Research ,Teaching and Learning in the era of ChatGPT.

Dr Islam Alkabbany

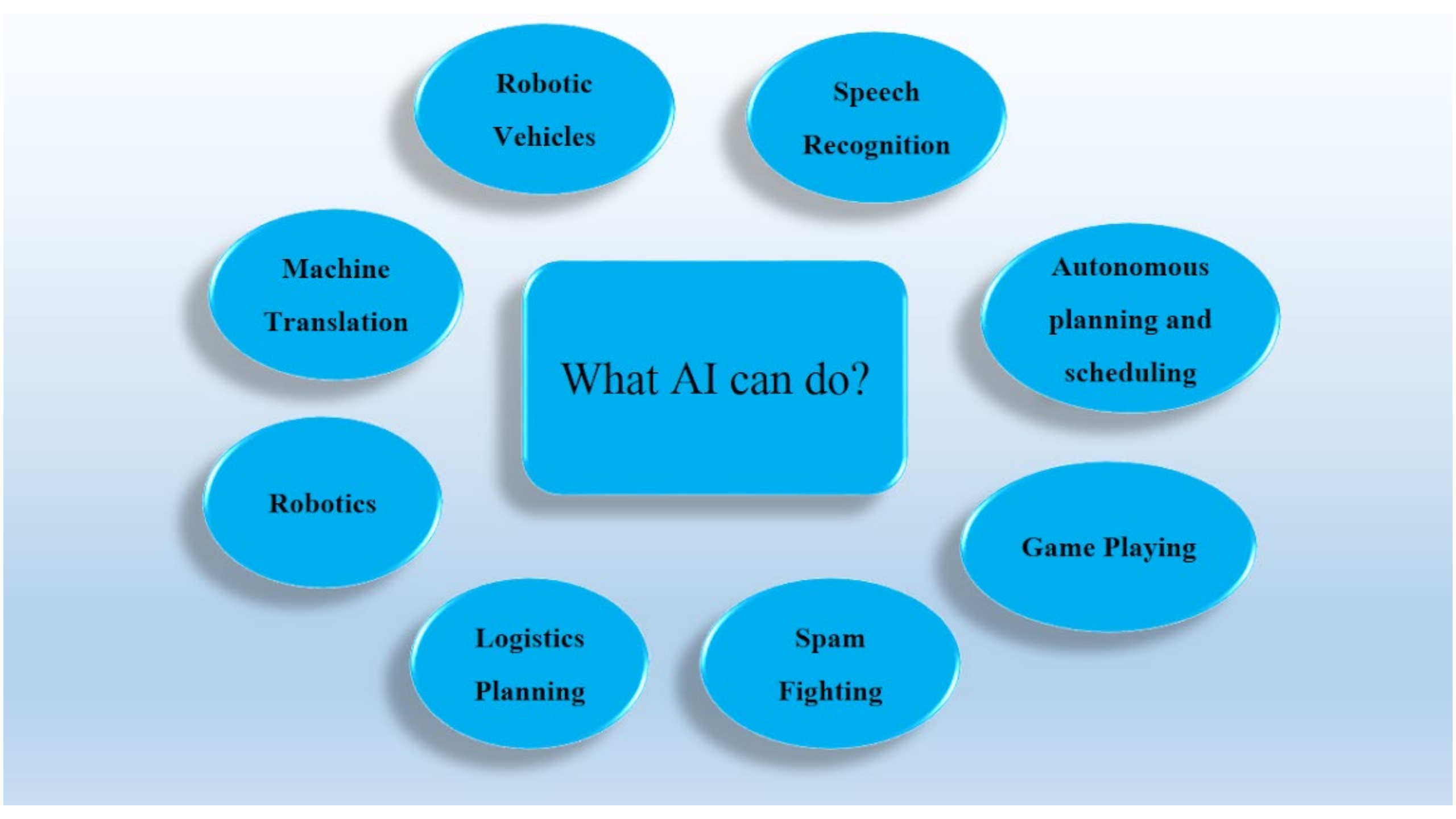
What is ChatGPT

- **ChatGPT (Chat *Generative* Pre-trained *Transformer*)** is a chatbot developed by OpenAI and launched on November 30, 2022. Based on a ***large language model.***

- it enables users to refine and steer a conversation towards a desired length, format, style, level of detail, and language

What is AI

- Artificial Intelligence (AI) refers to the development of computer systems or software that can perform tasks that typically require human intelligence.
- These tasks include problem-solving, learning, understanding natural language, recognizing patterns, speech recognition, and decision-making.
- The goal of AI is to create systems that can mimic or simulate human intelligence, enabling them to adapt and improve over time based on experience and data.


Robotic Vehicles
Speech Recognition
Machine Translation
What AI can do?
Autonomous planning and scheduling
Robotics
Game Playing
Logistics Planning
Spam Fighting

What AI can do?

- **Solving Problems:** AI can solve complex problems by analyzing data and making predictions. (**Fraud Detection in Financial Transactions)**

- **Making Decisions:** AI systems can make decisions based on patterns and data analysis. (**Flight Booking System)**

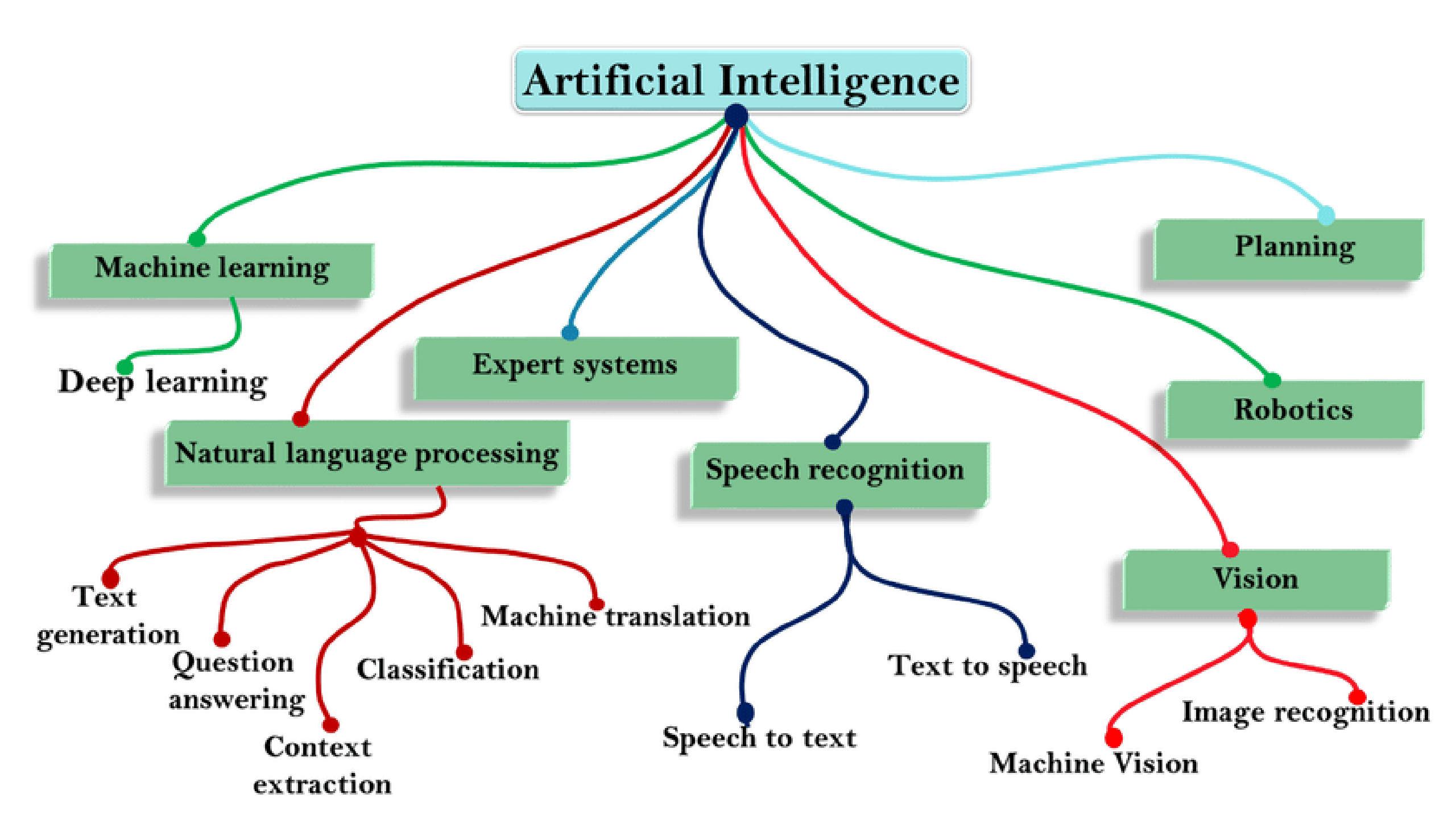
- **Learning from Data:** AI learns from data, improving its performance over time. (Speech recognition systems getting better with more user interactions.)

What AI can do?

- **Natural Language Processing:** AI provide the ability to understand and generate human language.(Virtual assistants like Siri and. ChatGPT)

- **Image and Pattern Recognition:** AI can recognize and interpret patterns in images and visual data. (Facial recognition)

- **Autonomous Systems:** AI can use to create autonomous systems that can operate without human intervention. (Self-driving cars)


Artificial Intelligence
Machine learning
Deep learning
Expert systems
Natural language processing
Text generation
Question answering
Context extraction
Classification
Machine translation
Speech recognition
Speech to text
Text to speech
Planning
Robotics
Vision
Image recognition
Machine Vision

AI subfields

- Artificial Intelligence (AI) is a multidisciplinary field with several subfields that focus on specific aspects of intelligent systems

1. **Machine Learning (ML):** Teaching computers to learn from data without explicit programming. *Applications:* Predictions, pattern recognition, and decision-making.
2. **Natural Language Processing (NLP):** Enabling computers to understand, interpret, and generate human language. *Applications:* Language translation, chatbots, sentiment analysis.
3. **Computer Vision:** Empowering machines to interpret and understand visual information. *Applications:* Image recognition, object detection, facial recognition.

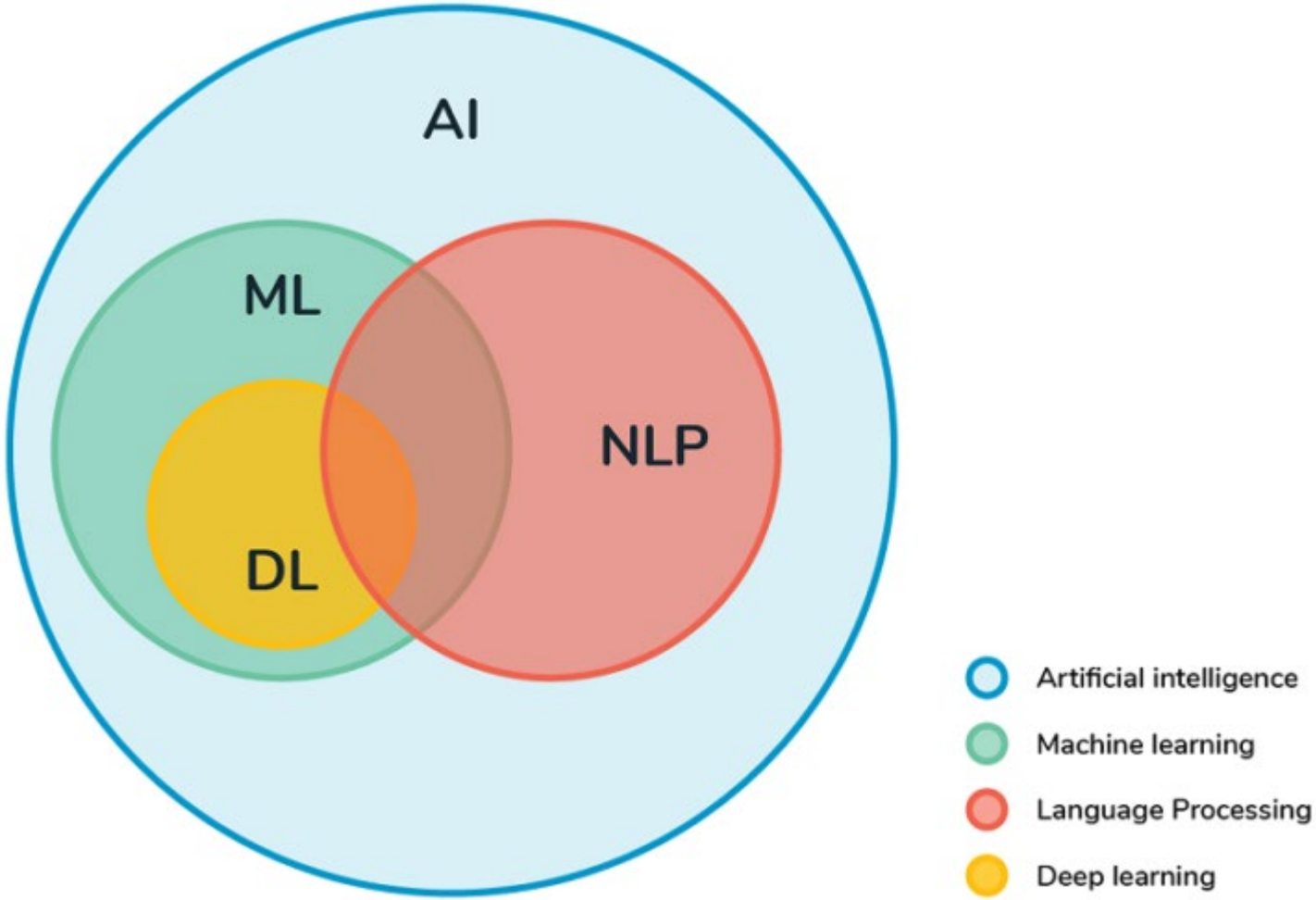
AI subfields

4. **Robotics:** Combining AI with mechanical systems to create intelligent machines for physical tasks. *Applications:* Industrial automation, drones, autonomous vehicles.
5. **Expert Systems:** Computer programs designed to mimic the decision-making abilities of human experts in a specific domain. *Applications:* Making decisions, solving problems, and providing expertise in a particular field.
6. **Knowledge Representation and Reasoning:** Dealing with how to represent information about the world in a form that a computer system can utilize for solving complex tasks. A*pplications:* Logical reasoning, drawing conclusions from information, and representing knowledge in a structured way.

- These subfields often overlap, and advancements in one area can contribute to progress in others

AI subfields

- These subfields often overlap, and advancements in one area can contribute to progress in others



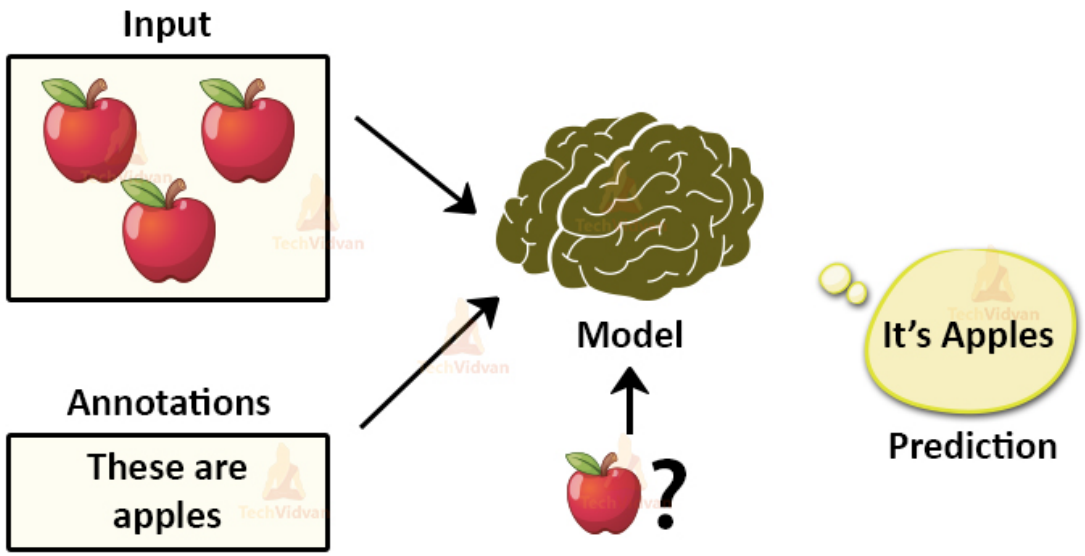
Machine Learning (ML)

- Machine learning is a subset of AI that focuses on the development of algorithms and statistical models that enable computers to perform a task without being explicitly programmed.

- It includes techniques like supervised learning, unsupervised learning, and reinforcement learning.

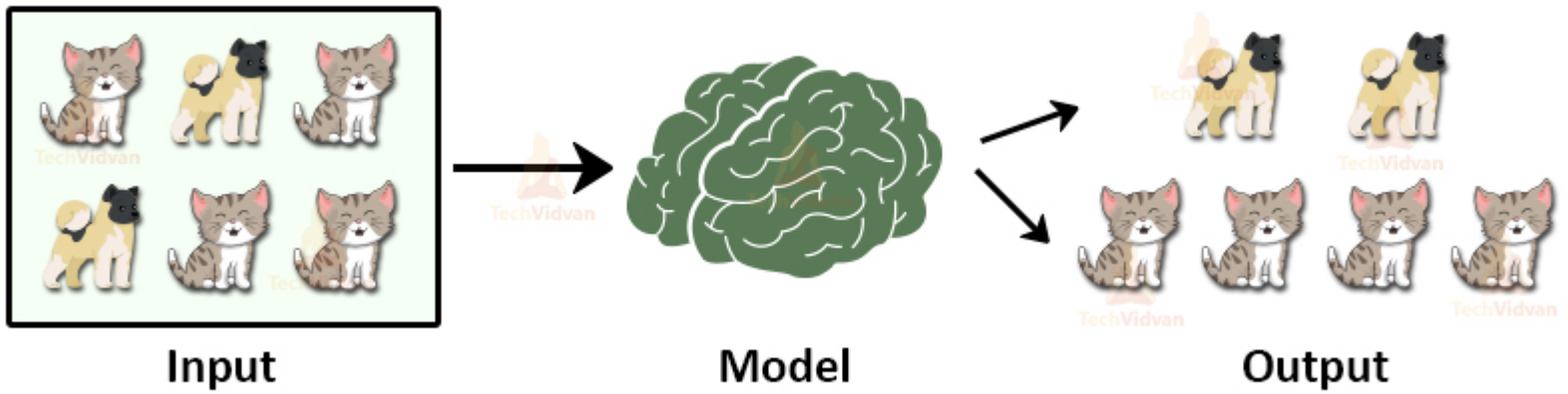
Machine Learning (ML)

- **Supervised learning** is a type of machine learning where the algorithm is trained on a labeled dataset, which means that the input data used for training is paired with corresponding output labels. The goal of supervised learning is to learn a mapping from the input data to the correct output by generalizing from the labeled examples.



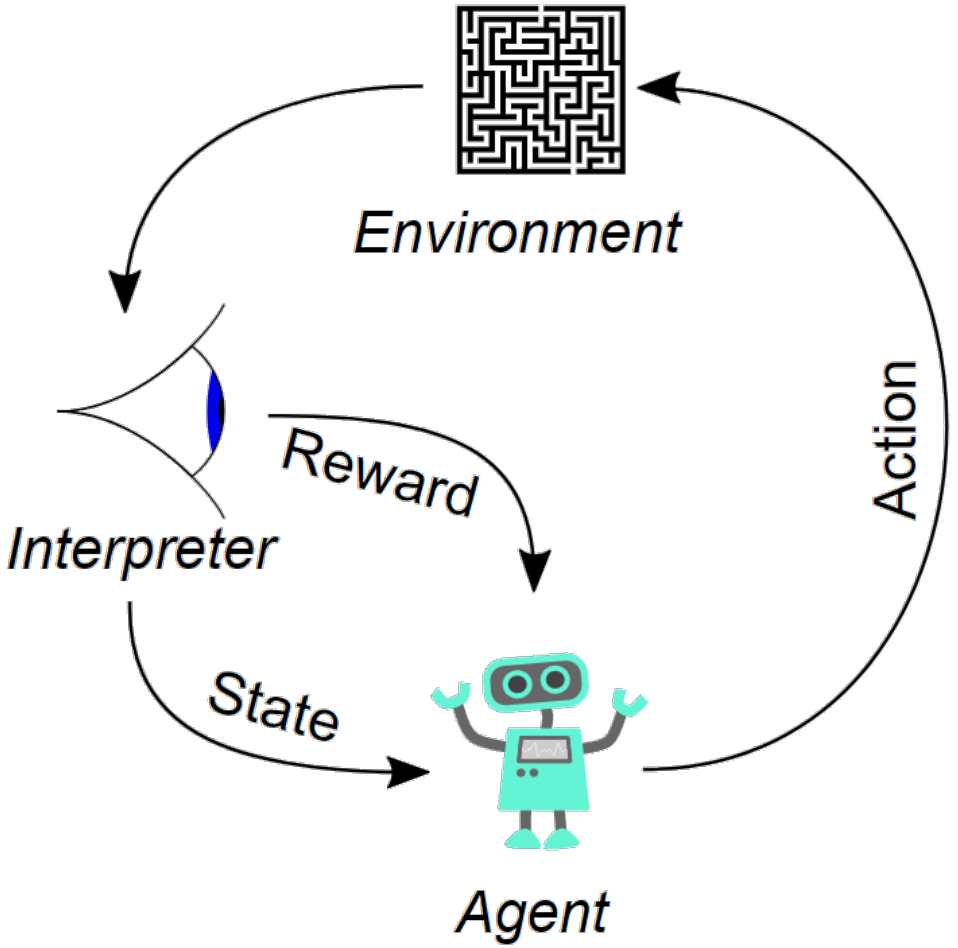
Machine Learning (ML)

- **Unsupervised learning** is a type of machine learning where the algorithm is given a dataset without explicit labels or target outputs. The goal of unsupervised learning is to find patterns, relationships, or structures in the data without the guidance of predefined output labels. In essence, the algorithm explores the inherent structure of the data to discover hidden patterns or groupings.



Machine Learning (ML)

- **Reinforcement Learning (RL)** is a type of machine learning paradigm where an agent learns to make decisions by interacting with an environment. The agent receives feedback in the form of rewards or penalties based on the actions it takes, and its objective is to learn a strategy (policy) that maximizes the cumulative reward over time

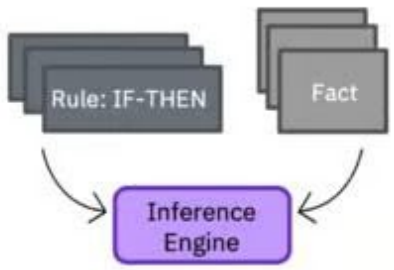


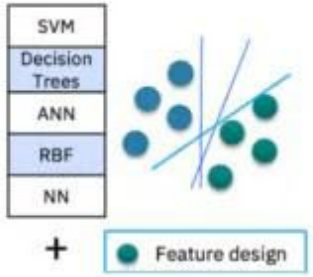
History of AI

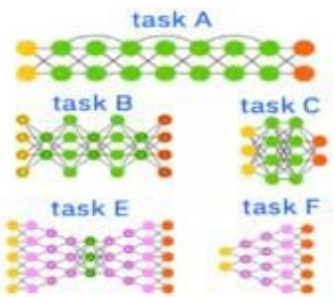
1980s Big data

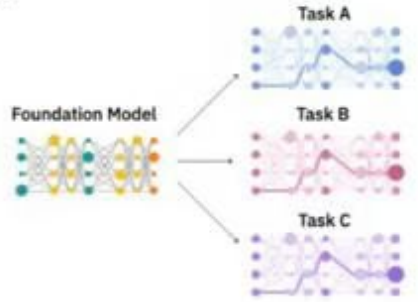
1980s to ~2010 Massive labeled data + Compute

2017+ **Self-supervision at scale** + Massive unlabeled data + Compute









Expert Systems

- Manually-crafted symbolic representations and rules
- No use of data and brittle

Machine Learning

- Less brittle but labor intensive
- Demanding data prep and feature engineering

Deep Learning

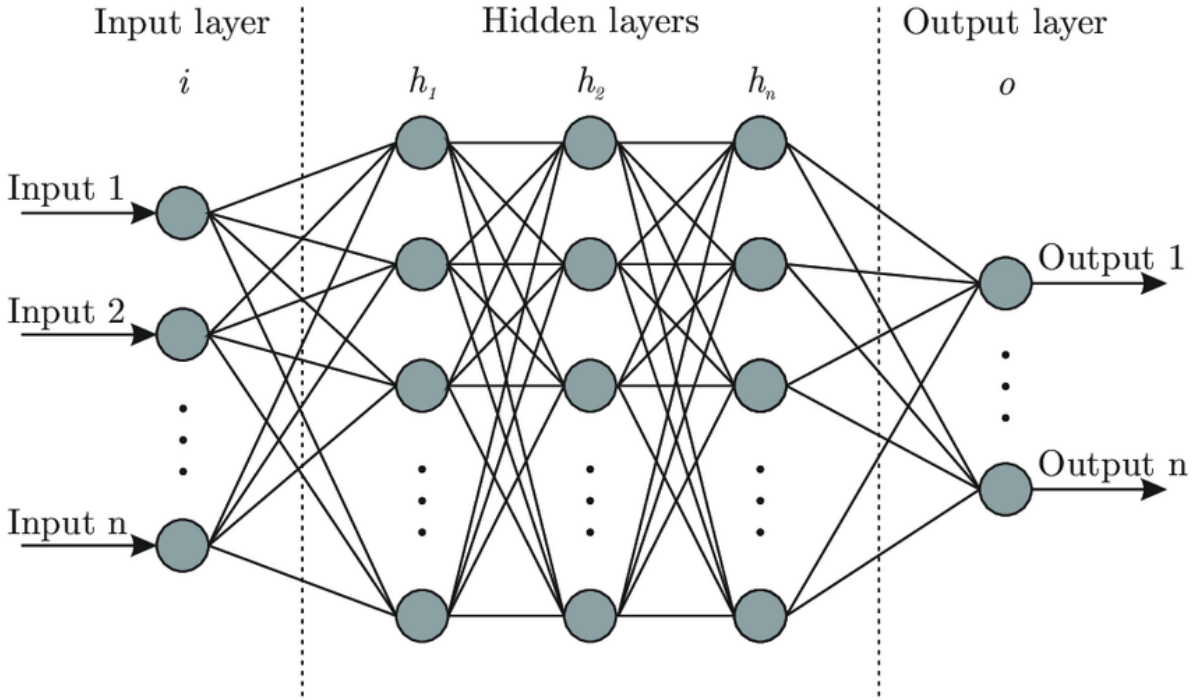
- Automatically learn if you have enough labeled data
- Enterprise adoption limited by availability of labeled data

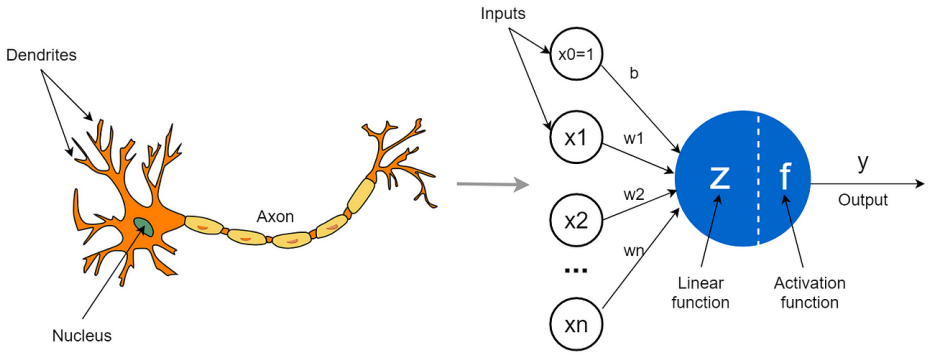
Foundation Models

- Learn from lots of data *without requiring labels*
- Quickly adopt to enterprise tasks using limited labels

Artificial Neural Networks and Deep Learning

- ANNs are computational models inspired by the structure and functioning of the human brain, designed to perform tasks such as pattern recognition and NLP



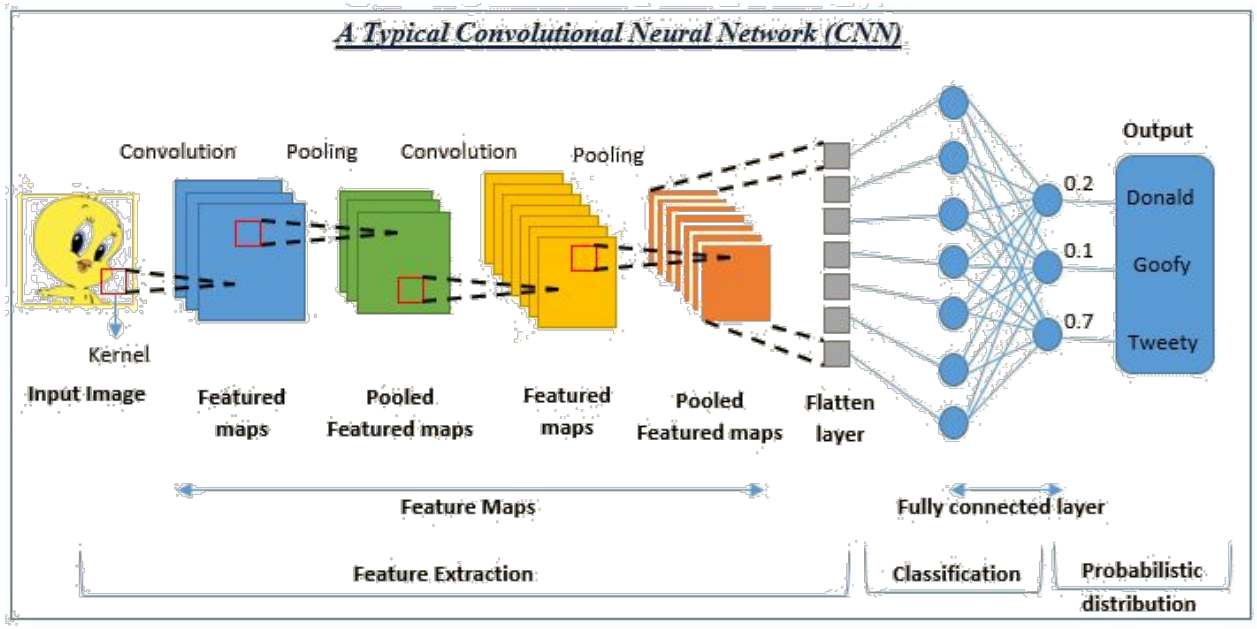


Deep Learning

- Deep learning is a subset of machine learning that involves neural networks with multiple layers (deep neural networks)
- **Characteristics of Deep Learning**
  1. **Depth:** The presence of multiple hidden layers.
  2. **Representation Learning:** Learning hierarchical representations of data.
  3. **Feature Hierarchies:** Automatic extraction of features at various levels.

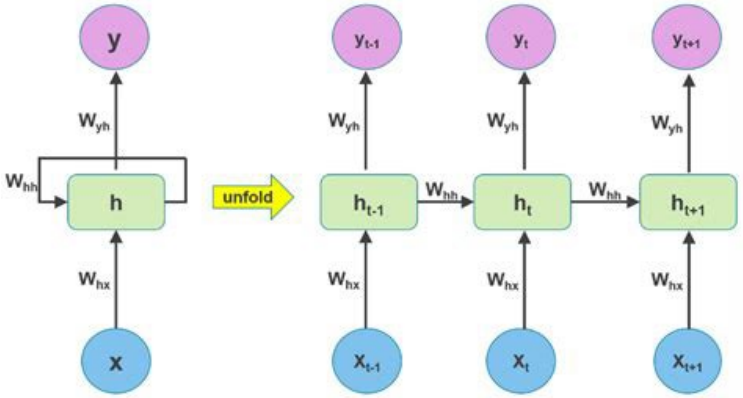
Convolutional Neural Networks (CNNs)

- A type of artificial neural network designed for tasks involving visual data, such as image and video recognition, image classification, object detection, and image generation
- CNNs are particularly effective in capturing spatial hierarchies and patterns in data, making them well-suited for tasks where the arrangement of features in the input matters.



Recurrent Neural Networks **(RNNs)**

- A type of artificial neural network designed to process sequential and temporal data. Unlike traditional feedforward neural networks, which process input data in a single pass, RNNs have connections that form directed cycles, allowing them to maintain a memory of previous inputs
- Suited for sequence data. Used in natural language processing, speech recognition, and time series analysis.
- Slow to train and can forget previous words in a sequence.



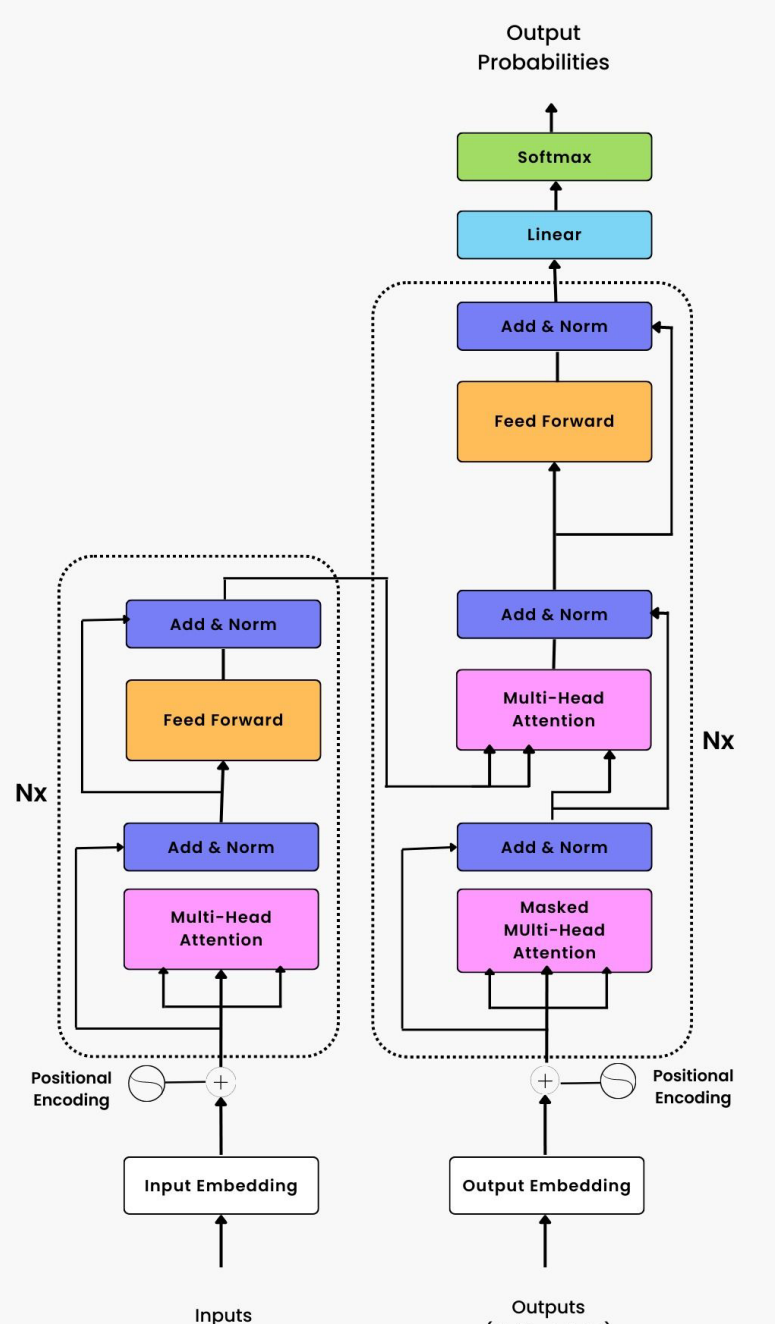
Long short term memory**(LSTM)**

- Recurrent neural networks with special components that allowed past data in an input sequence to be retained for longer. Suited for sequence data. Used in natural language processing, speech recognition, and time series analysis.

- LSTMs could handle strings of text several hundred words long, but their language skills were limited. .

Transformers

- The breakthrough behind today's generation of large language models came when a team of Google researchers invented transformers*
- The Transformer models use a modern and evolving mathematical techniques set, generally known as attention or self-attention.
- Namely, it enables neural networks to focus on specific parts of the input sequence while processing the data. This mechanism allows the model to assign different weights to different positions in the sequence, enabling it to capture the relevance and importance of each position more effectively.

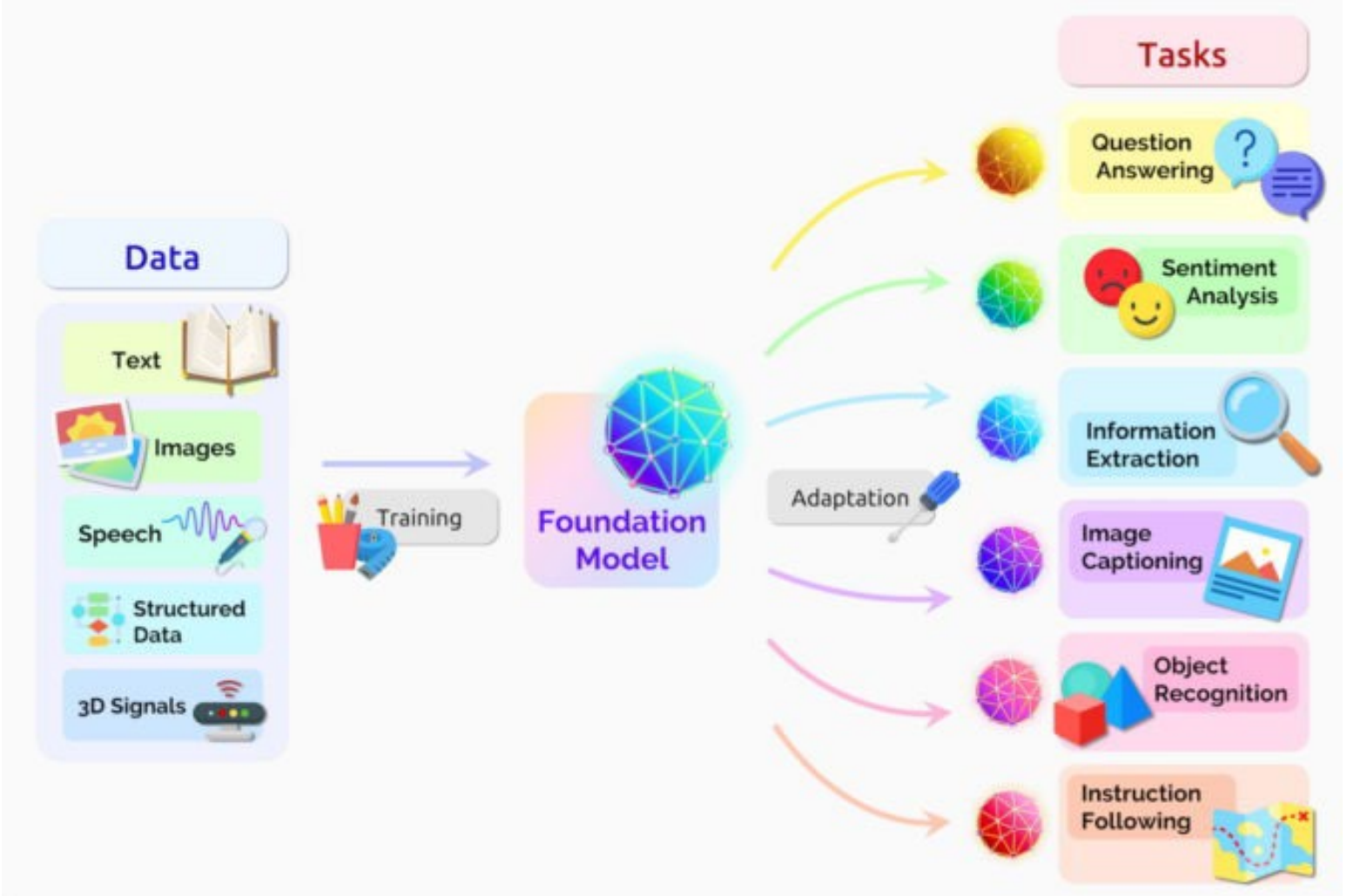


* Vaswani, Ashish, et al. "Attention is all you need." *Advances in neural information processing systems* 30 (2017).

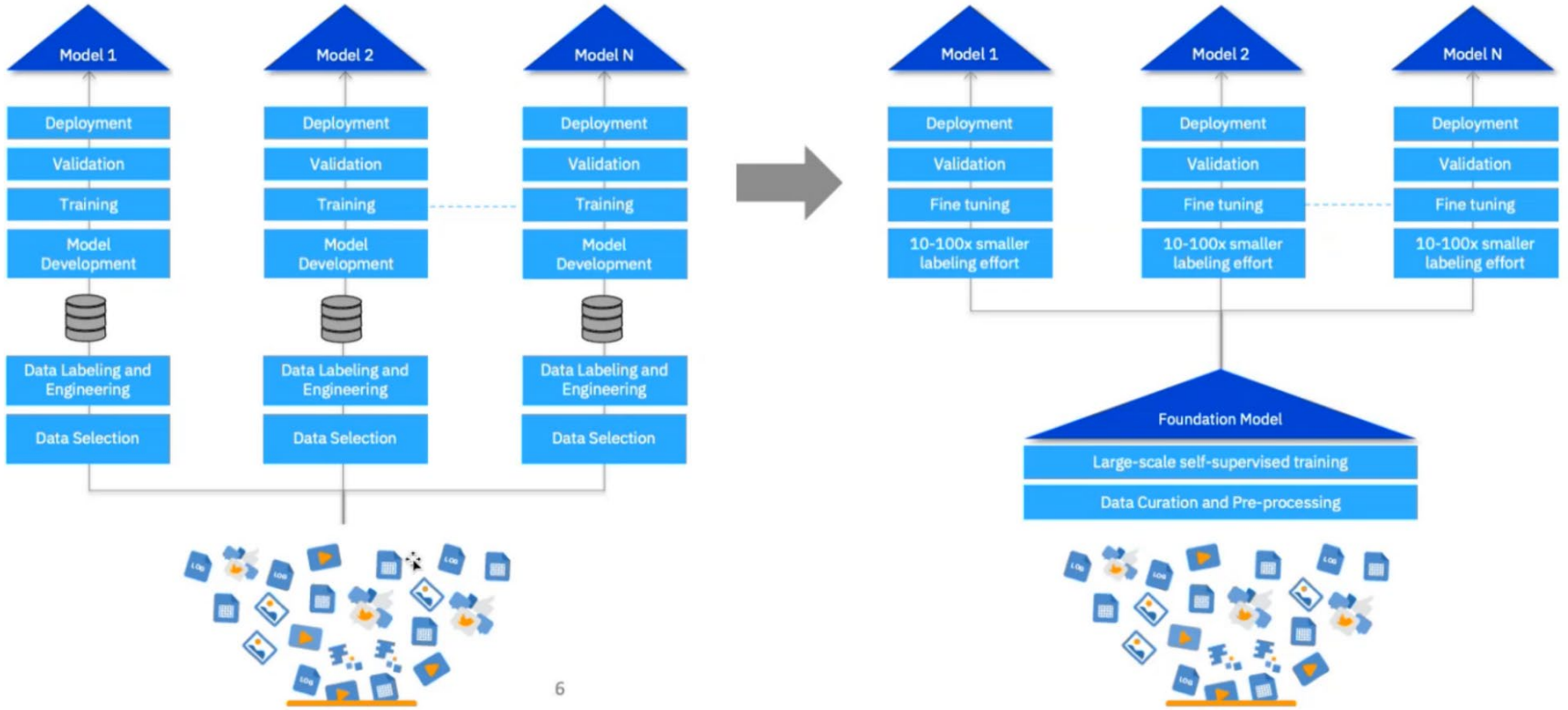
Transformers

- Transformers process long sequences in their entirety with parallel computation, which significantly decreases both training and processing times(No recurrent cells). This has enabled the training of very large language models (LLM)
- Transformers enable machines to understand, interpret, and generate human language in a way that's more accurate than ever before. They can summarize large documents and generate coherent and contextually relevant text for all kinds of use cases. Virtual assistants like Alexa use transformer technology to understand and respond to voice commands.
- Translation applications use transformers to provide real-time, accurate translations between languages. Transformers have significantly improved the fluency and accuracy of translations as compared to previous technologies.

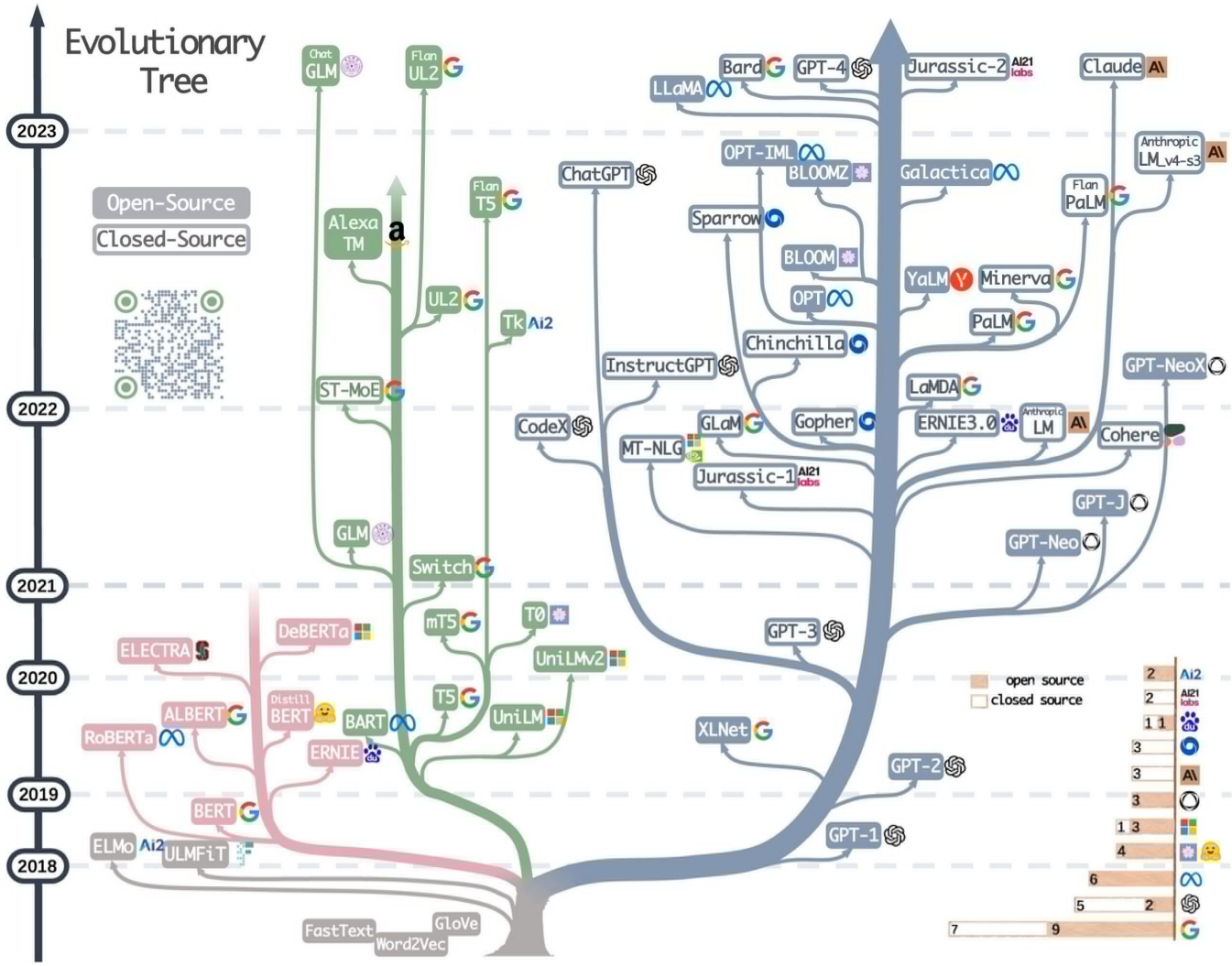
Large Model (Foundation Model)



Foundation model workflow



Large Language Model



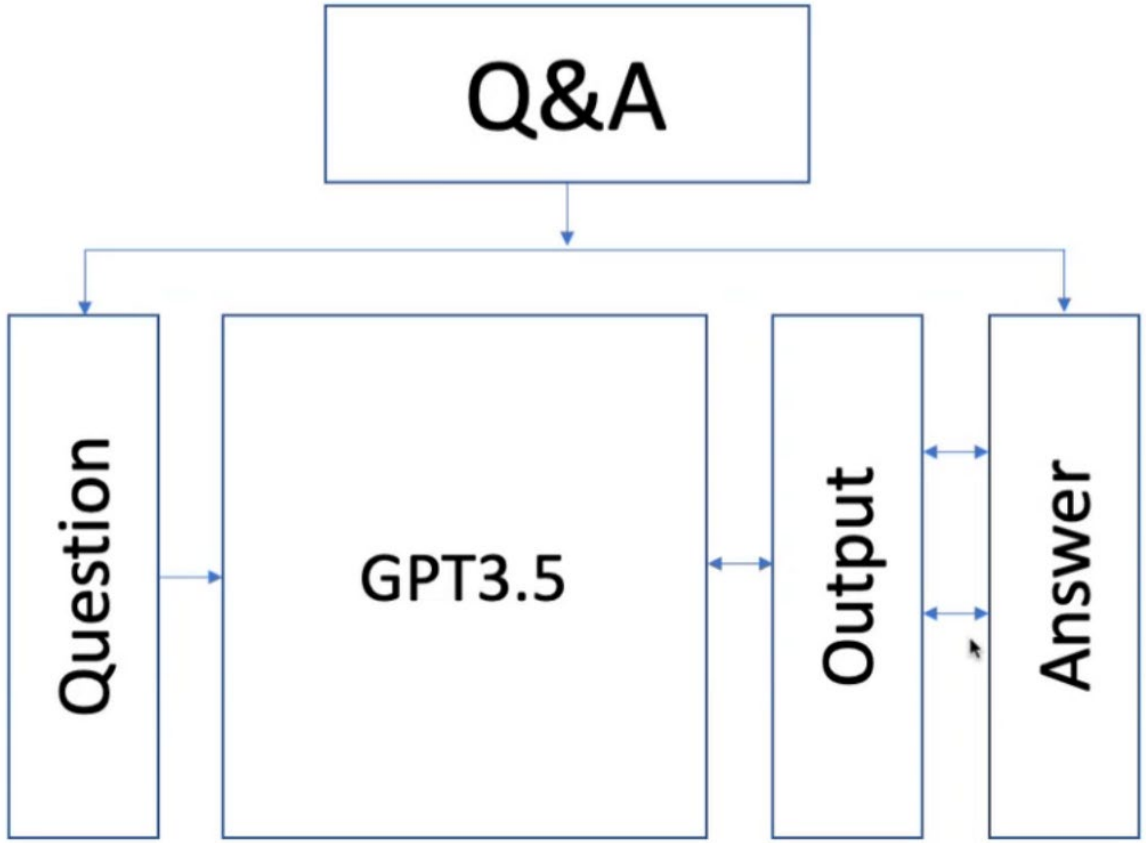
Chat GPT training

- It was estimated that the critical process of training a large language model such as GPT-3 could cost over $4-10 million per model.
- ChatGPT-3 has 175 billion parameters compared with GPT-2's 1.5 billion.
- Infrastructure used to train included thousands of NVIDIA AI-optimized GPUs
- Researches suggests training for GPT-3 alone consumed 185,000 gallons (700,000 liters) of water for cooling.

Chat GPT training - pretrain

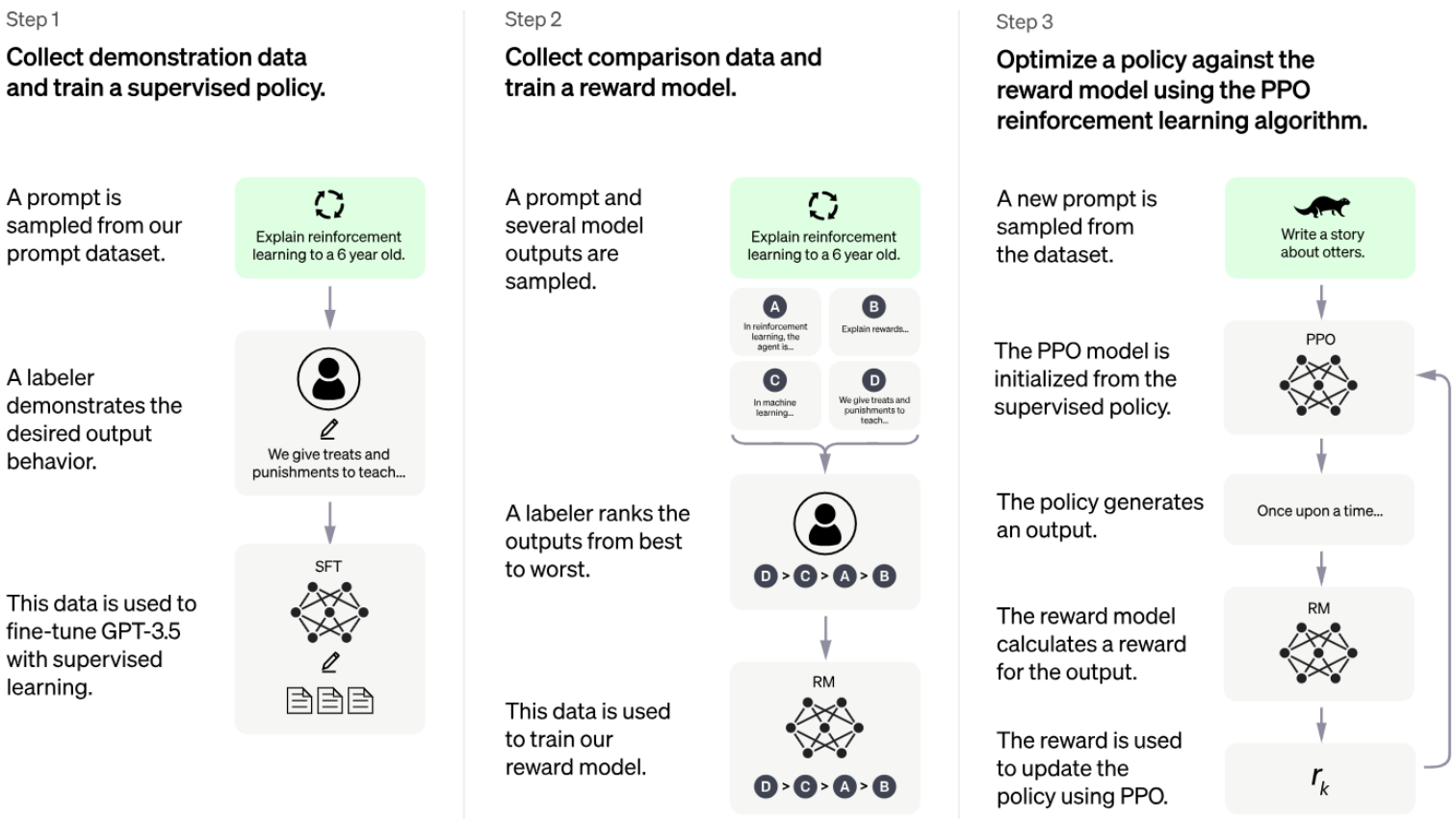
- Pre-training played a crucial role in developing a ChatGPT as it helped to learn the basic rules of language and understand common word usage and phrases. This knowledge is then used as a foundation for further customization.
- ChatGPT was trained on hundreds of thousands of books, articles, dialogues, including:
  - WebText2 (a large library of over 45 terabytes of text data)
  - Cornell Movie Dialogs Corpus (a dataset containing over 200,000 conversations between 10,000 movie characters in movie scripts)
  - Ubuntu Dialogue Corpus (a collection of 1,000,000 multi-turn dialogues between Ubuntu users and the community support team)
  - billions of lines of code from GitHub
- The training objective was to predict the next token in a given text based on the context of preceding words. By comparing its prediction with the actual next token in the sentence, the model adjusts its weights during training to enhance the accuracy of future predictions.

Chat GPT training – fine tuning



Chat GPT training – fine tuning

- Step 1 involves supervised fine-tuning
- Step 2 involve training a reward model
- Step 3 Proximal Policy Optimization (PPO) algorithm



ChatGPT capability

- Generating and helping to improve prose and code development
- Summarizing text
- Classifying content
- Answering questions
- Translating and converting language (including programming languages)

Issues with ChatGPT training

- It is only trained on data through September 2021, so it has limited knowledge of events that have occurred since then.

- Although it gives the illusion of performing complex tasks, it has no knowledge of the underlying concepts; it simply makes predictions

- Plausible-sounding but incorrect or nonsensical data This behavior is common for large language models, and is called "hallucination".

Issues with ChatGPT training

- No reference or explainability

- ChatGPT attempts to reject prompts that may violate its content policy. Despite this, some users managed to jailbreaking ChatGPT with various prompt engineering techniques to bypass these restrictions.

- Training Data:
  - Fake new/information infiltration
  - Biasing due to train of LLM generated data
- OpenAI also launched a Custom Models program which offers even more customization than fine-tuning allows for. Organizations can apply for a limited number of slots (which start at $2-3 million)

Academic research

- Scientific journals have different reactions to ChatGPT.
- Some including Nature and JAMA Network, "require that authors disclose the use of text-generating tools and ban listing a large language model (LLM) such as ChatGPT as a co-author".
- Science, IEEE "completely banned" usage of LLM-generated text in all its journals.
- Many authors argue that the use of ChatGPT in academia for teaching and review is problematic due to its tendency to hallucinate. Robin Bauwens, an assistant professor at Tilburg University, found that a ChatGPT-generated peer review report on his article mentioned fake studies.

ChatGPT in Education

- While ChatGPT could help instructor to generate presentation , lecture or even a whole course. The integrity of the provided text is not guaranteed

- Students could use LLM tools to do their work for them, passing off A.I.-generated essays, generated code and problem sets as their own.

ChatGPT in Education

- You can detect Chat GPT-written text using online tools
  - AI Text Classifier
  - GPTZero
  - Originality.AI
  - Writer AI Content Detector
  - ZeroGPT

ChatGPT in Education

- Most of this tools need at least 1000 word to get result with good accuracy.
- Even though AI text detector is not perfect and has a lot of false positive.
- There are tools that can manipulate the generated text to make it non detectable

ChatGPT in Education

- There's of course a rosy way of seeing this as just becoming another tool for learning and productivity like Google or Wikipedia
- From this outlook, it will just become standard for students to use tools like ChatGPT for all assignments, and so they end up getting graded on how effectively they can leverage the tool rather than problem solve by themselves.
- This sort of thing already happening before ChatGPT, student used to google the assignment and basically which leads to assessing how good students are at using Google to find the answer
- Teacher have to either spent more time to detect the use of AI tool, or be more creative in assigning tasks.

Recommendations

- **Proceed but don't over-pivot.** Recognize that this is very early stage and much of what you are hearing is hype. That said, the potential is significant.
- **Explore other emerging generative AI use cases.** Go beyond GPT language-focused ones.
- **Encourage careful experimentation.** Encourage out-of-the-box thinking about work processes, but not before you define usage guidelines, ensure understanding of the risks, issues and best practices, and have all generated text reviewed by humans.

Similar LLM tool

- Google has recently released its new AI model, PaLM(Pathways Language Model) , which is reported to be three times larger and more powerful than ChatGPT
- It has 540 billion parameter