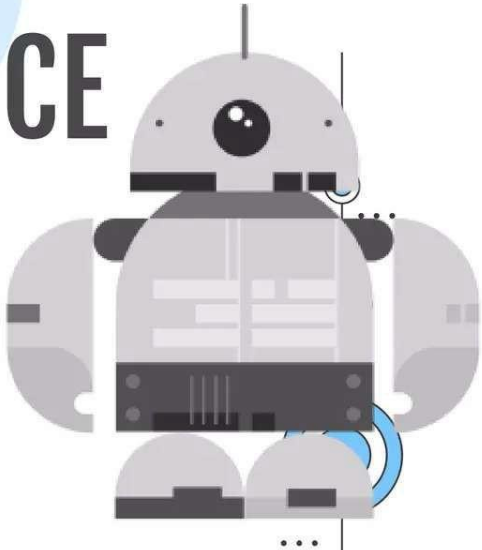
HISTORY
OF ARTIFICIAL INTELLIGENCE



A Journey Through the Evolution of Artificial Intelligence

1. Introduction and inception of AI
2. Early enthusiasm, great expectations
3. A dose of reality
4. Expert systems
5. Return of neural networks
6. Probabilistic reasoning and ML
7. Big data and deep learning

1. Introduction and inception of AI

The first work that is now generally recognized as AI was done by Warren McCulloch and Walter Pitts (1943). Inspired by the mathematical modeling work of Pitts's advisor Nicolas

Alan Turing's gave lectures on the topic as early as 1947. He introduced:

- ➔ Turing test
- ➔ Machine Learning
- ➔ Genetic algorithms

1943: Birth of Neural Networks

Warren McCulloch and Walter Pitts introduced the first mathematical model of neural networks in their paper "A Logical Calculus of the Ideas Immanent in Nervous Activity."

This laid the foundation for artificial neurons and showed that the brain could be modeled using logic circuits.

1950: Turing Test and Machine Intelligence

Alan Turing published "Computing Machinery and Intelligence," proposing the Turing Test, a way to measure machine intelligence by determining if a computer could mimic human responses.

1951: First AI Program

Christopher Strachey developed a checkers-playing program for the Ferranti Mark I, one of the earliest stored-program computers.

1956: The Birth of AI – Dartmouth Conference

The Dartmouth Summer Research Project on Artificial Intelligence, organized by John McCarthy, Marvin Minsky, Nathaniel Rochester, and Claude Shannon, officially coined the term "Artificial Intelligence" and established AI as a formal field of study.

1957-1958: Perceptron and AI Expansion

Frank Rosenblatt developed the Perceptron, an early neural network model capable of learning from data.

John McCarthy created LISP, the first AI programming language, which became widely used in AI research.

2. Early enthusiasm, great expectation

1952 - 1969
AI researchers focused in particular on tasks considered indicative of intelligence in humans, including games, puzzles, mathematics, and IQ tests

In 1958, John McCarthy defined the high-level language Lisp.
Stressed representation and reasoning in formal logic.
Applications of logic included:

- ➔ Cordell Green's question-answering and planning systems
- ➔ The Shakey robotics project at the Stanford Research Institute (SRI).

1961: Unimate – The First Industrial Robot

Unimate, developed by George Devol and Joseph Engelberger, became the first industrial robot.

Installed at a General Motors factory, it performed repetitive tasks like moving hot metal parts, revolutionizing automation.

1964-1966: ELIZA – The First Chatbot

ELIZA, developed by Joseph Weizenbaum at MIT, became the first chatbot.

It simulated a Rogerian psychotherapist, responding to users by recognizing patterns in their text input and replying accordingly.

ELIZA showed how simple natural language processing (NLP) could create the illusion of human-like conversation.

1969: Shakey the Robot – AI and Robotics

Shakey, developed by SRI International, became the first AI-powered robot that could move, perceive its environment, and make decisions.

It combined computer vision, natural language processing, and problem-solving, setting the stage for modern robotic

3. A dose of Reality(1966-1973)

Early translation relied on simple word replacement and grammar rules.
Lacked contextual understanding, leading to humorous mistranslations.

U.S. government stopped funding AI translation projects in 1966.
Early AI programs worked in simple “microworlds” but failed with real-world complexity.

A Dose of Reality (1966-1973)

Challenges and Setbacks in Early AI Research

Overconfidence in Early AI

Herbert Simon predicted AI breakthroughs within 10 years, but actual progress took decades.

Computational Complexity

Early AI underestimated the difficulty of scaling up due to combinatorial explosion.

Flawed Assumptions

AI relied on human introspection rather than rigorous problem analysis, leading to unrealistic expectations.

Disillusionment and Funding Cuts

The Lighthill Report (1973) led the British government to reduce AI funding significantly.

Many problems required testing too many possibilities, making solutions impractical.

Lighthill Report (1973) criticized AI's scalability, leading to reduced UK funding.

Perceptrons (early neural networks) had major limitations, in that what they could represent was limited as shown by Minsky & Papert (1969).

Funding for neural networks declined, despite key techniques (like backpropagation) already existing.

4. Expert Systems in AI (1969 - 1986)

What are they?
Computer programs designed to mimic human experts in decision making. Used rules and knowledge to solve problems

How they worked:

- ➔ Knowledge base: stored facts and rules
- ➔ Inference engine: Applied logic to draw conclusions.
- ➔ User interface: Allowed interaction with users

Key developments:

DENDRAL(1969): Analyzed molecular structures for chemists.

MYCIN(1970s): Diagnosed bacterial infections in medicine.

XCON(1980s): Helped configure computer systems for DEC.

Impacts and limitations:

Success: Improved efficiency in medicine, business, and engineering.

Challenges: Required manual updates, couldn’t adapt to new situations.

Decline & Evolution

By the late 1980s, expert systems were replaced by machine learning and more advanced AI techniques.

5. Return of Neural networks

1986 - Present
A neural network is a computational model inspired by the human brain consisting of layers of neurons (nodes) connected by weighted links.

Back-propagation is a learning algorithm for training neural networks by adjusting connection weights using gradient descent to minimize errors.

Connectionist Models vs. Symbolic AI

In the mid-1980s at least four different groups reinvented the back-propagation learning algorithm first found in 1969 by Bryson and Ho.

These neural network-based models, called connectionist models, were seen by some as competitors to the symbolic AI. Symbolic AI emphasizes rule-based, logical reasoning. Connectionism on the other hand focuses on learning from data by adjusting connections in a neural network.

The modern perspective is that symbolic and connectionist approaches complement each other rather than compete.

Branches of Neural Network Research

Over time, neural network research has split into two major fields:

Algorithmic and mathematical study – Focuses on improving neural network architectures and understanding their properties.

Biological modeling – Aims to replicate the empirical properties of real neurons and brain structures.

6. Probabilistic reasoning & ML

Probabilistic reasoning refers to the use of probability theory to model and manage uncertainty in decision making processes.

Machine Learning(ML) is a type of AI that teaches machines to imitate the way humans learn, to perform tasks autonomously.

Brittleness of expert systems led to a new, more scientific approach. This incorporated:

- ➔ Probability rather than boolean logic
- ➔ Machine learning rather than hard-coding
- ➔ Experimental results rather than philosophical claims

Shared benchmark problem sets became the norm for demonstrating progress, including:

- ➔ the UC Irvine repository for machine learning data sets,
- ➔ the LibriSpeech corpus for speech recognition,
- ➔ the MNIST data set for handwritten digit recognition.

In the 1980s, approaches using hidden Markov models (HMMs) came to dominate the area.

Two relevant aspects of HMMs:

- ➔ Are based on rigorous mathematical theory
- ➔ Are generated by a process of training on a large corpus of reel speech data

1988 was an important year for the connection between AI and other fields, statistics, operations research, decision theory, and control theory.

7. Big data and Deep learning

Big data, is a phenomenon created by the advancement of world wide web (www).
WWW has facilitated the creation of large datasets, words(text), images, audio and video

Big data acts as fuel while AI acts as the engine enabling the analysis of huge datasets.

Deep Learning refers to machine learning using multiple layers of simple adjustable computing elements

Experiments dating as far back as the 1970s have been carried out in various networks and specifically in the form of convolutional neural networks.
In the 1990s, successes in handwritten digit recognition were found.
It was in 2011 that deep learning methods took of.
The process occurring first in speech recognition followed by visual object recognition.
Deep learning systems have exceeded human performance on some vision tasks.

LARGE LANGUAGE MODELS

LLMS:

GPT (OpenAI) – Released: 2020 (GPT-3), 2023 (GPT-4)

Gemini (Google DeepMind) – Released: 2023

Claude (Anthropic) – Released: 2023

Mistral 7B (Mistral AI) – Released: 2023

DeepSeek (DeepSeek AI) – Released: 2023

Qwen (Alibaba Cloud) – Released: 2023

Perplexity AI – Released: 2022

Grok (xAI) – Released: 2023