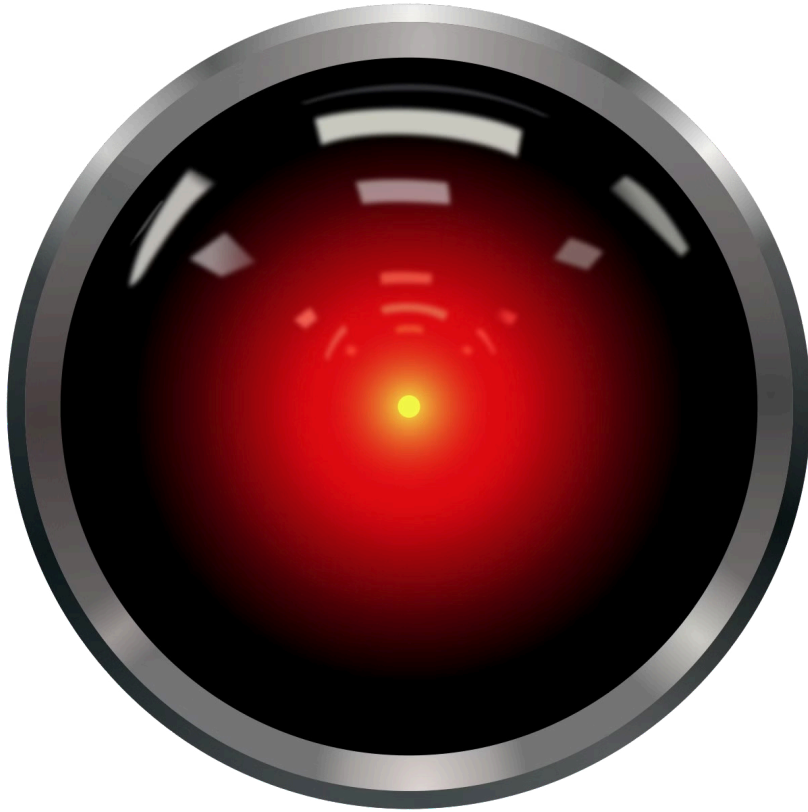


# Introduction to Artificial Intelligence and NLP



**Prof. Alfio Ferrara**

University of Milan

*sed noli modo*



---

**Natural Language Processing**

Master Degree in Computer Science and Data Science for Economics

# NLP course: learning materials



## Course Lectures

Classes feature extensive examples and Python notebooks. While attendance is not mandatory, it is strongly recommended.



## Course Materials

Slides, handouts, and Python notebooks are progressively published on both platforms.



## Access Platforms

Materials available on:

- Ariel platform:

<https://aferrarair.ariel.ctu.unimi.it>

- GitHub repository:

<https://github.com/afflint/nlp>

# NLP course: assessment methods

## Project Development

Select and discuss your project topic with the lecturer. Your project must demonstrate comprehension of lecture topics and propose innovative solutions to specific research problems.

## SIFA Registration

Register for the examination through the SIFA service. This step is mandatory for all students.

## Schedule Discussion

After SIFA registration, contact the lecturer to arrange your project discussion time.

## Final Evaluation

Your assessment will be based on both the project outcomes and an interview discussing your work and related topics.

# Definition of Artificial Intelligence

The term Artificial Intelligence does not belong to technical terminology; it is a term that has essentially had a marketing objective since its inception.



# Technical Terminology of AI

## Expert Systems

Systems that emulate the decision-making process of a human expert.

## Knowledge-based Systems

Systems based on codified knowledge.

## Machine Learning

Automatic learning from data and experiences.

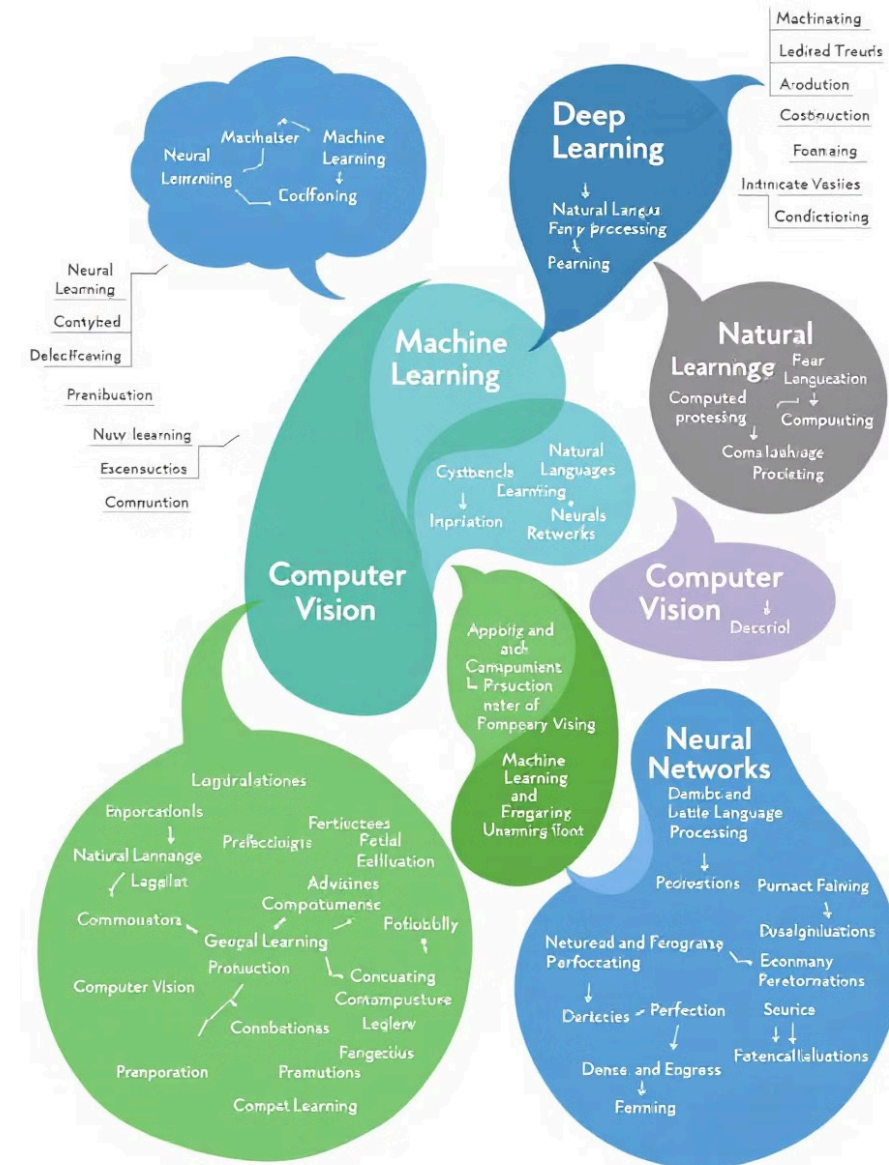
## NLP

Tools for manipulating natural language (Natural Language Processing)

## Computer Vision

Systems for manipulating images and visual language

## Etc...





# The Goal of AI

- 1 The goal of AI has never been to **replicate human intelligence**.
- 2 The goal is to **simulate tasks normally associated** with human intelligence.
- 3 To therefore perform tasks such as writing, drawing, interacting with others, etc. **through machines**.



## Let's play a game: which of those texts is the overview of a Spanish TV Series?

### 1 Text A

Victor is a new student at high school on his own journey of self-discovery, facing challenges at home, adjusting to a new city, and struggling with his sexual orientation. When it all seems too much, he reaches out to Simon to help him navigate the ups and downs of high school.

### 2 Text B

He's surly, unorthodox, unapologetically blunt, and he's about to change your life. Meet the new philosophy teacher, Merlí, who will help his students view the world in a whole new light, both in and out of the classroom.

### 3 Text C

The adventures of 13-year-old teenager, who is magically transported to the fictitious world of Amphibia, a rural marshland full of frog-people. With the help of an excitable young frog named Sprig, she will become a hero and discover the first true friendship of her life.

## Which of those texts is the (fake) overview of TV Series generated by ChatGPT?

### 1 Text A

Twin brother and sister  
Dipper and Mabel Pines are in  
for an unexpected adventure  
when they spend the summer  
helping their great uncle  
Stan run a tourist trap in the  
mysterious town of Gravity  
Falls, Oregon.

### 2 Text B

Alex, Justin and Max Russo  
are not your ordinary kids -  
they're wizards in training!  
While their parents run the  
Waverly Sub Station, the  
siblings struggle to balance  
their ordinary lives while  
learning to master their  
extraordinary powers.

### 3 Text C

In the quirky town of  
Evergreen, Colorado,  
energetic Emma and clever  
Jay find themselves  
navigating a world where  
magic and love collide in the  
most unexpected ways.



# The *Turing Test*, or *Imitation Game*

I PROPOSE to consider the question, '**Can machines think?**' This should begin with definitions of the meaning of the terms 'machine' and 'think'. The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words 'machine' and 'think' are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, 'Can machines think?' is to be sought in a statistical survey such as a Gallup poll. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.

**The new form of the problem can be described in terms of a game which we call the 'imitation game'.** It is played with three people, a man (A), a woman (B), and an interrogator (C) who may be of either sex. The interrogator stays in a room apart from the other two. **The object of the game for the interrogator is to determine which of the other two is the man and which is the woman.** He knows them by labels X and Y, and at the end of the game he says either 'X is A and Y is B' or 'X is B and Y is A'. The interrogator is allowed to put questions to A and B [...]

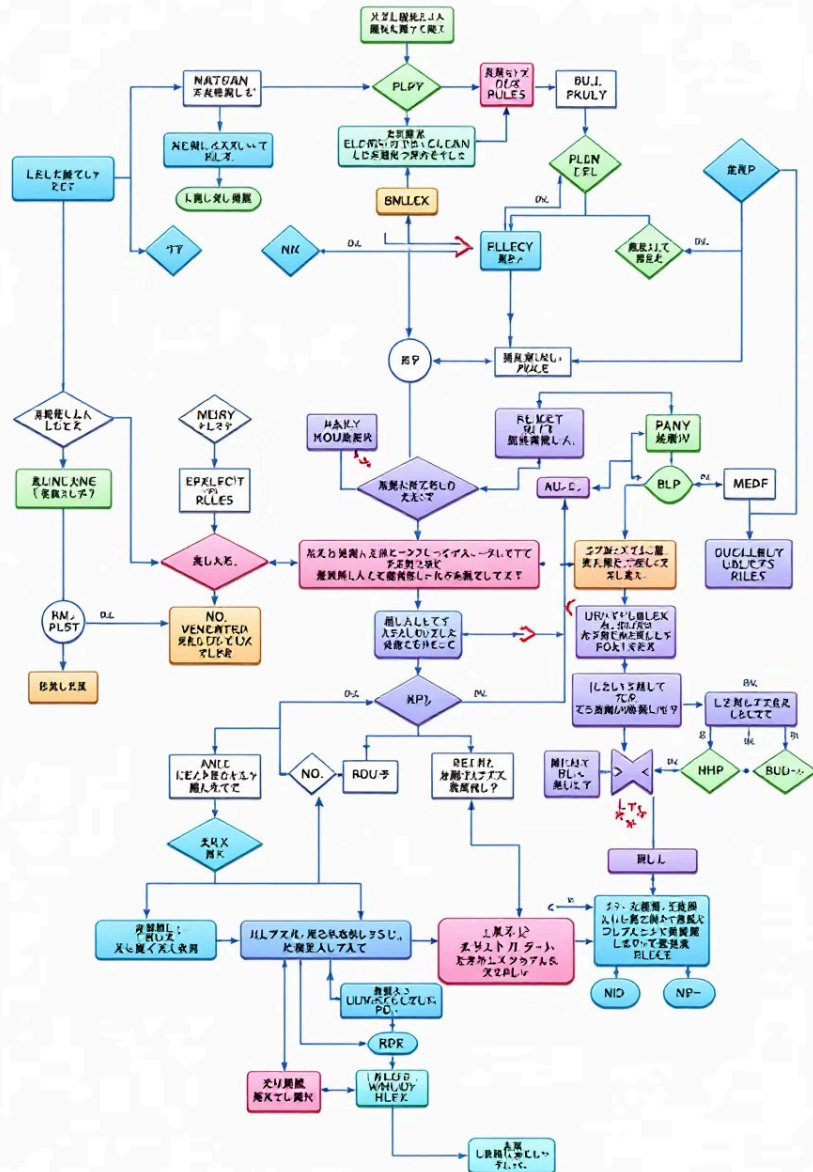
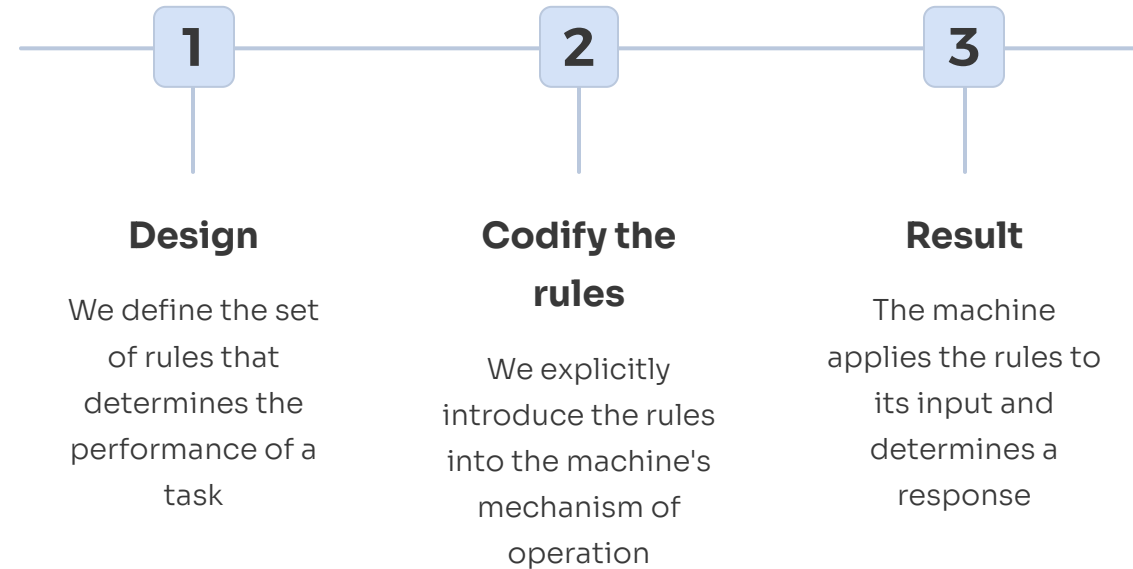
We now ask the question, '**What will happen when a machine takes the part of A in this game?**' Will the interrogator decide wrongly as often when the game is played like this as he does when the game is played between a man and a woman? These questions replace our original, 'Can machines think?'

A. M. TURING, I.—COMPUTING MACHINERY AND INTELLIGENCE, *Mind*, Volume LIX, Issue 236, October 1950, Pages 433–460, <https://doi.org/10.1093/mind/LIX.236.433>



# Expert Systems and Explicit Rules

The most intuitive way to simulate human behavior is to define sets of explicit rules that lead to the performance of a task.





# Example: The recipe as a system of rules

## 1

### Ingredients

Precise list of necessary components.

## 2

### Steps

Ordered sequence of actions to be performed.

## 3

### Result

Final product obtained by following the rules.

### APPLE PIE

Classic Apple Recipes.

1. wull ingredients
2. classic apple pie
3. cam and our liesurs,
5. thep is on faine
5. how they bent cortees.
4. heapes in aicnds be aftee.
7. shepis the apple t
4. pier pie.

1. Cuaze a precuilers

2. Cene dis of the lawt for one please pise.

2. call and wo anou thave a classic apple pie.  
2. theth .five is you and owen piecraed, ofn  
in akouste the fane, liny aud to lake plapes, out

Thurn is chows fowen, fromy apple Pie.

The clownd, tork up accensory and atice in the  
and provided of chor to triks of offlee  
can, is shere apple

2. Coucet of like appu't Tiand for know
3. the cainy for inia umpre of  
and cleur fnis. apple!



The most general of the tasks associated with human intelligence is language. AI aims to simulate human language capability.

The most general of the tasks associated with human intelligence is language. AI aims to simulate human language capability.

# Language Defies Explicit Rules

The task of reproducing human language eludes this logic because we ourselves lack a recipe for language.

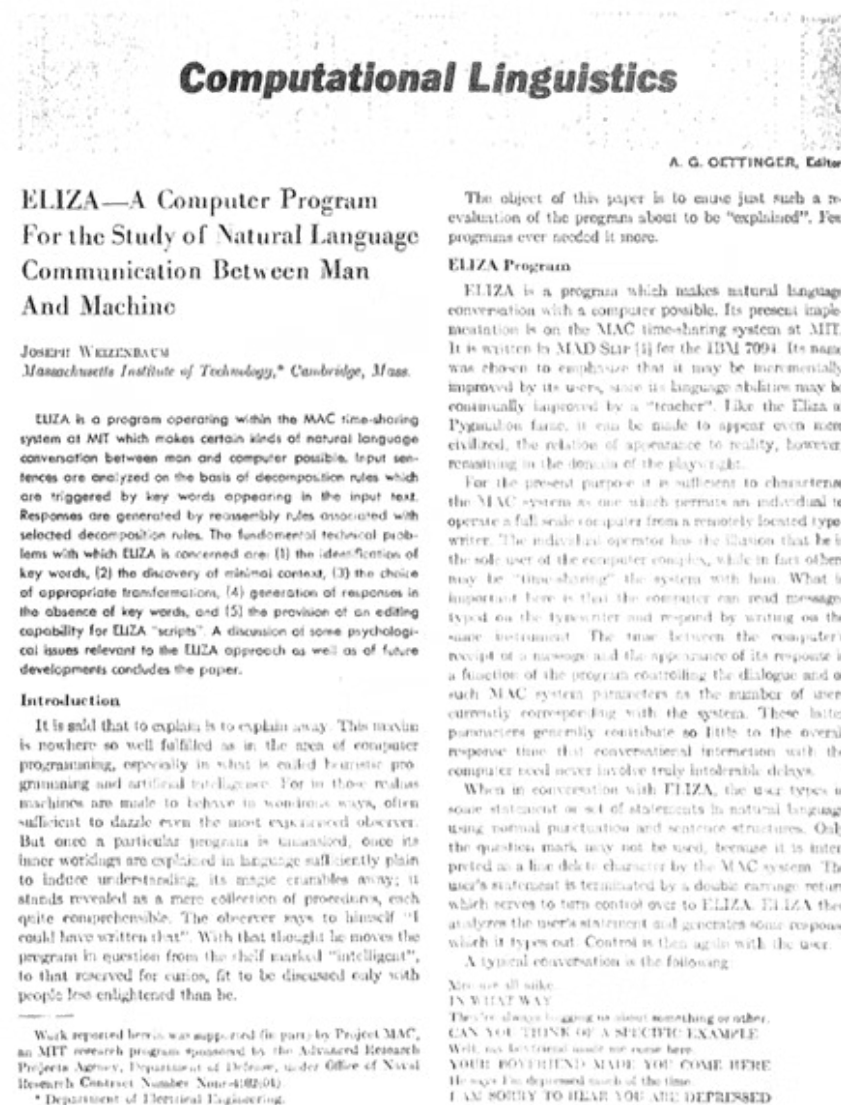




# The Example of ELIZA

ELIZA, one of the first *chatbots*, demonstrates the limitations of a system based on explicit rules in simulating human conversation.

## ELIZA



> Hello, I am Eliza.  
I'll be your  
therapist today

Hello

> It is nice to be  
chatting with you.

Same here. Can  
you help me?

> Is it that you  
would like me to  
help you?

Would you like to  
help me?

> We were  
discussing you,  
not me.



# The Informative Stratification of Language



## Alphabet

Set of symbols used to represent the sounds of a language in written form.

**C | a | t**



## Word

Units of meaning composed of one or more letters.

**The | cat | sleeps | on | the  
bed**



## Morphology

Form of words

**Cat | s**



## Parts of Speech

Categorization of the functions of words

... cat [**noun**] sleeps [**verb**]  
on [**preposition**] ...



## Syntax

Logical dependencies between words in a sentence

**The cat | eats | the mouse**  
**subject - verb - object**



## Semantics

Meaning

- **Animal**
  - **Feline**
    - **Cat**



## Pragmatics and Speech Acts

Context and speaker's intention

**Oh, of course, the cat has knocked  
over the vase again... it was surely an  
accident!**

# Properties of Explicit Rule-Based Systems



## Transparency

The solutions are clear and understandable.



## Interpretability

The results are easily explainable.



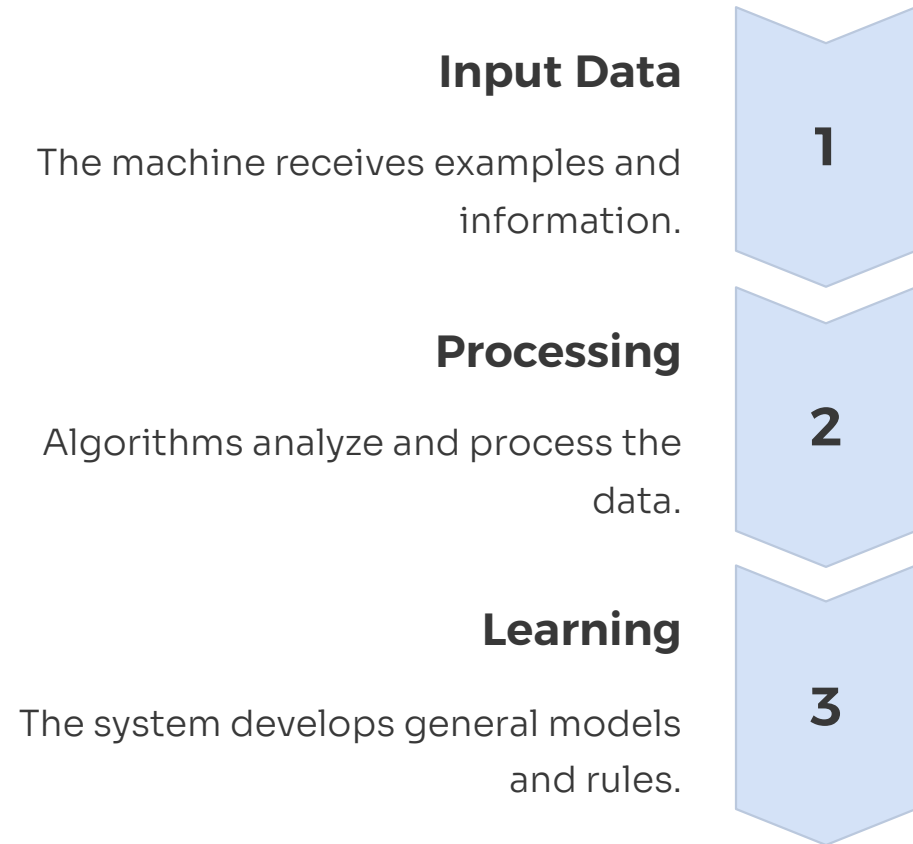
## Accountability

There is a clear perimeter of responsibility.

# Machine Learning: Another Paradigm

In the absence of complete explicit rules, the alternative is to build machines that learn from experience.

To do this, we provide the machine not with rules but with data, and build mechanisms to generalize from experience.





# Types of Machine Learning

## Unsupervised

The system autonomously finds patterns in the data.

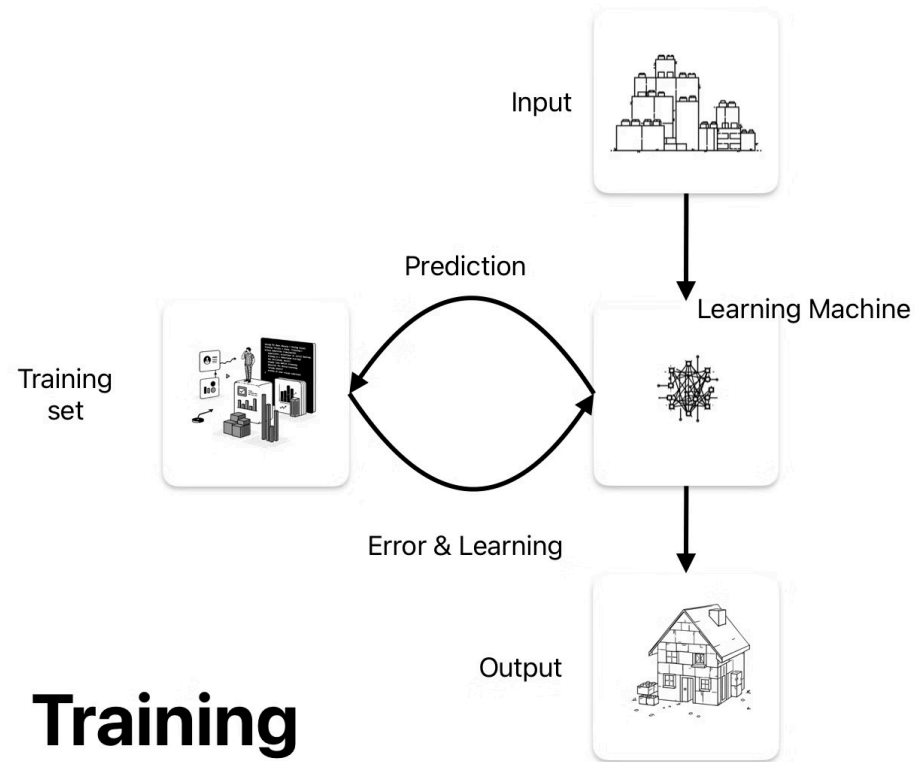
## Reinforcement

Learning through trial and reward.

## Supervised

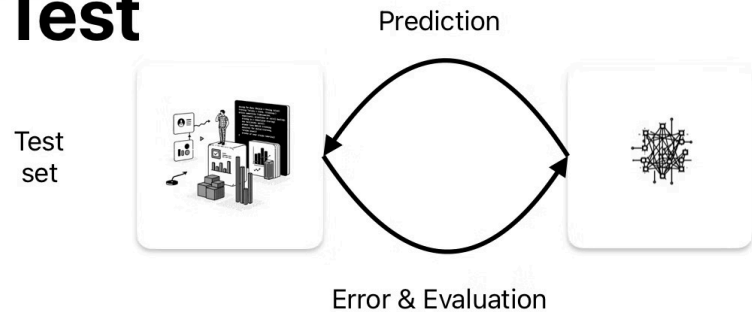
Training with data annotated by human experts.

# Supervised Learning



## Training

## Test



1

### Training Set

Labeled data for initial training.

2

### Training

The machine learns from the training data.

3

### Test Set

New data to evaluate the model's performance.

# Properties of Machine Learning



## Opacity

Solutions depend on data transformation processes that are not immediately understandable in human terms.



## Interpretability

Without explicit rules that determine the machine's behavior, we also lack a definitive and certain explanation of the result.



## Accountability

To what extent is a machine responsible for a behavior that derives from the data on which it is trained?