sẵn trên thị trường đã vượt qua các vấn đề về tốc độ đối với nhiều bài toán.

* Chương này sẽ bắt đầu bằng cách thảo luận cách phần cứng tiêu chuẩn có thể giải quyết được vấn đề hiệu suất và tiếp tục khám phá vai trò của phần cứng chuyên dụng trong việc nâng cao khả năng của AI.