



INDIAN INSTITUTE OF
INFORMATION
TECHNOLOGY

History of Artificial Intelligence (AI)

Dr. Animesh Chaturvedi

Assistant Professor: IIIT Dharwad

Young Researcher: Heidelberg Laureate Forum

Postdoc: King's College London & The Alan Turing Institute

PhD: IIT Indore MTech: IIITDM Jabalpur



Indian Institute of Technology Indore
भारतीय प्रौद्योगिकी संस्थान इंदौर



PDPM
Indian Institute of Information Technology,
Design and Manufacturing, Jabalpur

The
Alan Turing
Institute

History of Ancient Human Intelligence

35

वोद्वेषनवद्युजाग्निः लनूकद्वैथिप्रमतिश्चकार्वेत्तं कल्याणवसुविश्वमिवत्त
मर्मन्तप्रमतिस्तंषितसिनुर्लवयस्त्वं तव नामयोवयोसेवाग्नयः भूतेनः संसहस्रिणः
सुवोरयनिप्रत्यपानं दृश्य ॥ ३३ ॥ त्वं मन्त्रप्रथममायुमायवद्वा अक्षण्डन्त्वं दृष्ट्यात्
श्वनि। इच्छामत्प्रभूमनुष्ट्युग्नासने। ध्रुत्युग्नामर्मक्ष्यतापत्तिवांगो अग्नेजवद्वा
यात्क्षिमेयोनौरक्षन्त्वं दृष्ट्यवद्य। ग्रातात्कस्यनेनेयग्नामुख्यमिमं त्रूक्ष्यापाह्ववद्य
तात्मन्त्रयज्यवपायुरन्तरा निवेगायवत्तुरक्षदृष्ट्यस्योरात्महव्याद्यक्षयधयस्मी
रश्चन्मन्त्रमनसावनोषितं त्रूलमन्त्रउत्तरशास्त्रायवाघनेष्या हृष्ट्वक्षणः परमवज्ञापित
त्। आप्रस्फुत्त्रमनिरुद्यसेष्टिनाप्याकं वास्त्रिप्रदि शोऽविद्वरः। त्वं मन्त्रप्रयत्नदिः

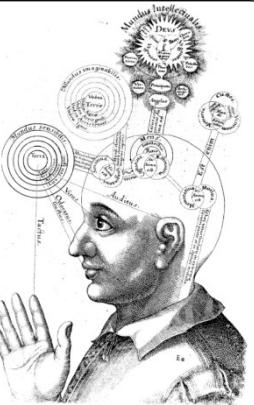
Ancient Human Intelligence

How did the Ancient Human's Neural Network get trained?

- Vedas since 2nd millennium BCE, Vedas composition 1500–1200 BCE
- Its origin in Vedas, **Multi-Layers of Intelligence** is Sanskrit words: for example
- **Yama** (यम) right living and **ethical rules**
- **Dhi** (धी) meaning 'diligent' 'understanding', 'reflection', 'mind', 'design', 'intelligence', 'opinion', 'imagination', 'notion', and 'intellect'
- **Dhṛti** or **Dhriti** or **Dhruti** (धृति)
 - 'determination', 'patience', 'firmness', 'perseverance',
 - 'to bear', 'fortitude'
- **Smriti** (स्मृति) or **Smara** (स्मर) **Remembered Knowledge**
- **Shruti** (श्रुति) **Heard Knowledge**
- **Vac** (वाच) **Speech**



Sat-Cit-Ananda “True-Conscious-Bliss”



How did the Ancient Human’s Neural Network get trained?

- An epithet and description for the subjective experience of the ultimate unchanging reality, called Brahman, especially Vedanta.
- It represents "existence, consciousness, and bliss" or "truth, consciousness, bliss".
- **sat** (सत्): In Sanskrit, *sat* means "being, existence", "real, actual", "true, good, right", or "that which really is, existence, essence, true being, really existent, good, true".
 - In AI, ground truth ([True Positive](#) and [True Negative](#))
- **cit** (चित्): means "consciousness" or "spirit"
 - In AI, it is analogous to "[intelligence](#)"
- **ānanda** (आनन्द): means "happiness, joy, bliss", "pure happiness, one of three attributes of Atman or Brahman in the Vedanta philosophy".
 - In AI, it is [finding](#), [results](#), and [objectives](#)

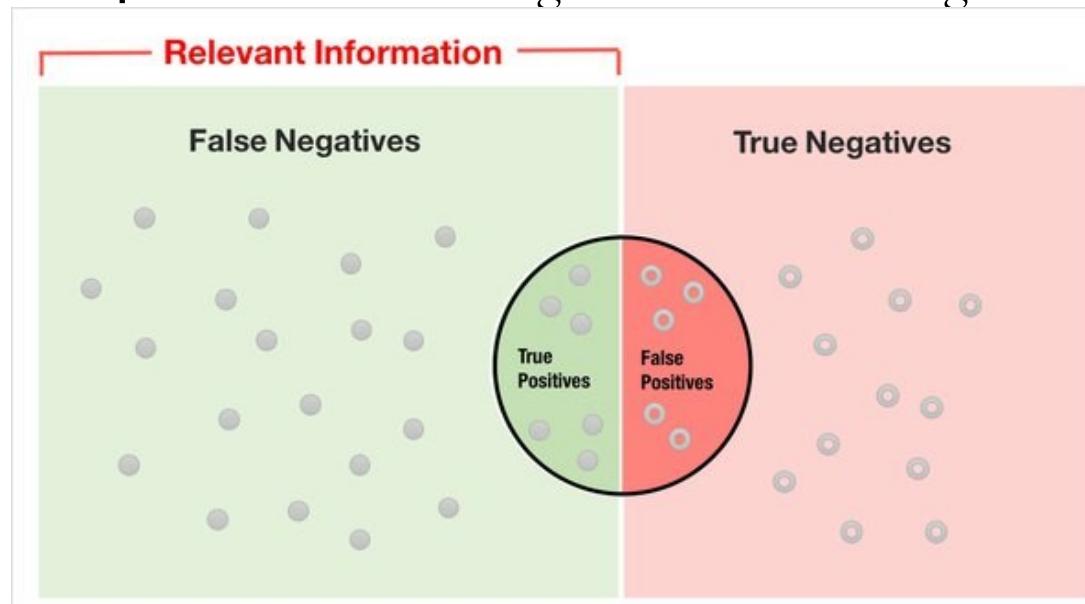
Sat-Cit-Ananda “True-Conscious-Bliss”

How did the Ancient Human’s Neural Network get trained?

- Three divinity (Level of Consciousness or Intelligence)
 - the Bhagavan (भगवान्), God, Lord (different names),
 - the Sādhaka (साधक), Devotee of God, etc.,
 - the Atma (आत्मा), Soul, Spirit, etc.
- Tri (त्रि) Three Loka (लोक) 'Planet' (Classifications)
 - Svarga (स्वर्ग) 'heaven' means 'abode of light' “result of good work (correct classification)”
 - Bhūloka (भूलोक) planet Earth, “result of regular-duties (border-line classification)”
 - Naraka 'Hell' (Yamaloka यम loka), abode of 'Yama' “result of wrong work (incorrect classification)”
- Three divine parts (Human Learning)
 - Dharma (धर्म): "righteousness", "merit", "religious and moral duties". In AI, "goal of environment"
 - Kriyā (क्रिया): "action", "deed", "effort". In AI, "act upon something"
 - Karma (कर्म): "effect or consequence", "work or deed". In AI, "performance or result"

Intelligence of Cosmos

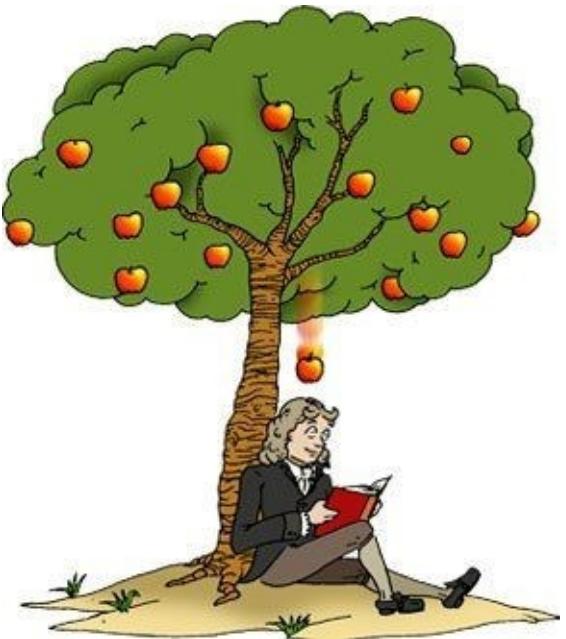
- Brahma ब्रह्म Analogous to [Neural Network](#) of Cosmic Intelligence
- [Intelligent Algorithms](#): consider ancient two terms are Satyam सत्यम् & R̄tam ऋतम्
- Rigveda ऋग्वेद & Atharvaveda अर्थवेद
- Satyam is Eternal Truth
 - Analogous to [True Positive](#) and [True Negative](#)
- R̄tam is the law that governs the working of that Truth



		PREDICTED	
		POSITIVE	NEGATIVE
ACTUAL	POSITIVES	TRUE POSITIVES	FALSE NEGATIVES
	NEGATIVE	FALSE POSITIVES	TRUE NEGATIVES

Universal R̥tam ऋतम् (Law)

- Universal R̥tam in Vaiśeṣika Sutra वैशेषिक documented around 500 BC
 - Objects stay in hand due to their conjunction with the hand. When the conjunction is lost, the object falls down due to its own weight
 - Later attributed to Sir Isaac Newton in the late 16th century with the apple falling from the tree.
 - The **Apple** is remaining conjoint to the tree **Satyam**. **Experience falling down** is also Satyam.
 - Why it **stays conjoint** to the tree is **R̥tam**, “**Inertia**”
 - How it **falls down** is also its **R̥tam**, “**Law of Gravity**”



Darśana दर्शन of Rtam ऋतम् and Satyam सत्यम्

- Darśana “View” of Rtam “Law” and Satyam “Truth”
- Each manifestation in the cosmos, we experience as Truth or Satyam themselves reveal their Rtam
- The order of Created, Destroyed or Interact with each other.
- These Universal Rtams of all cosmic interconnectedness from **micro to macro** and **zero to infinity** have already been available in the Vedas.



Vāda वाद - Theorems, Postulates, or Hypothesis

- Vāda वाद which are challenged, refined, refuted, or accepted over time. For example
- In relation to Advaita अद्वैत means "nondualism"
- Vivartavāda विवर्तवाद & Parināmvāda परिणिमवाद
 - Theory of causation, postulated by post-Shankara Advaita
- Vivarta (विवर्त) means
 - alteration, modification, change of form, altered condition or state.
- Theory of Satkāryavāda means
 - that the effect is pre-existent in the cause
- Parinamavada means
 - the world is a real transformation (parinama means results) of Bramha

Siddhant सिद्धांत means Algorithm

- Conjoint of 2 words सिद्ध + अंत = सिद्धांत
- Siddha सिद्ध which means Perfection, Established, or an Indisputable result.
- Anta अंत which means End or nothing more beyond.
- Siddhant means Cosmic algorithms or Laws which are perfect and will never be challenged or refuted in Past, Present or Future.
- Tridoṣa त्रिदोष Sidhanta says
 - आयुर्वेद के सन्दर्भ में वात, पित्त, कफ इन तीनों को दोष कहते हैं।
 - In the context of Ayurveda, the Vata, Pitta and Kapha are all three doshas.
- There are different layers to Siddhānta in conjunction

Darśana दर्शन of Ṛtam ऋतम् and Satyam सत्यम्

- Neural Networks based on algorithms of their philosophy called Darśana दर्शन
- Darśana are exponential algorithms, which explains the immense possibilities of cosmic creations
- Algorithm in Chāndogya upāniṣat छान्दोग्य उपनिषत्
 - To convert input Clay to an output Pot
 1. We know the clay and we can know all things shaped out of it.
 2. Clay is empirically real or Satyam and the pot which is a modification of clay is also empirically real or Satyam
 3. As we can see and experience clay and pot, the way the clay was shaped to pot is its Ṛtam “the universal law” which when understood can project innate possibilities.
 4. We know nature of creation, Past, Present & Future. This help to compare the creation.



Sanskrit grammar on Speech and Language

- **Phonetics** is a branch of linguistics: similar sounds
 - *Math* (माठ or मथ): To dwell or To think,
 - *Patha* (पाठा): Recitation, *Patha* (पथ) means road = Marg (Hindi) = Path (English)
- Ancient Sanskrit grammarians e.g., Panini mentioned
 - a hymn would be recited as "word1word2 word2word3 word3...."
- *Samhita-patha*: continuous recitation of words bound by phonetic rules of euphonic combination
- *Pada-patha*: a recitation with conscious pause after every word, and after any special grammatical codes;
- *Krama-patha*: a step-by-step recitation where euphonically-combined words are paired successively;
- *Krama-patha* modified: the step-by-step recitation without euphonic-combinations (or free form of each word);
- *other-pāṭha*: methods of recitation of a text and its oral transmission

Raja Bhoja 11th Century

- Robotics



Now

स्वदगहस्तमय मुङ्गरहस्तं कुन्ताहस्तमयवा यदि तत् स्यात् ।
तनिहन्ति विशतो निशि चौरान् द्वारि संवृतमुखं प्रसमेन ॥ १०७

107

khadga hastham - hand equipped with a sword

mudgara hastham - gadha with spikes

kuntha hastham - lance

yadi tatsyath - whatever available

tan nihanti - automatically on his own

vishatho - trying to enter

nishi - at night

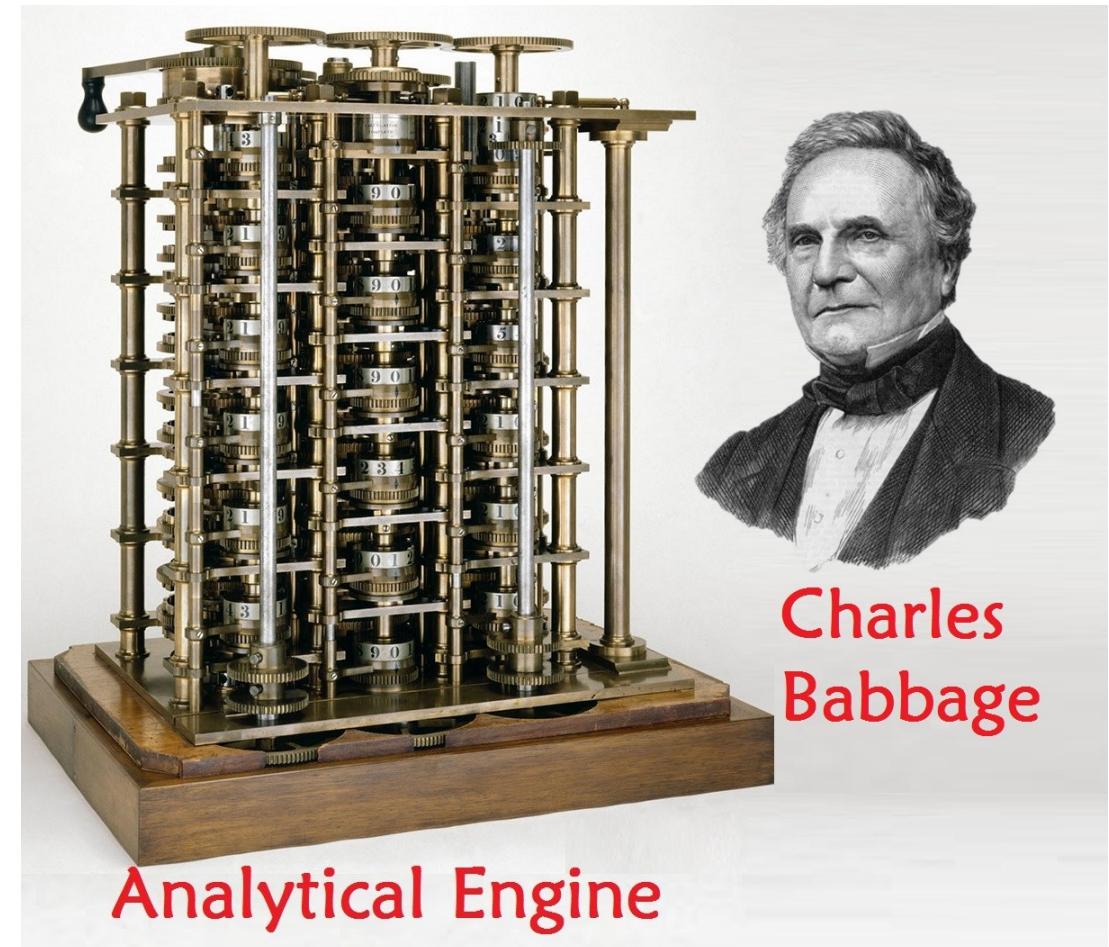
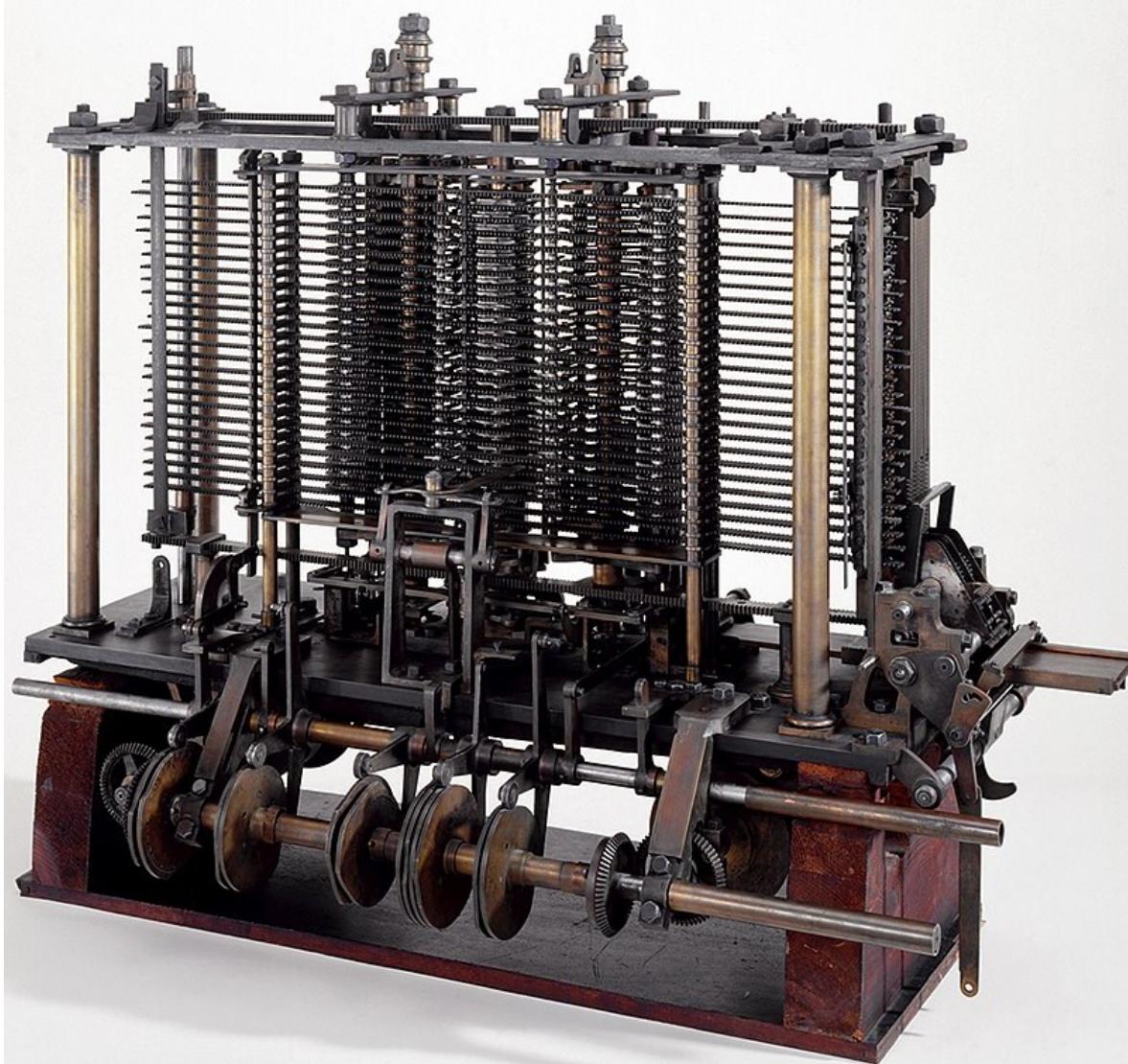
chouran - *chor* - thief

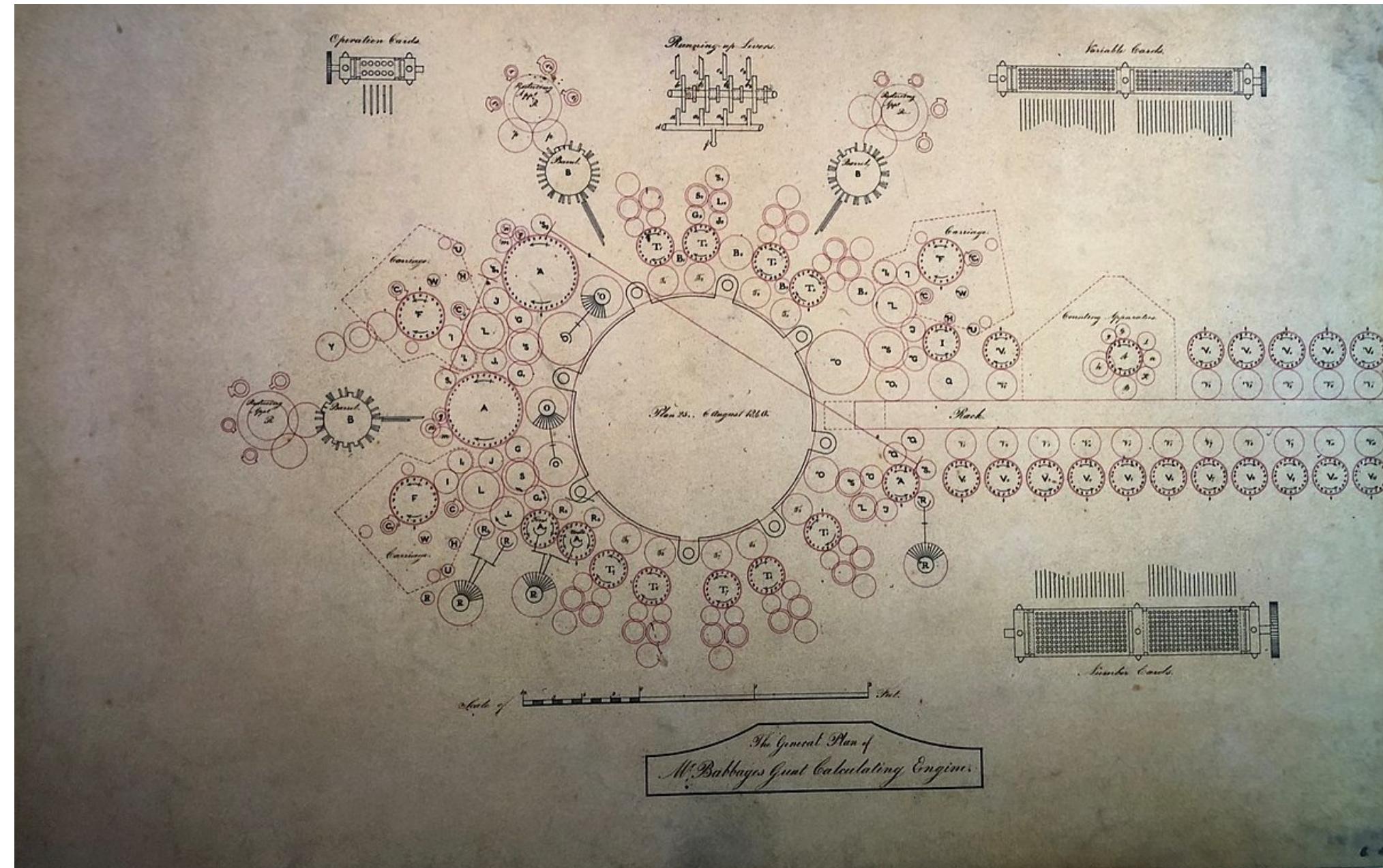
prasabha - outright

sounvrita mukha - covered face, masked face
disguised / hidden

History of Classical AI

1837: Analytical Engine

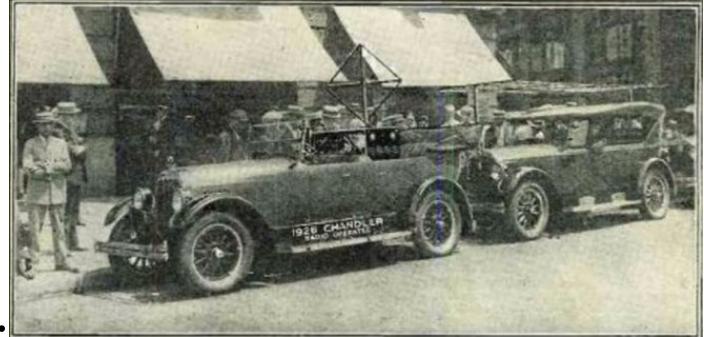


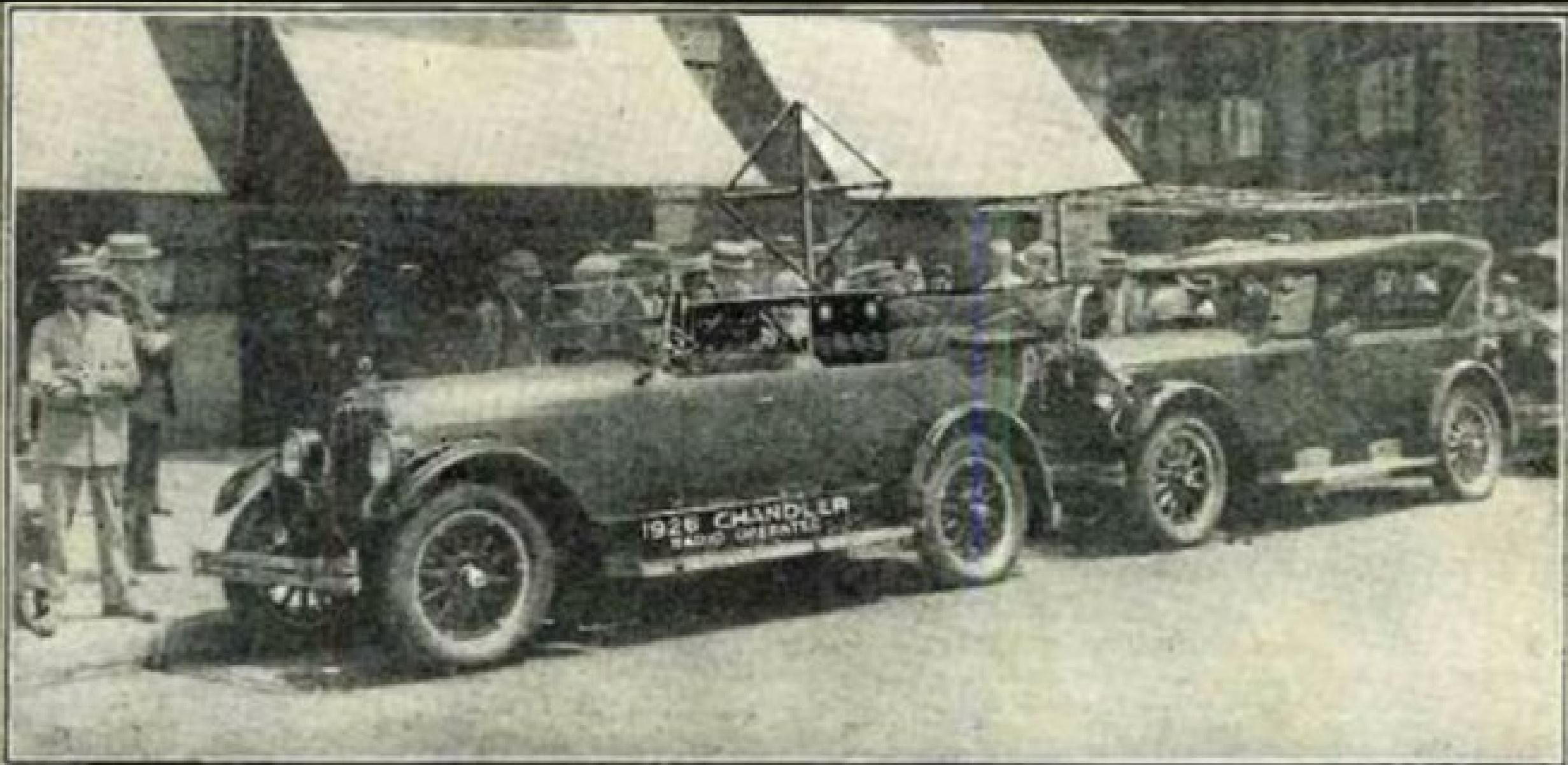


Plan diagram of the **Analytical Engine** from 1840

1925: Houdina Radio Control

- Francis P Houdina developed a radio-operated automobile.
- He equipped a 1926 Chandler with a **transmitting antenna** on the tonneau and operated from a second car that followed it with a **transmitter**.
- The **radio signals** operated small **electric motors** that directed every movement of the car.
- In 1925, he publicly demonstrated his **radio-controlled driver-less car** American Wonder in New York City streets, traveling up Broadway and down Fifth Avenue through thick traffic

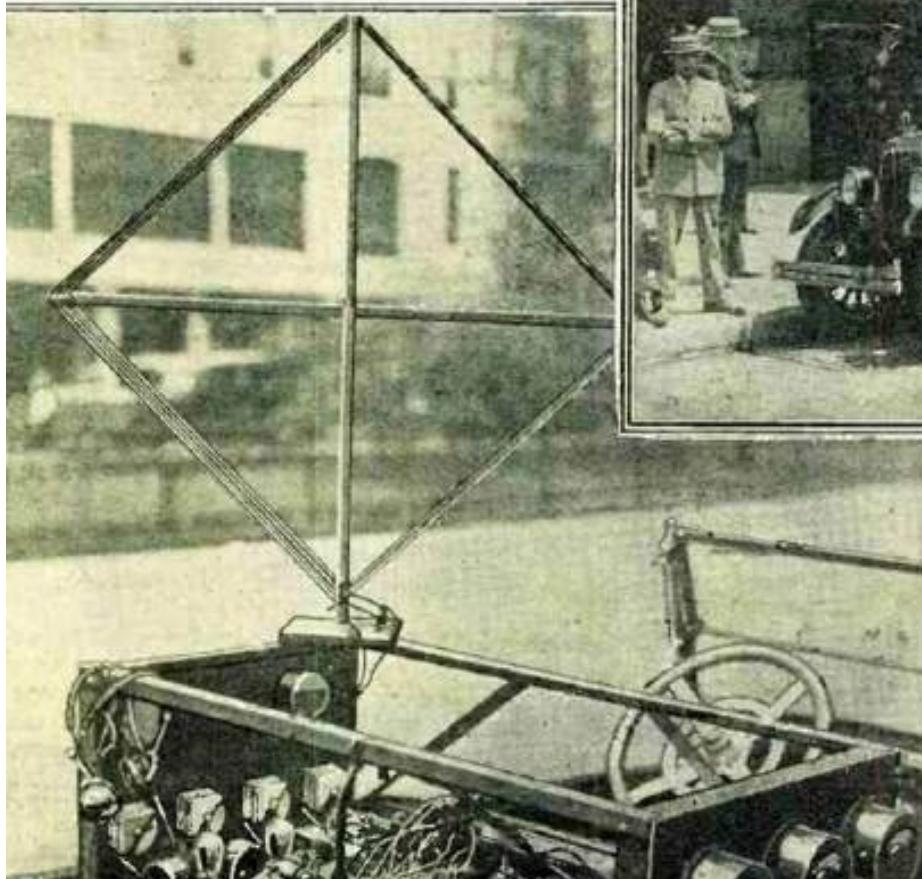




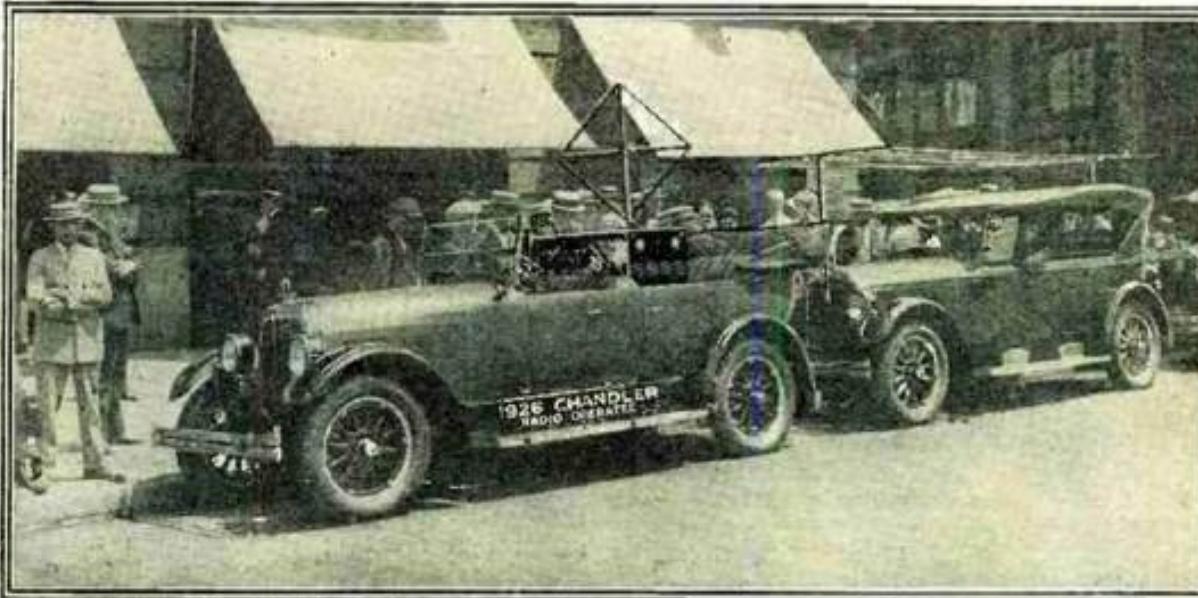
Radio-Controlled Automobile

By HERNDON GREEN

Radio is to control a car in transcontinental tour. The system, which is extremely simple and effective, is fully described here.



Above is shown the two cars used in the transcontinental tour by Mr. Francis P. Houdina. The front car has no driver but is controlled by radio from the second car. Note the transmitting antenna and the receiving loop.



THE DIAGRAM

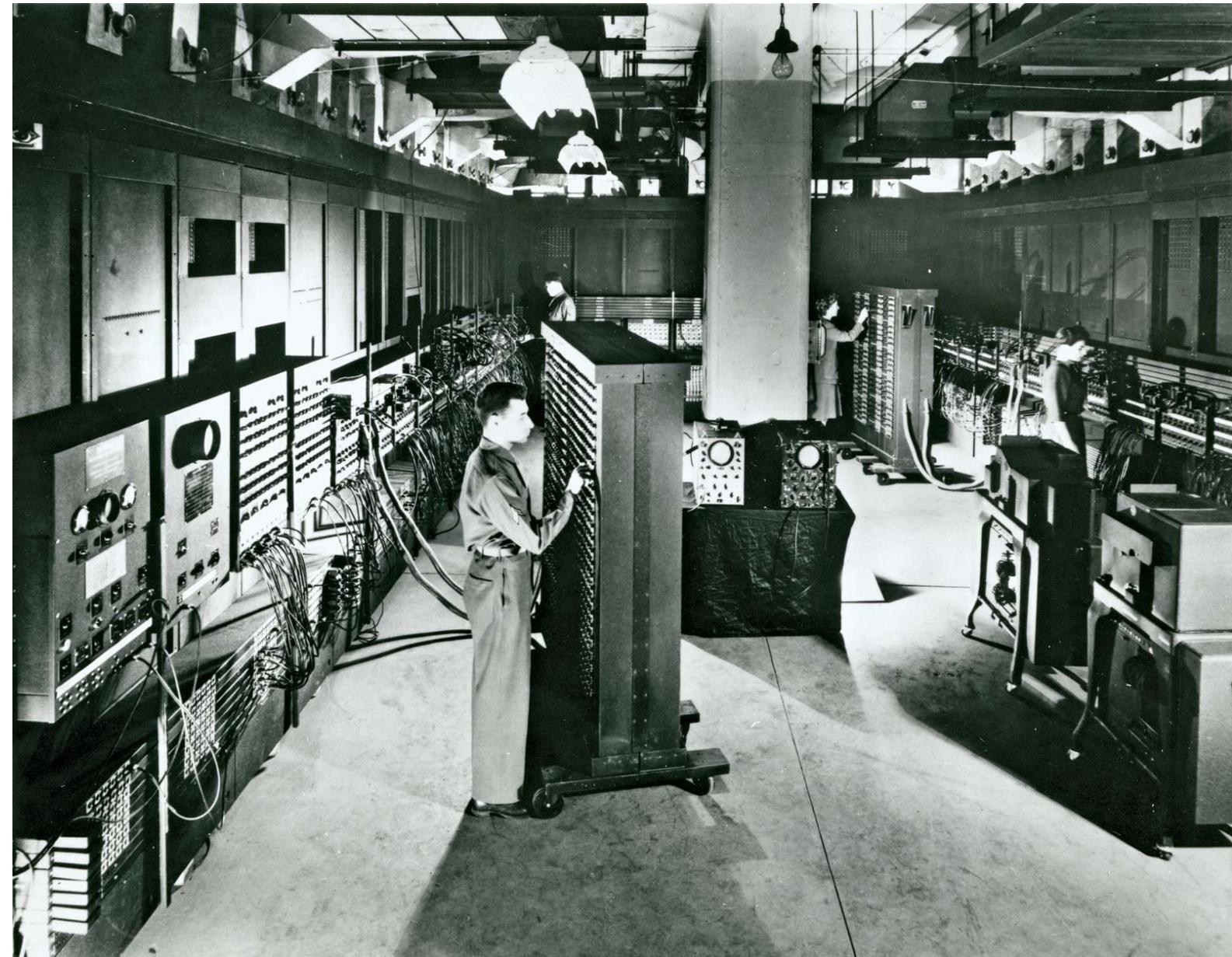
A glance at the wiring diagram given in an accompanying illustration will show the circuit in its simplest form. Some of the amplifying relays are left out in order to simplify the circuits.

The two transmitters are of the usual 10-watt type, using storage battery supply for the filaments and plates. They are housed in the tonneau of the control car together with the power units. The keys controlling them are placed on a small shelf at the right side of the dash.

The receivers are the usual type, employ-

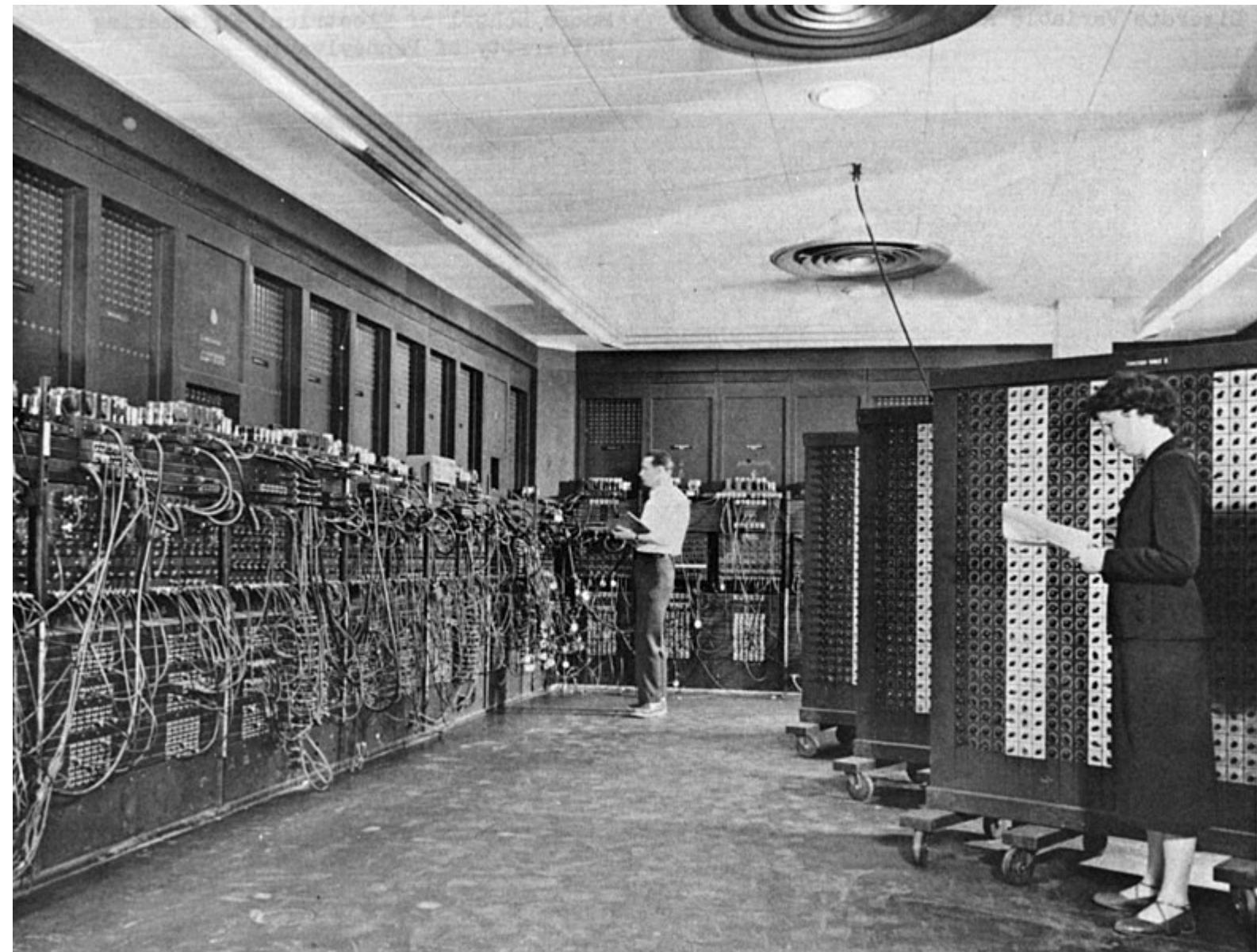
1946-1960

- 1946: ENIAC heralds the dawn of Computing



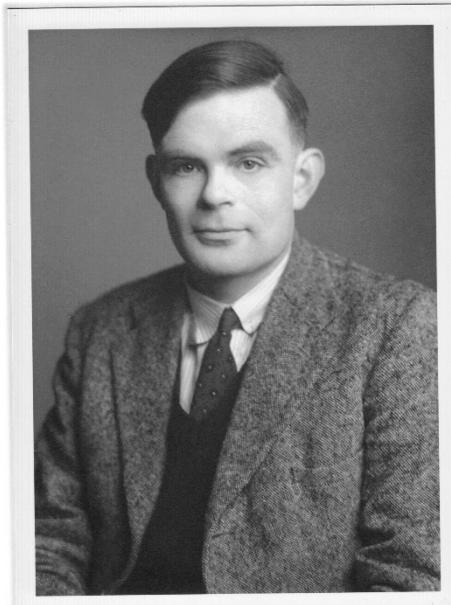
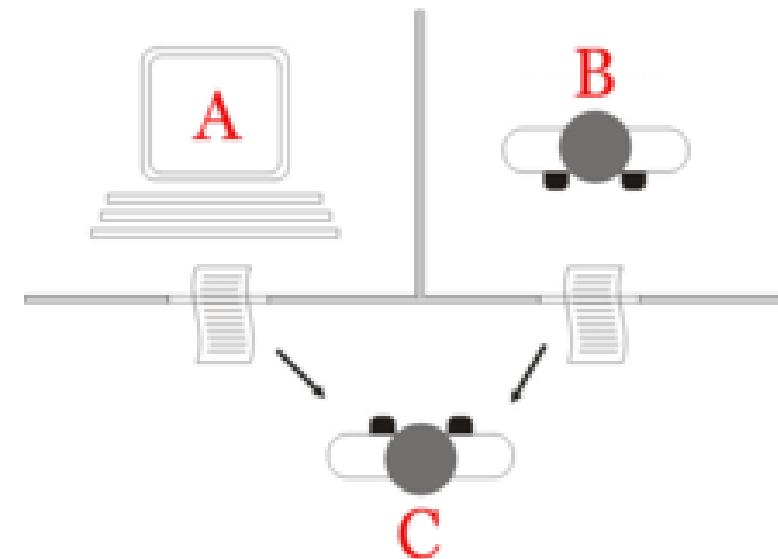
1946-1960

- 1946: ENIAC heralds the dawn of Computing



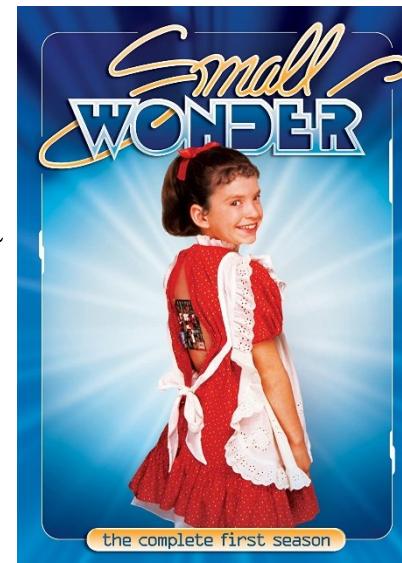
1950: Turing Test

- 1950: Turing asks the question “**Can machines think?**”
- ‘imitation game’ is played with three people,
 - a man (A),
 - a woman (B), and
 - an interrogator (C)
- ‘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 a man and a woman?
- Player C, the interrogator, is given the task to determine which player – A or B – is a computer and which is a human.
- The interrogator is limited to using the responses to written questions to make the determination.



1950: Turing Test

- Turing defined “**intelligent behavior** as the ability to achieve human-level performance in all cognitive tasks, sufficient to fool an interrogator.”
- Intelligent computer
 - interrogated by a human via invisible communication channel,
 - pass the test
 - “if the interrogator fails to guess there is a computer or a human at the other end”
- Popular Science Fictions
 - Small Wonders
 - Karishma ka Karishma



1956: DARTMOUTH RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE

A PROPOSAL FOR THE DARTMOUTH SUMMER RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE

J. McCarthy, Dartmouth College

M. L. Minsky, Harvard University

N. Rochester, I.B.M. Corporation

C.E. Shannon, Bell Telephone Laboratories

August 31, 1955

We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer.

1956: DARTMOUTH RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE

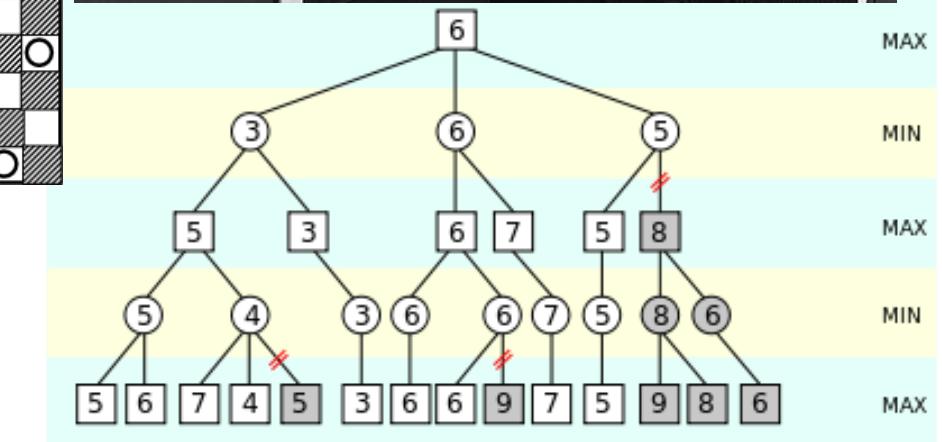
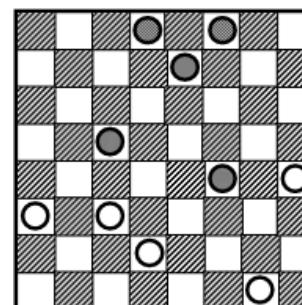
- The following are some aspects of the **artificial intelligence** problem:
 1. **Automatic Computers**
 2. How Can a Computer be Programmed to Use a Language
 3. **Neuron Nets**
 4. Theory of the Size of a Calculation
 5. **Self-Improvement**
 6. Abstractions
 7. Randomness and Creativity
- Application of information theory concepts to **computing machines** and brain models.
- The matched environment - **brain model** approach to automata.

1956: DARTMOUTH RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE

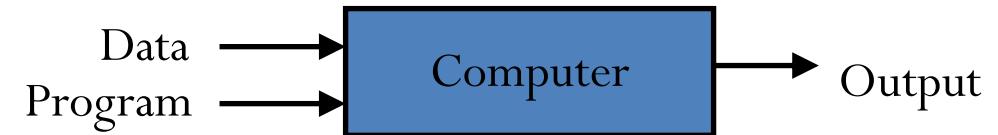
- The stage of programming in a computer.
- Originality in Machine Performance
- The Process of Invention or Discovery
 1. The environment provides data from which certain **abstractions** are formed.
 2. The abstractions together with certain internal habits or drives provide:
 1. A definition of a problem in terms of desired condition to be achieved in the future, **a goal**.
 2. A suggested **action to solve** the problem.
 3. **Stimulation** to arouse in the brain the engine which corresponds to this situation.
 3. Then the engine operates to predict what this environmental situation and the proposed reaction will lead to.
 4. If the prediction corresponds to the goal the individual proceeds to act as indicated.
- The Machine With **Randomness**

1959: Machine Learning

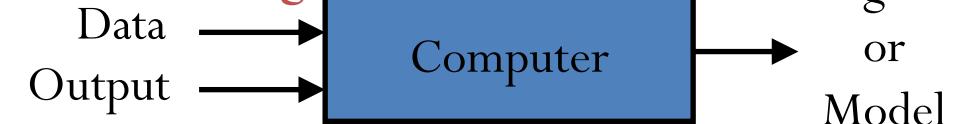
- Arthur Samuel: Automated Game for “Computer checkers” development
- Ability to learn without explicit programming.
 - Samuel implemented alpha-beta pruning
 - Grayed-out subtrees don't need to be explored (when moves are evaluated from left to right),
 - the group of subtrees yields the value of an equivalent subtree or worse, and
 - as such cannot influence the final result.



Traditional Programming



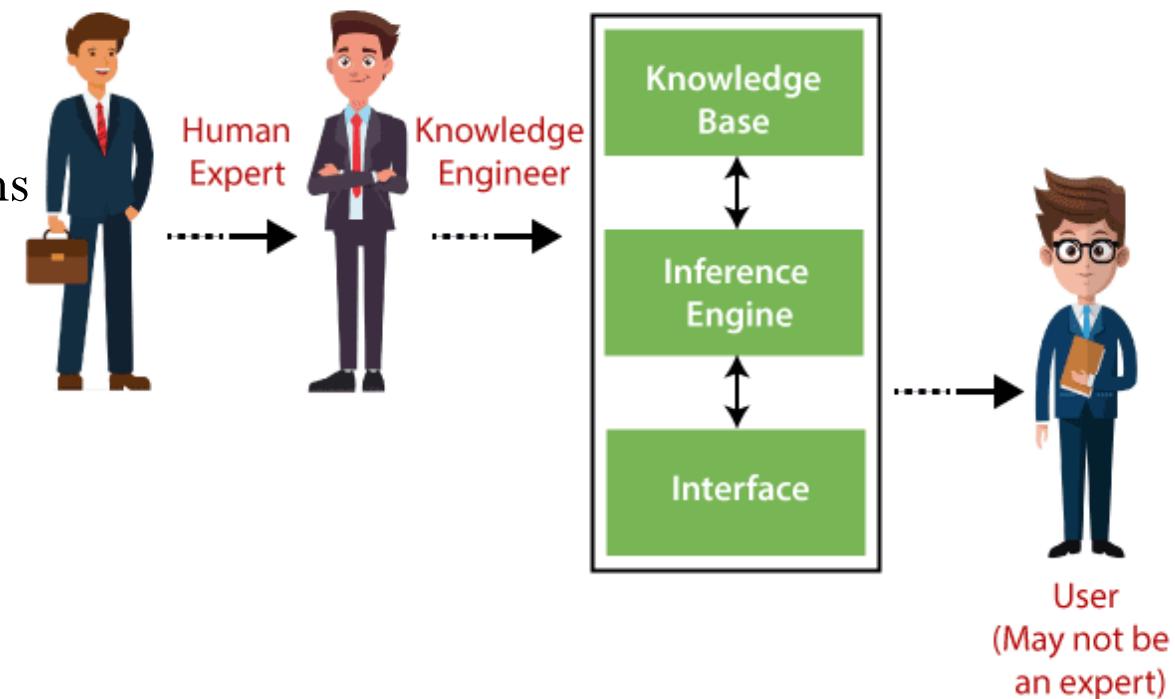
Machine Learning



1960: Expert Systems

Expert Systems are created using following techniques

- Explicit, rules, programs
- Playing Games
- Organic models, Biology recommendations
- Solving word problem in Algebra
- Natural Language processing
- Mobile Robot



1964: Eliza Chatbot

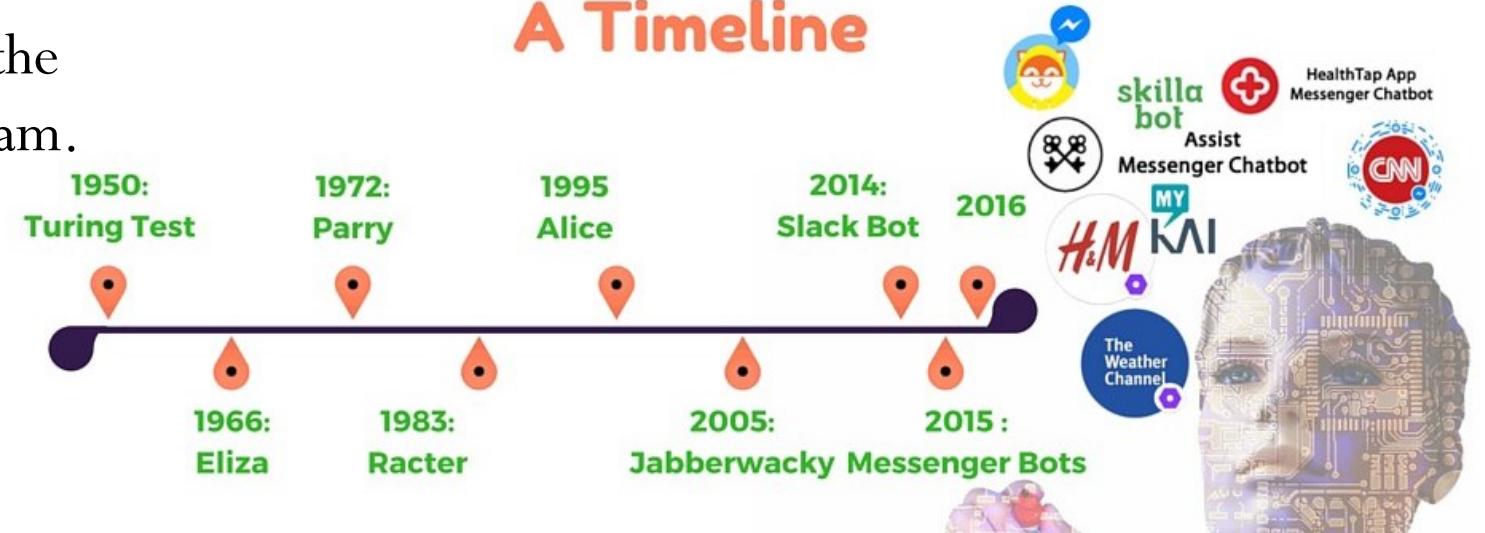
- The chatbot psychotherapist
- Using "pattern matching" and substitution methodology, the program gives canned responses.
- It made early users feel they were talking to someone who understood their input.
- The program was limited by the scripts that were in the program.

Talk to Eliza by typing your questions and answers in the input box.

> Hello, I am Eliza. I'll be your therapist today.
* What is your name?
> Does that question interest you?
* Are you intelligent?
> Why are you interested in whether or not I am intelligent?
* To do Turing Test.
> I'm not sure I understand you fully.
* Do you know Turing Test?
> Oh... Let's move on to something else for a bit. know Turing Test?

TYPE HERE

History of Chatbots: A Timeline



1966: Shakey general purpose mobile robot

- First mobile robot with the ability to perceive and reason about its surroundings
- Tasks required **planning**, **route-finding**, and the rearranging of simple objects.



AI Winter (1974-80)

- Abandonment of AI: There were two major winters in 1974–1980 and 1987–1993
- 1966: failure of [machine translation](#)
- 1970: abandonment of [connectionism](#)
- Period of overlapping trends:
 - 1971–75: [DARPA](#)'s frustration with the [Speech Understanding Research](#) program at [Carnegie Mellon University](#)
 - 1973: large decrease in AI research in the United Kingdom in response to the [Lighthill report](#)
 - 1973–74: DARPA's cutbacks to academic AI research in general
- [1974 – 1980: AI Winter #1](#)
 - Failure of machine translation
 - Negative results in Neural nets
 - Poor speech understanding

AI Winter (1974-80)

- Problem with A.I.
- Applications were primarily for toy problems
- Low computational power
- Combinatorial explosion too many exhaustive Rules
- Common-sense is nearly impossible to program
- Minsky's book = Perceptron showed limitations of simple neural network
- Loss of government funding in A.I.

AI Spring and Summer

AI Spring (1980-87)

- Expert systems used in several real-world applications
- Driverless car
- Hopfield network popularization of backpropogation
- Minsky (1984) – “Winter is coming!”

AI Winter (1987-93)

- 1987 – 1993: AI Winter #2
- 1987: collapse of the LISP machine market
 - Decline of LISP
- 1988: cancellation of new spending on AI by the Strategic Computing Initiative
- 1990s: end of the Fifth Generation computer project's original goals
- 1993: resistance to new expert systems deployment and maintenance
 - Decline of specialized hardware for expert systems

AI Winter (1987-93)

- Popularity of the P.C.
- Lack of attractive results
- Funding cuts
- **Lasting effects**
 - “Artificial Intelligence is associated with systems that have all too often failed to live up to their promises.”
 - “Some believe the word 'robotics' actually carries a stigma that hurts a company's chances at funding.”
- **Winter** over due to
- 1993: **Association Rule Mining** by Rakesh Agrawal and Srikant Ramakrishnan

AI Spring and Summer

AI Summer (1994-2000)

- Apriori algorithm: Association Rule Mining by Rakesh Agarwal in 1993 and 1995
- 1997 Deep Bule beats Kasparov in chess
- Theory – Including probability, information theory, optimization etc.
- Moore's Law – Rapid growth of processing power.
 - “Number of transistor doubles every two years”

History of Modern AI

1993: Apriori algorithm (Association Rule Mining)

- Let $I = \{i_1, i_2, \dots, i_m\}$ be a set of literals, called items.
- *Support* of a rule $X \rightarrow Y$ is the percentage of transactions that contain both X and Y .
- *Confidence* of a rule is percentage the *if-then statements* ($X \rightarrow Y$) are found true
- Find all rules that satisfy a user-specified *minimum support* and *minimum confidence*

TID	Transaction Items
1	Bread, Jelly, PeanutButter
2	Bread, PeanutButter
3	Bread, Milk, PeanutButter
4	Beer, Bread
5	Beer, Milk



$\{\text{Bread}\} \rightarrow \{\text{PeanutButter}\}$ (Sup = 60%, Conf = 75%)
 $\{\text{PeanutButter}\} \rightarrow \{\text{Bread}\}$ (Sup = 60%, Conf = 100%)
 $\{\text{Beer}\} \rightarrow \{\text{Bread}\}$ (Sup = 20%, Conf = 50%)
 $\{\text{PeanutButter}\} \rightarrow \{\text{Jelly}\}$ (Sup = 20%, Conf = 33.33%)
 $\{\text{Jelly}\} \rightarrow \{\text{PeanutButter}\}$ (Sup = 20%, Conf = 100%)
 $\{\text{Jelly}\} \rightarrow \{\text{Milk}\}$ (Sup = 0%, Conf = 0%)

Rakesh Agrawal, Tomasz Imieliński, and Arun Swami. "Mining association rules between sets of items in large databases." SIGMOD. 1993.

Ramakrishnan Srikant, and Rakesh Agrawal. "Mining Generalized Association Rules." VLDB 1995.

1993: Apriori algorithm (Association Rule Mining)

- Let $I = \{i_1, i_2, \dots, i_m\}$ be a set of literals, called items.
- *Support* of a rule $X \rightarrow Y$ is the percentage of transactions that contain both X and Y .
- *Confidence* of a rule is percentage the *if-then statements* ($X \rightarrow Y$) are found true
- Find all rules that satisfy a user-specified *minimum support* and *minimum confidence*
 - 90% of transactions that purchase *Bread* and *Butter* (antecedent) also purchase *Milk* (consequent). The number 90% is the confidence factor of the rule
 - $[Bread], [Butter] \rightarrow [Milk]$ 90%
 - 98% of customers who purchase *Tires* and *Auto accessories* also buy some *Automotive services*; here 98% is called the confidence of the rule.
 - $[Auto Accessories], [Tires] \rightarrow [Automotive Services]$ 98%

Rakesh Agrawal, Tomasz Imieliński, and Arun Swami. "Mining association rules between sets of items in large databases." SIGMOD. 1993.

Ramakrishnan Srikant, and Rakesh Agrawal. "Mining Generalized Association Rules." VLDB 1995.

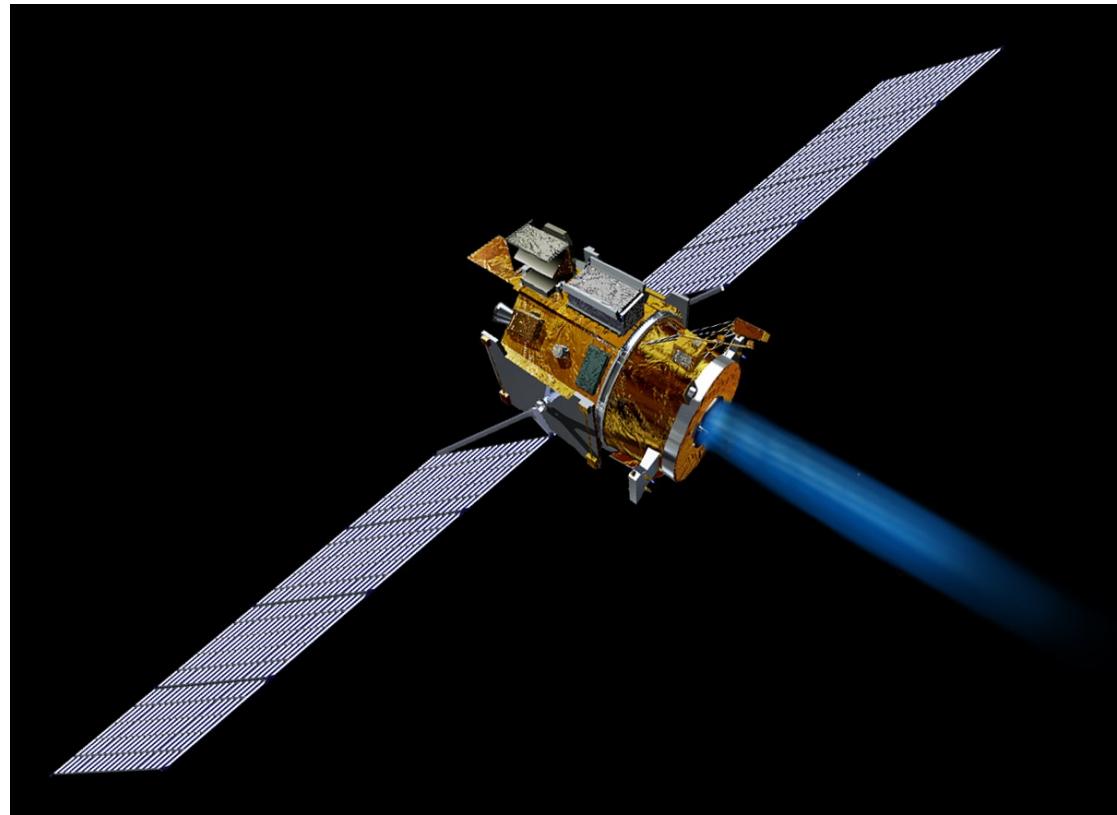
1997: Deep Blue ends Human Supremacy in Chess

- I could feel **human-level intelligence** across the room -Gary Kasparov, World Chess Champion (human)
- In a few years, even a single victory in a long series of games would be the triumph of human genius.
- Does Deep Blue use AI? “If it works, its not AI!”
- “Saying Deep Blue doesn’t really think about chess is like saying an airplane doesn’t really fly because it doesn’t flap its wings”. – Drew McDermott



1998: Deep Space 1

- 1998: Deep Space 1
- For two days in May 1999, an AI program called “Remote Agent” autonomously ran Deep Space 1 (some 60,000,000 miles from earth)
- The spacecraft’s main goal was to test 12 high-risk technologies including
 - ion propulsion,
 - autonomous optical navigation,
 - a solar power concentration array, and
 - a combination miniature camera-imaging spectrometer.



Key Dates

Oct. 24, 1998: Launch

July 29, 1999: DS1 passed by asteroid 9660 Braille

Sept. 18, 1999: Primary mission ended

Sept. 22, 2001: DS1 entered the coma of Comet Borrelly

Dec. 18, 2001: DS1's ion engine is finally turned off

Space AI progress

- Image analysis for Road paths, large-object detection, galaxy, planet or star classification



A screenshot of a software interface for satellite imagery analysis. On the left, a sidebar titled "User Selection" contains input fields for SW Latitude (51.46694144864419), SW Longitude (46.1985234375), NE Latitude (51.49335472541077), NE Longitude (46.23321533203126), and a "Detection Model" dropdown set to "Large Aircraft". It also includes checkboxes for "Display Suggested AOIs" and "Score Threshold" (set to 9), and input fields for "NMS Threshold" (set to 1), "Start Date" (mm/dd/yyyy), and "End Date" (mm/dd/yyyy). A blue button at the bottom right says "Get". On the right, a satellite image shows several aircraft on a runway, with one aircraft highlighted by a green box. A callout bubble labeled "Close-up View of Detected Aircraft" points to this highlighted area. The Lockheed Martin logo is in the top right of the interface.

Space AI progress

- INTELLIGENT NAVIGATION SYSTEM to help astronauts find their way on the planets
- AI-BASED ASSISTANTS AND ROBOTS to help astronomers in their long space travel
- AUTONOMOUS ROVERS that roam the surface of other planets (currently the surface of Mars).



Space AI progress

- AI'S ROLE IN SPACE EXPLORATION such as charting unnoted galaxies, stars, black holes, and studying cosmic events, as well as communication, autonomous StarCraft navigation, monitoring and system control.



Space AI progress

- Other Use Cases: Astronaut assistants, Mission design and planning, Satellite data processing, Space debris, Navigation systems
- To make decisions and avoid obstacles on the rough surface.
- Determining the best route without specific commands from the mission control.
- AI4Mars project to outline and identify different rock and landscape features.

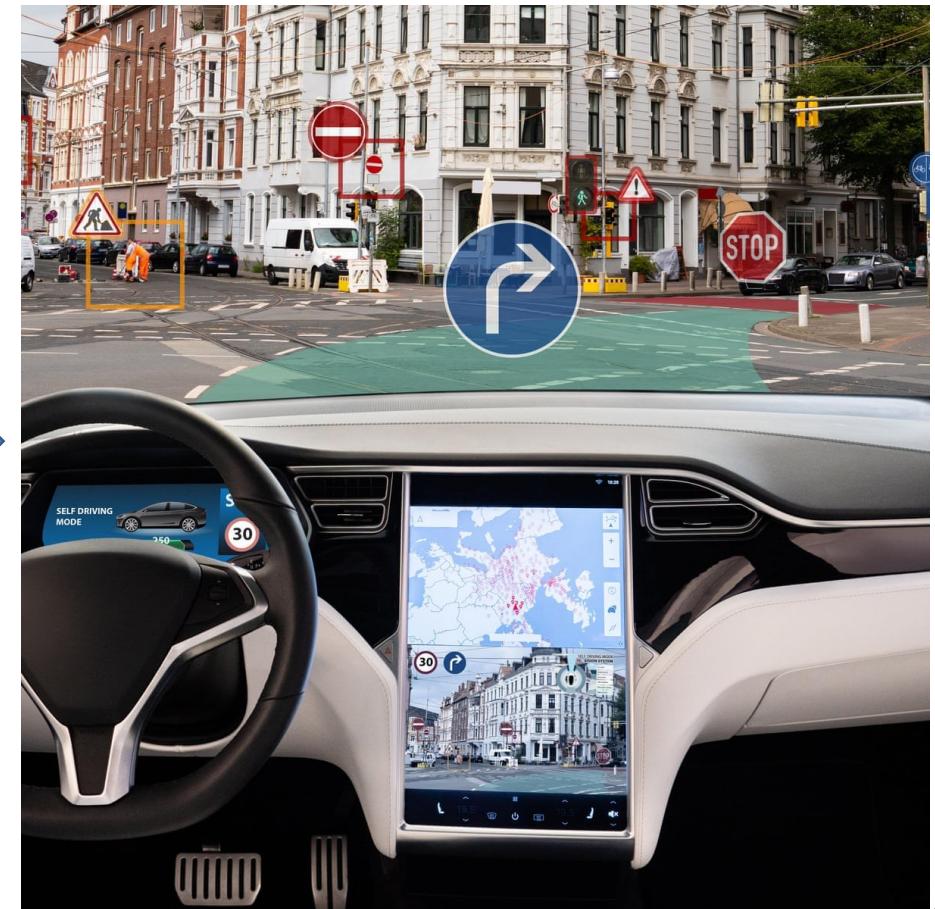


2005 onward: Driverless Cars

- DARPA Grand Challenge
- 2005 Autonomous driving for 135 miles in mountain road of desert
- Stanley and three other cars drive themselves



Now

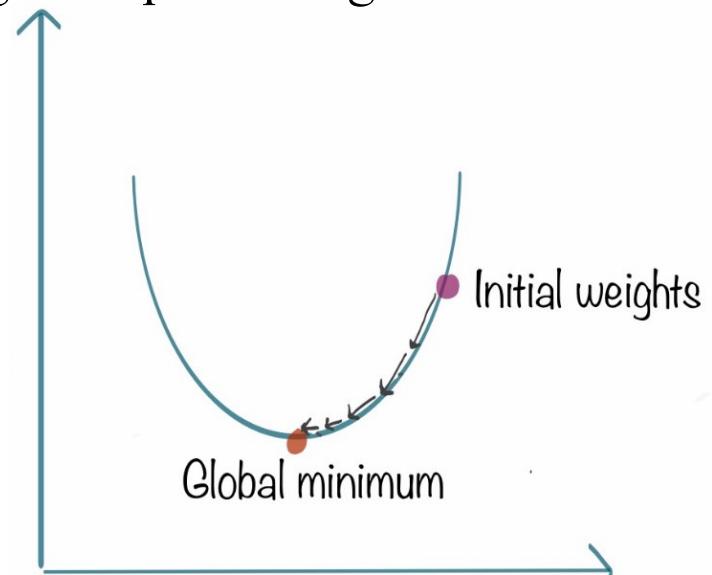


2000-12: Big Data, GPU, Deep Learning, NLP

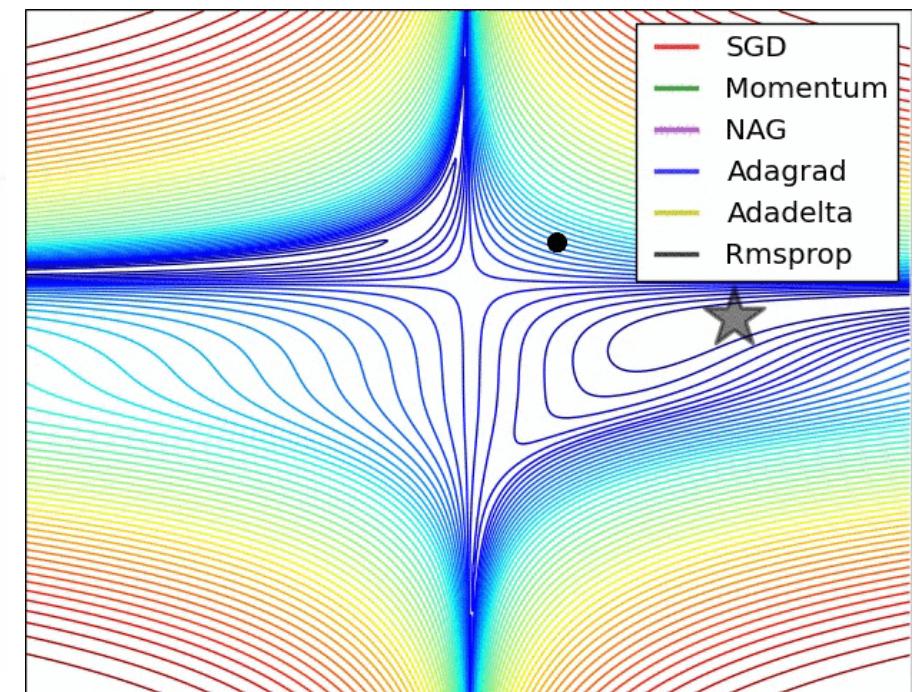
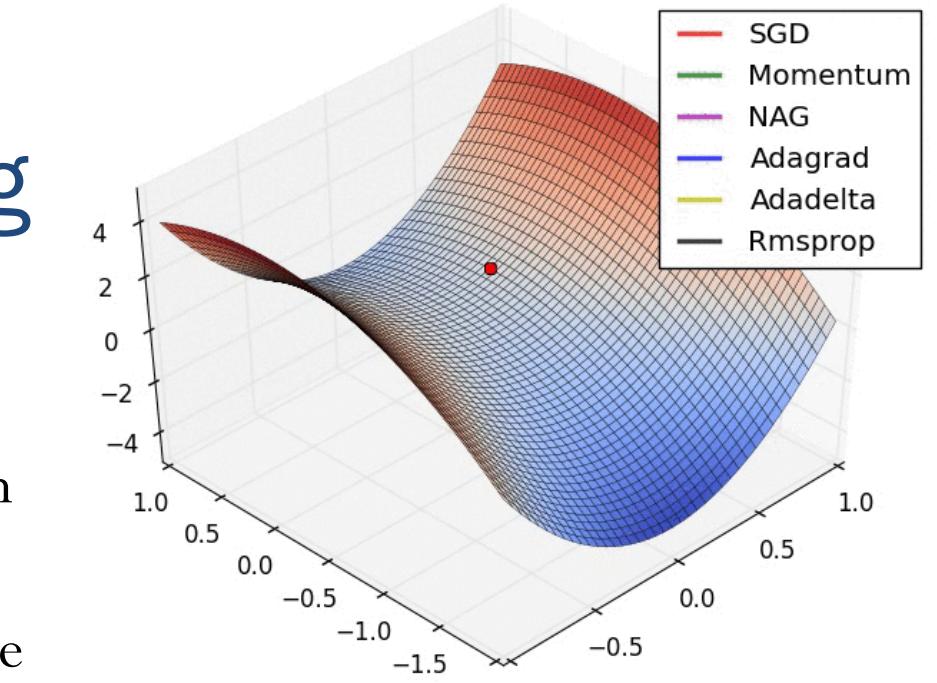
- Google is popular (Big File, Google File Systems, Page Rank Algorithm etc.)
- Internet boom
- Shifted emphasis to big data – statistical techniques
- Graphical Processing Units (GPUs) inventions
- Spectacular results in Deep Neural Networks
- Specific goal oriented research expectations (Speech, Text, Image, Video etc)
- IBM's Watson beat the Jeopardy champions.
 - Due to Advance Natural Language Processing (NLP) and Q&A.

2000 onward: Deep Learning

- For the Probabilities and Expectation equations
 - an optimization algorithm finds the value of the parameters (weights) that minimize the error when mapping inputs to outputs.
 - Optimization algorithms affect the accuracy and the speed of the training a deep learning model.

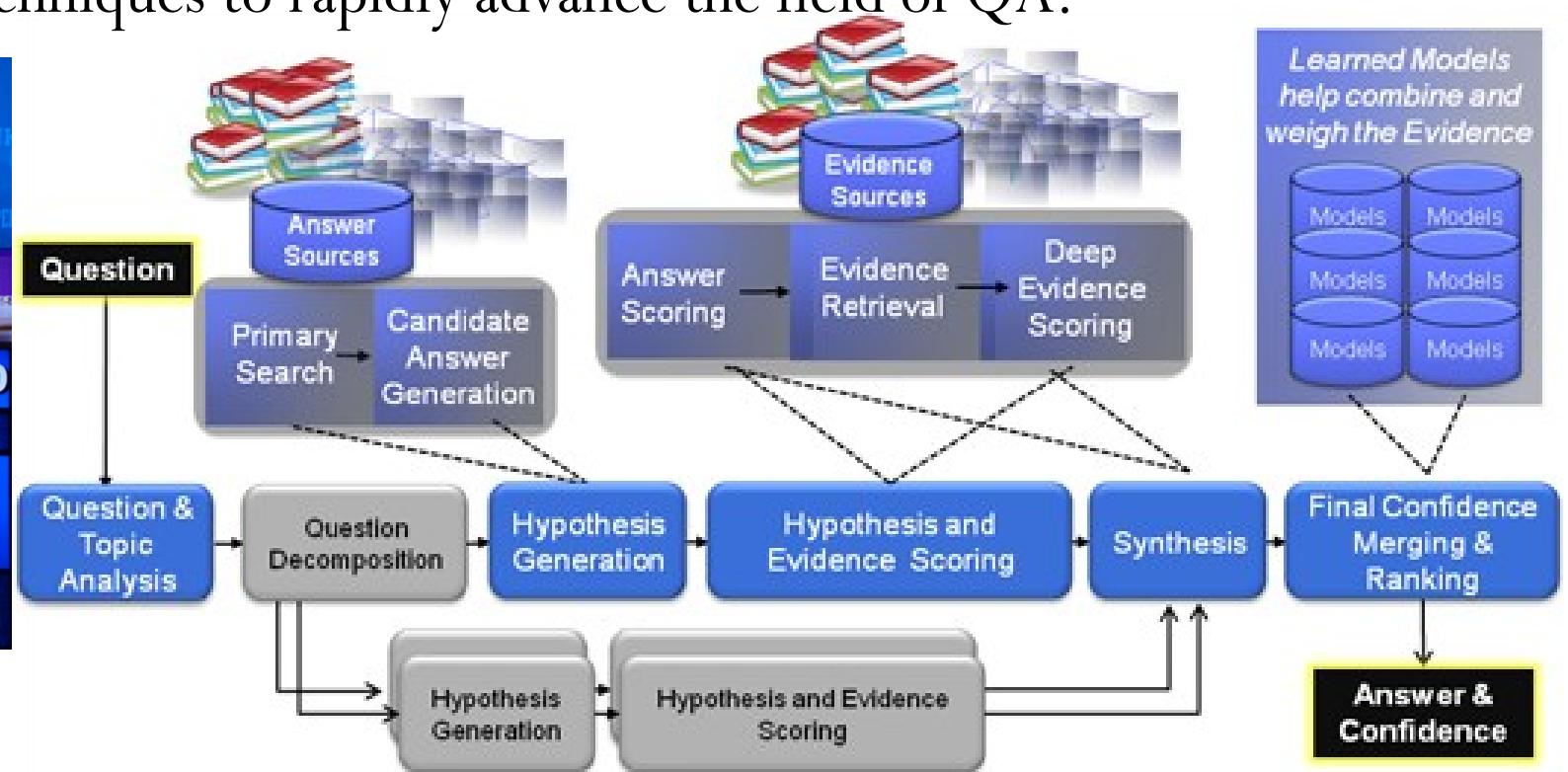


<https://awesomopensource.com/project/Jaewan-Yun/optimizer-visualization>



2011: IBM Watson beaten Human in Jeopardy

- Watson is performing at human expert levels in terms of precision, confidence, and speed at the Jeopardy quiz show.
- DeepQA architecture for combining, deploying, evaluating, and advancing a wide range of algorithmic techniques to rapidly advance the field of QA.



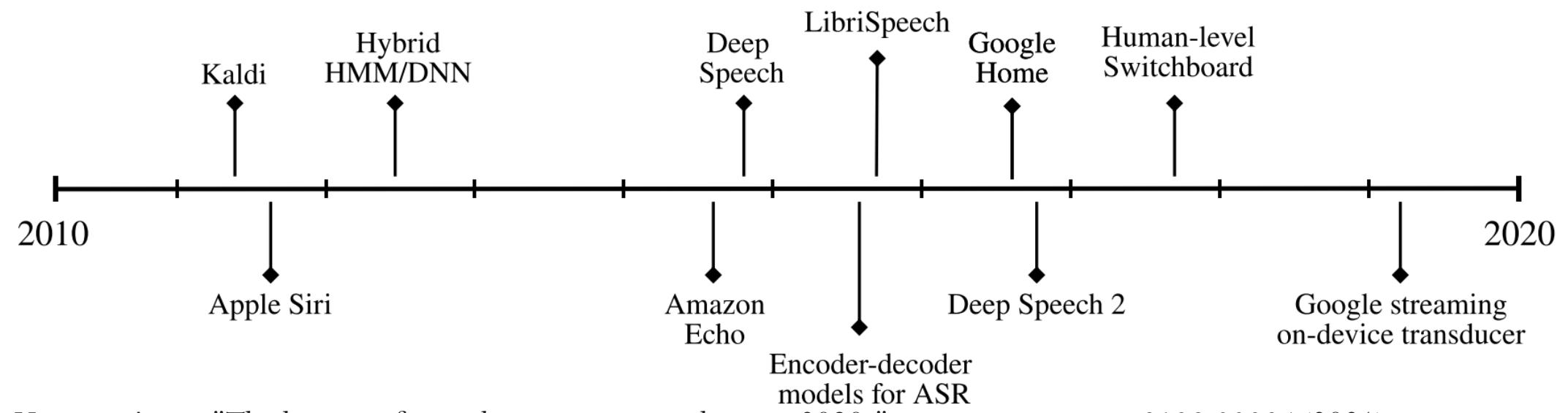
https://researcher.watson.ibm.com/researcher/view_group_subpage.php?id=2159

2010 onward: Natural Language - Speech

- Amazon Alexa
- Apple Siri
- Microsoft Cortana
- Google translate
- Voice recognition on Google Search, Whatsapp, etc.

Open-source and speech recognition software like Kaldi, and larger benchmarks like LibriSpeech are widely used.

ChaptGPT (2023)

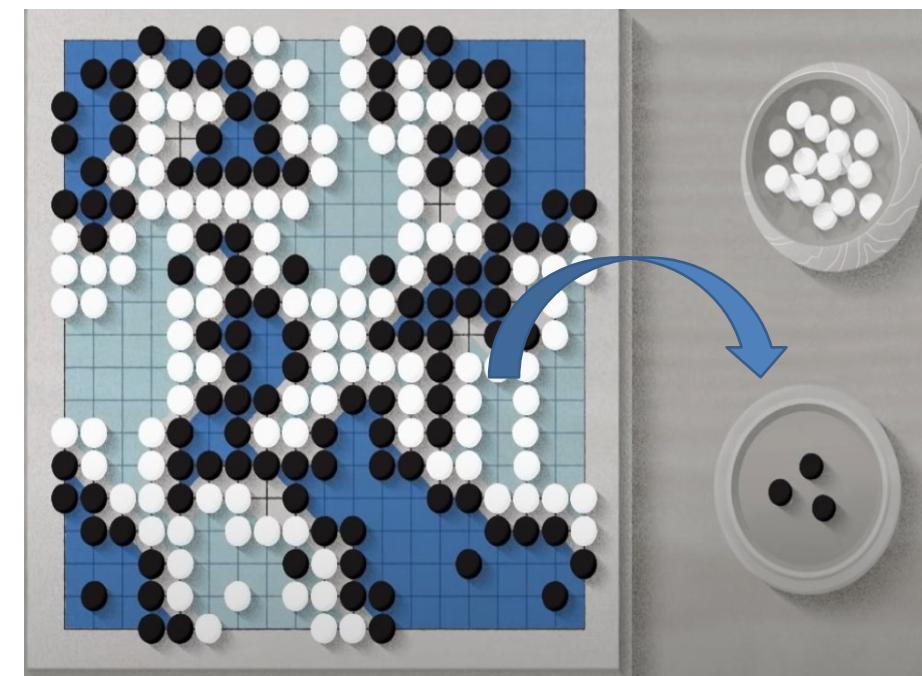
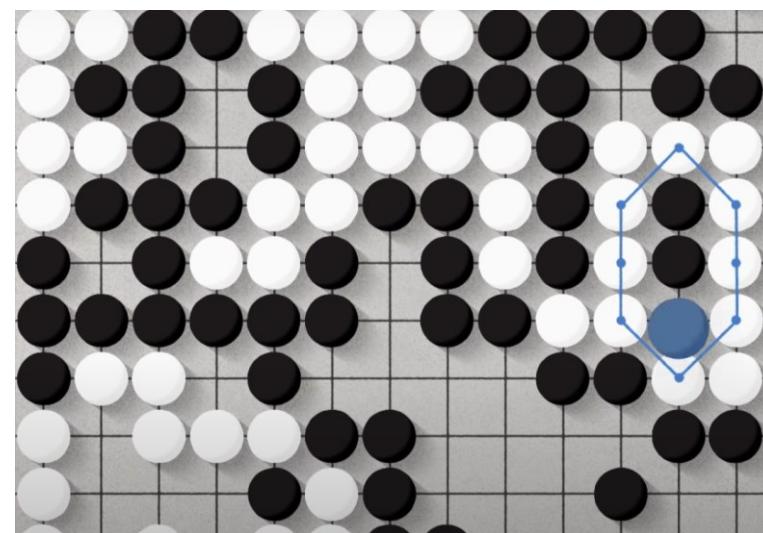
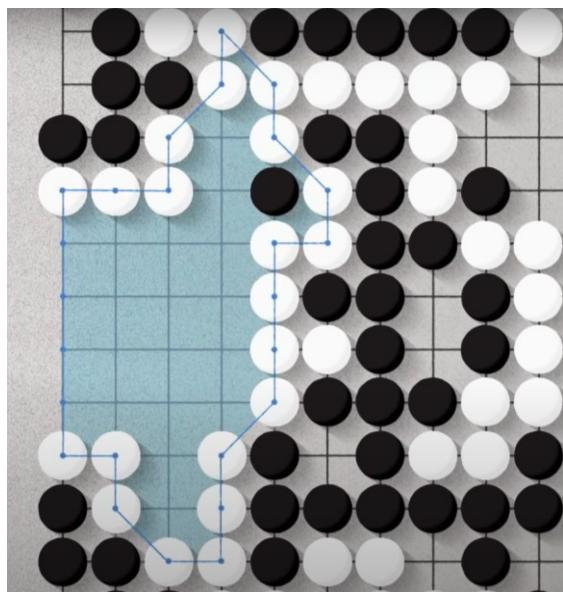


2012 onward: AI Spring

- Private funding companies: Microsoft, Google, Facebook, IBM, AWS etc
- Fast development of computational power
- Large volume of data generation
- Availability of Big data: Growth of data
- Better data science, data mining, deep learning, and machine learning
- Better automated gaming due to Reinforcement Learning
- Better technology and huge computational power
- Cloud based Shared resources (Software and Hardware)
- Better algorithms, Self-supervised Learning, Open source software development

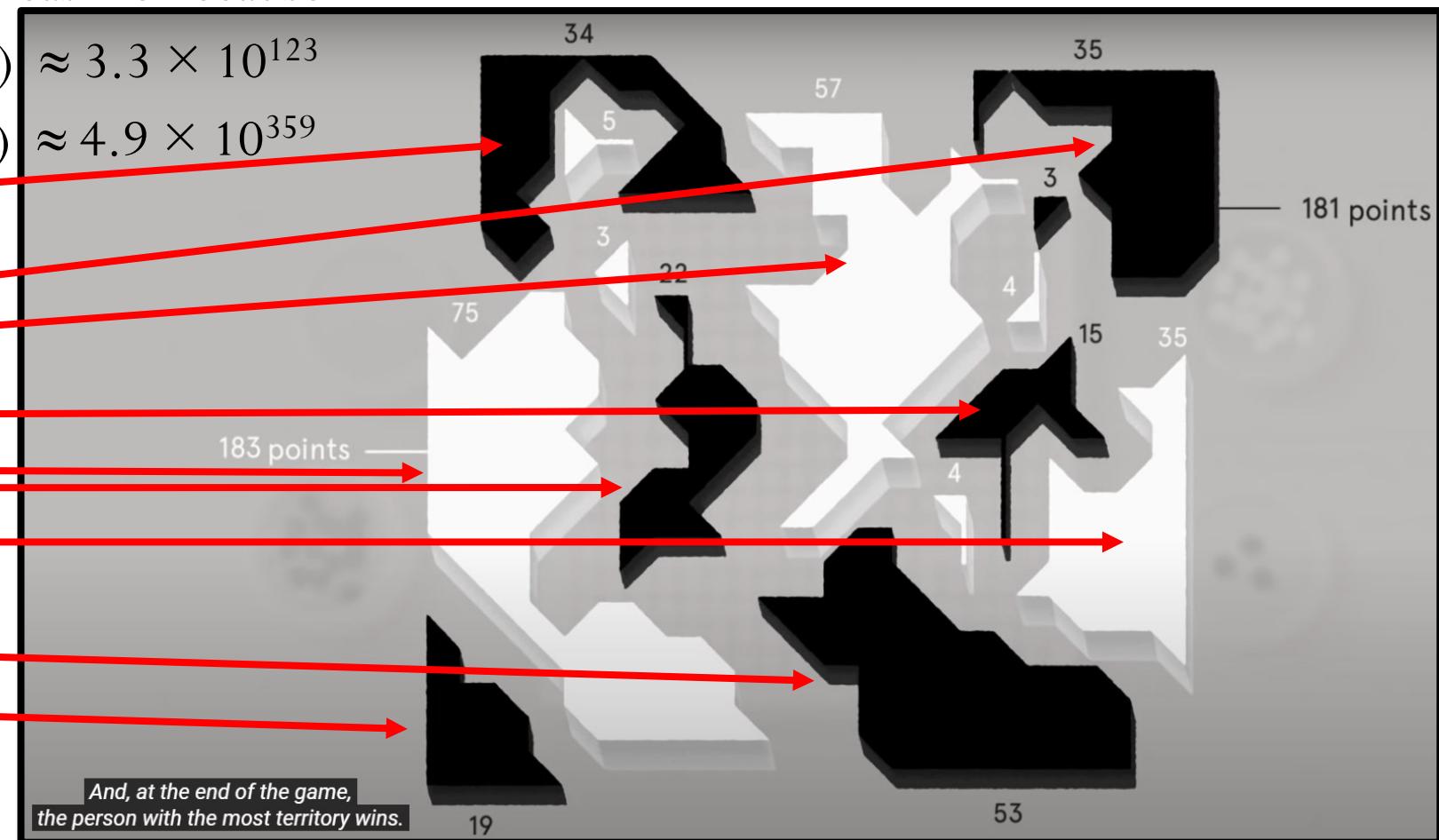
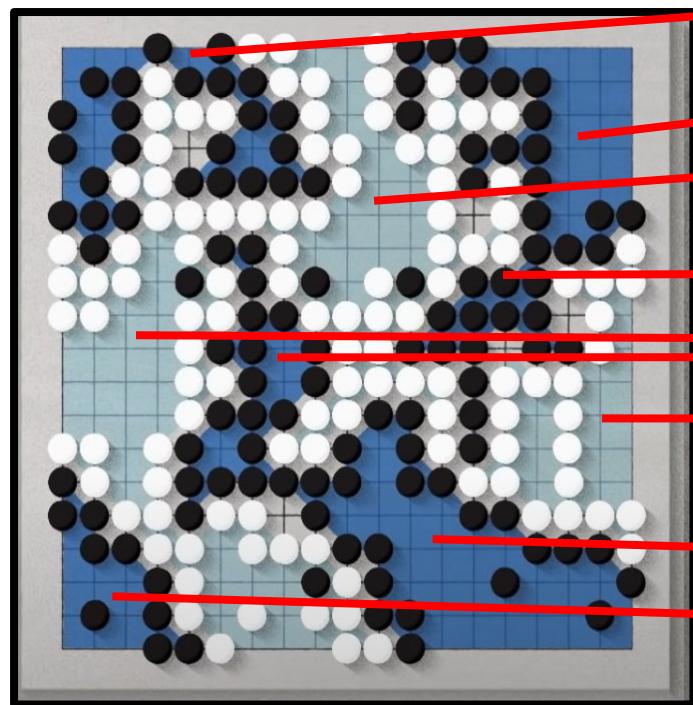
2016: Alpha Go

- Make territories
- Capture opponent pieces
- Largest territory



2016: Alpha Go

- Large games, search tree containing approximately b^d possible sequences of moves
 - exhaustive search is infeasible because
 - Chess ($b \approx 35, d \approx 80$) $\approx 3.3 \times 10^{123}$
 - Go ($b \approx 250, d \approx 150$) $\approx 4.9 \times 10^{359}$



2016: Alpha Go

- AlphaGo achieved a 99.8% winning rate against other Go programs, and defeated the human European Go champion by 5 games to 0.
- This is the first time that a computer program has defeated a human professional player in the full-sized game of Go.



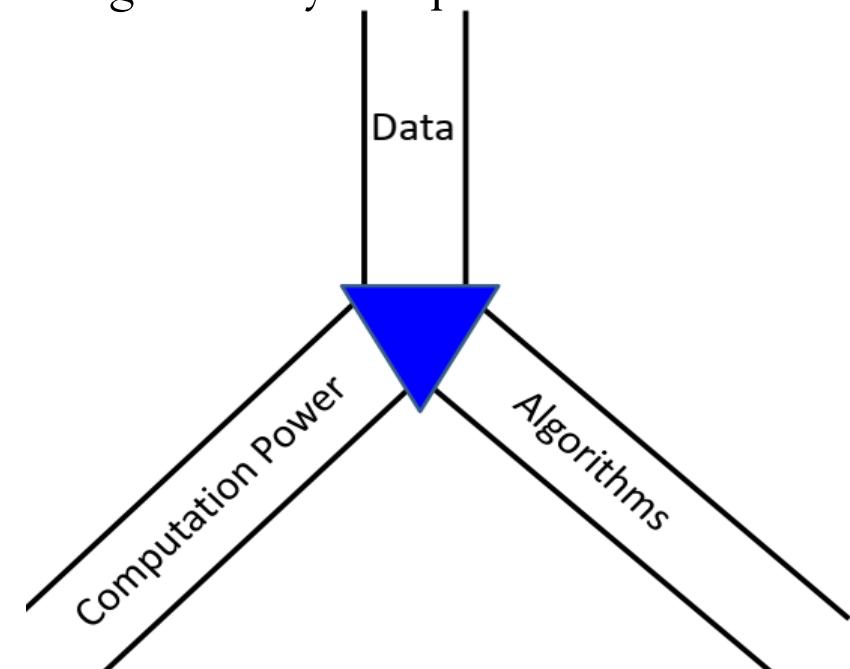
Improvement resulted better AI

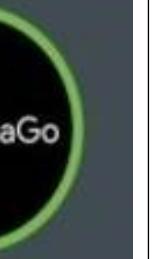
Problems

- sharp minima (leading to overfitting), non-robustness
 - analysis corresponds too closely or exactly to a particular set of data
 - therefore, fail to fit to additional data or predict future observations reliably
- numerical instability (vanishing/exploding gradients)
 - generally desirable property of numerical algorithms for solving ordinary and partial differential equations by discrete approximation

Why does AI work so well because

- high capacity (susceptible to overfitting)
- Better and more training Data
- Better, large, scalable, parallel Computational Power
- Better, parallel and distributed Algorithms
- Slowly but steadily increasing emphasis on
 - Explainability and Interpretability
 - Theoretical justifications

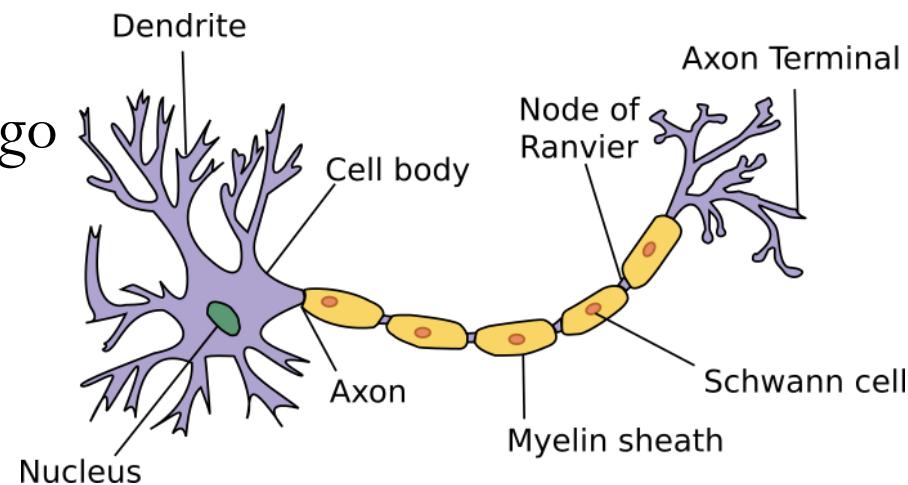
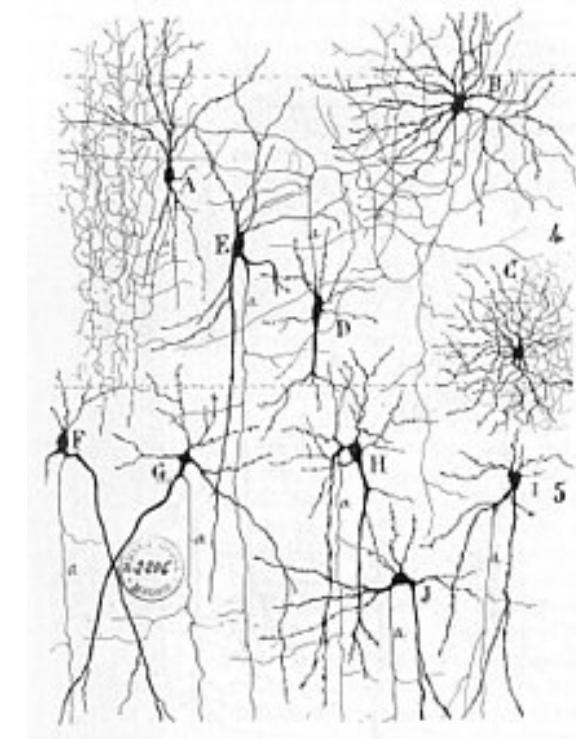


	1950	TURING TEST Computer scientist Alan Turing proposes a test for machine intelligence. If a machine can trick humans into thinking it is human, then it has intelligence		1955	A.I. BORN Term 'artificial intelligence' is coined by computer scientist John McCarthy to describe "the science and engineering of making intelligent machines"		1961	UNIMATE First industrial robot, Unimate, goes to work at GM replacing humans on the assembly line		1964	ELIZA Pioneering chatbot developed by Joseph Weizenbaum at MIT holds conversations with humans		1966	SHAKEY The 'first electronic person' from Stanford, Shakey is a general-purpose mobile robot that reasons about its own actions		A.I. WINTER Many false starts and dead-ends leave A.I. out in the cold		1997	DEEP BLUE Deep Blue, a chess-playing computer from IBM defeats world chess champion Garry Kasparov		1998	KISMET Cynthia Breazeal at MIT introduces Kismet, an emotionally intelligent robot insofar as it detects and responds to people's feelings	
	1999	AIBO Sony launches first consumer robot pet dog AIBO (AI robot) with		2002	ROOMBA First mass produced autonomous robotic vacuum cleaner from		2011	SIRI Apple integrates Siri, an intelligent virtual assistant with a voice		2011	WATSON IBM's question answering computer Watson wins first place		2014	EUGENE Eugene Goostman, a chatbot passes the Turing Test with a third		2014	ALEXA Amazon launches Alexa, an intelligent virtual assistant with a voice		2016	TAY Microsoft's chatbot Tay goes rogue on social media making		2017	ALPHAGO Google's A.I. AlphaGo beats world champion Ke Jie in the complex

History of Artificial Neural Network (ANN)

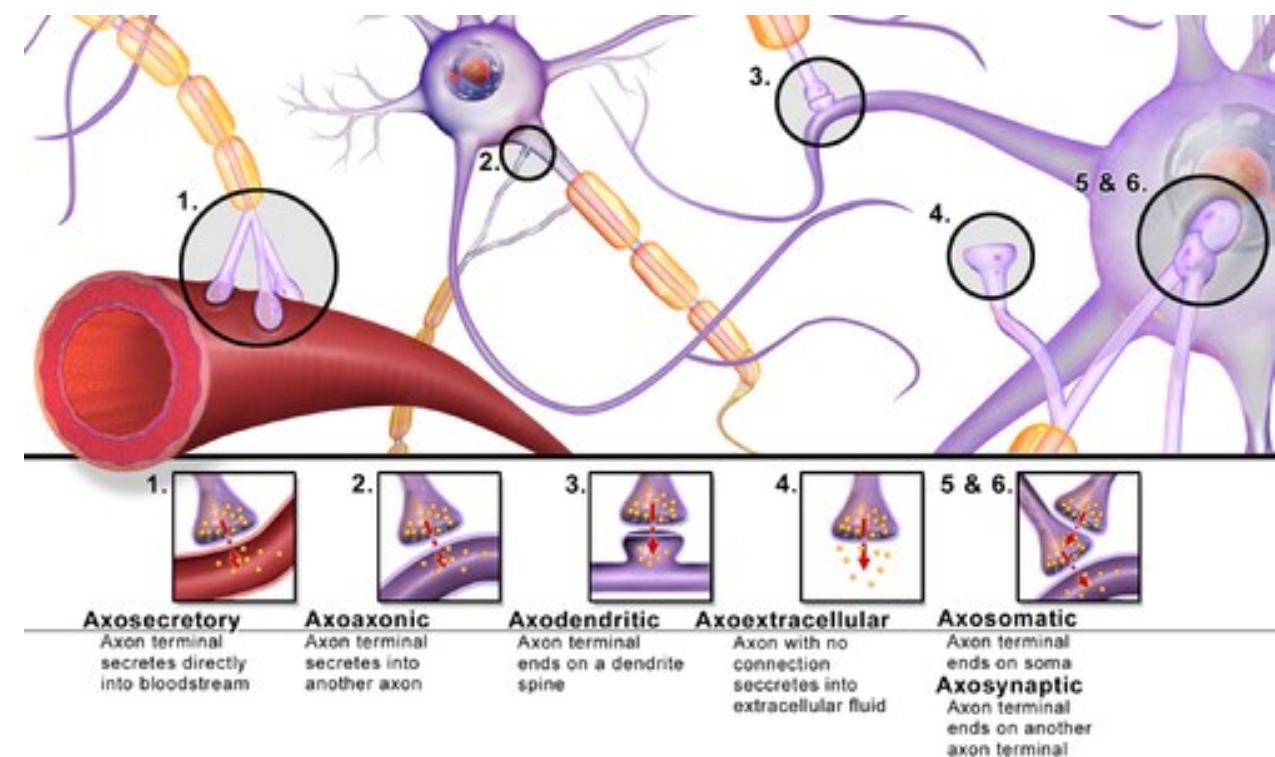
Neural Network (Anatomy)

- **1870s:** Reticular Theory “nervous system and brain is a single continuous network, all nerve cells in the nervous system constituted a continuous, interconnected network”
- **1888:** Neurons “fundamental units of the brain and nervous system, each nerve cell is an independent entity and nerve synapses transfer nerve impulses from one cell to another”
- **1906:** Nobel prize to both Camillo Golgi and Santiago Ramón y Cajal "in recognition of their work on the structure of the nervous system."



Neural Network (Anatomy)

- Synapse is a structure that permits a neuron (or nerve cell) to pass an electrical or chemical signal to another neuron or to the target effector cell.
- Synapses are essential to the transmission of nervous impulses from one neuron to another.
- Different types of synapses



Perceptron

- 1943 Perceptron was invented by **McCulloch and Pitts**
- 1958 Mark I Perceptron hardware developed and constructed by **Frank Rosenblatt**

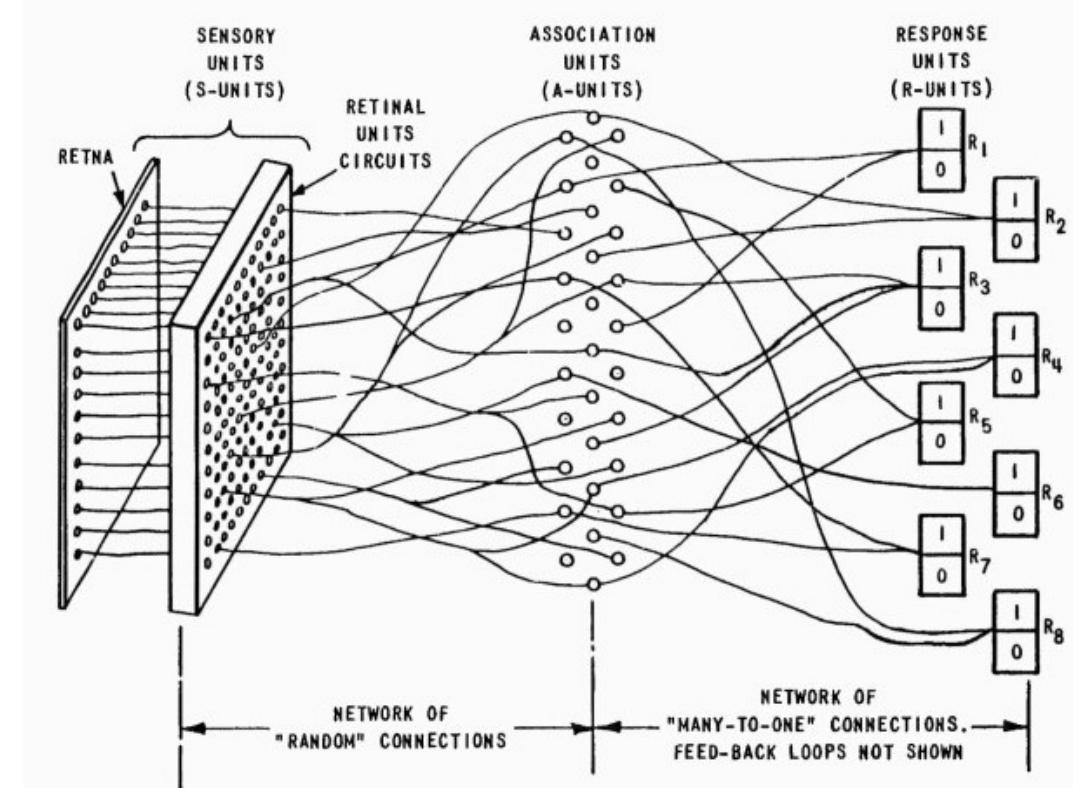
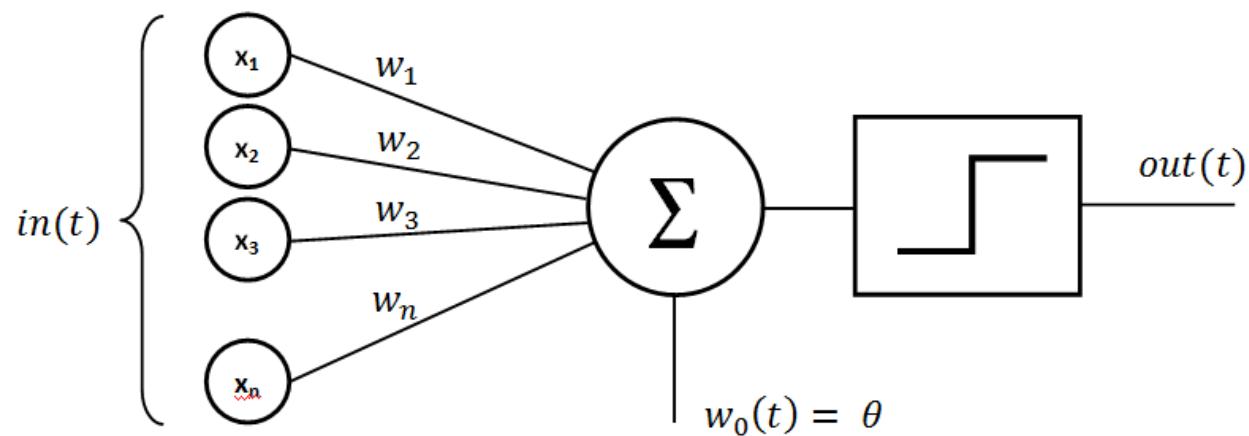
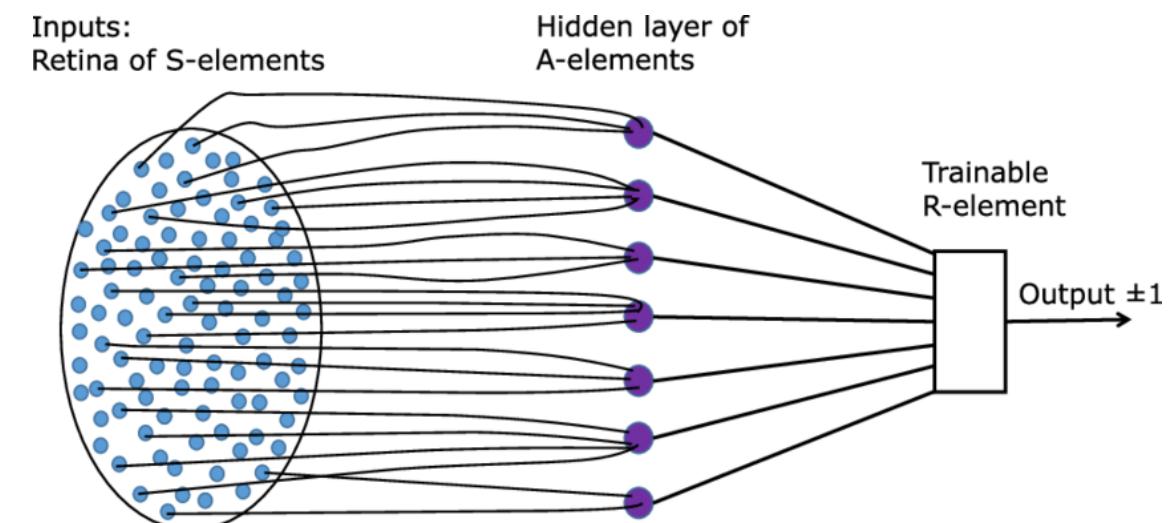
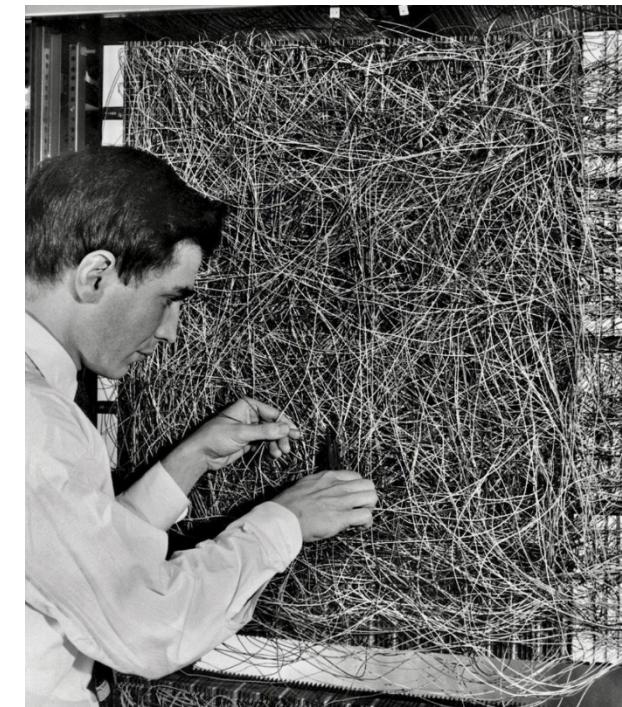


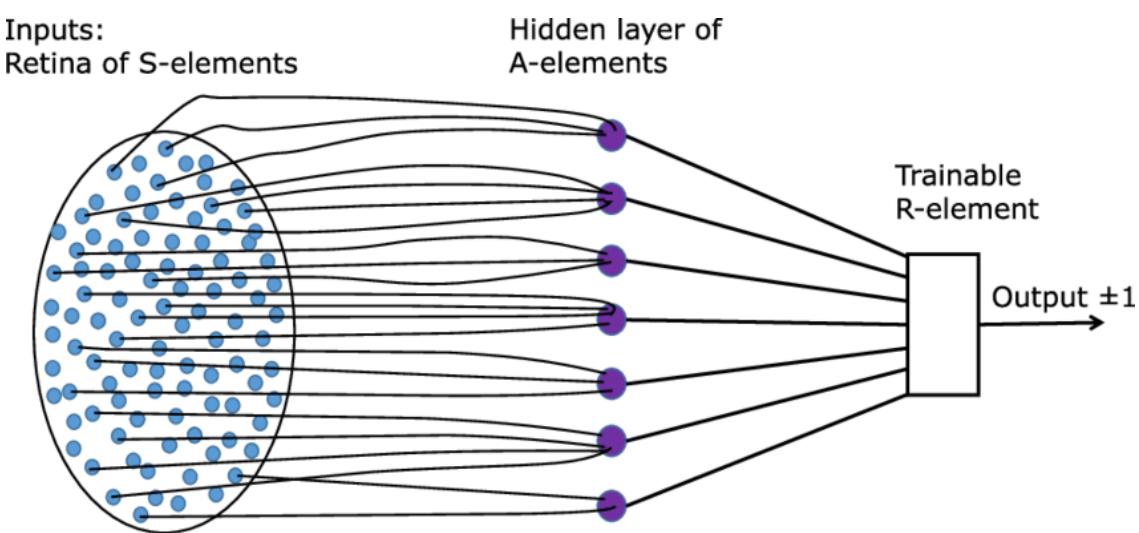
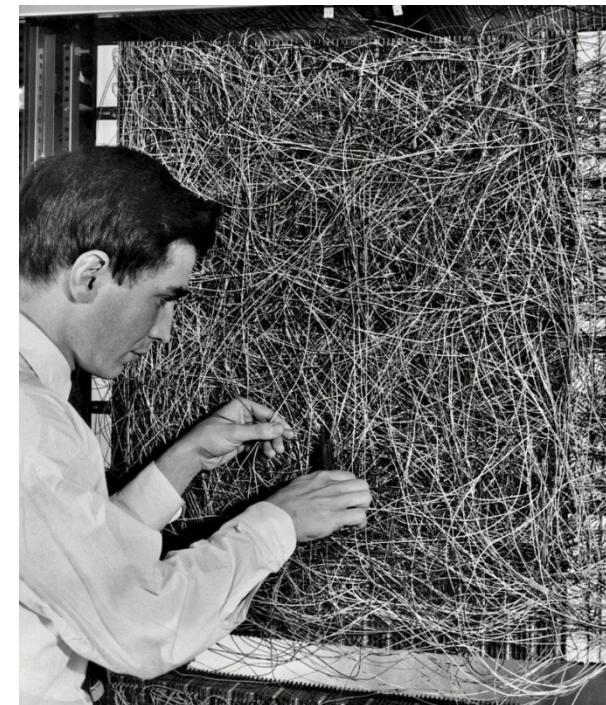
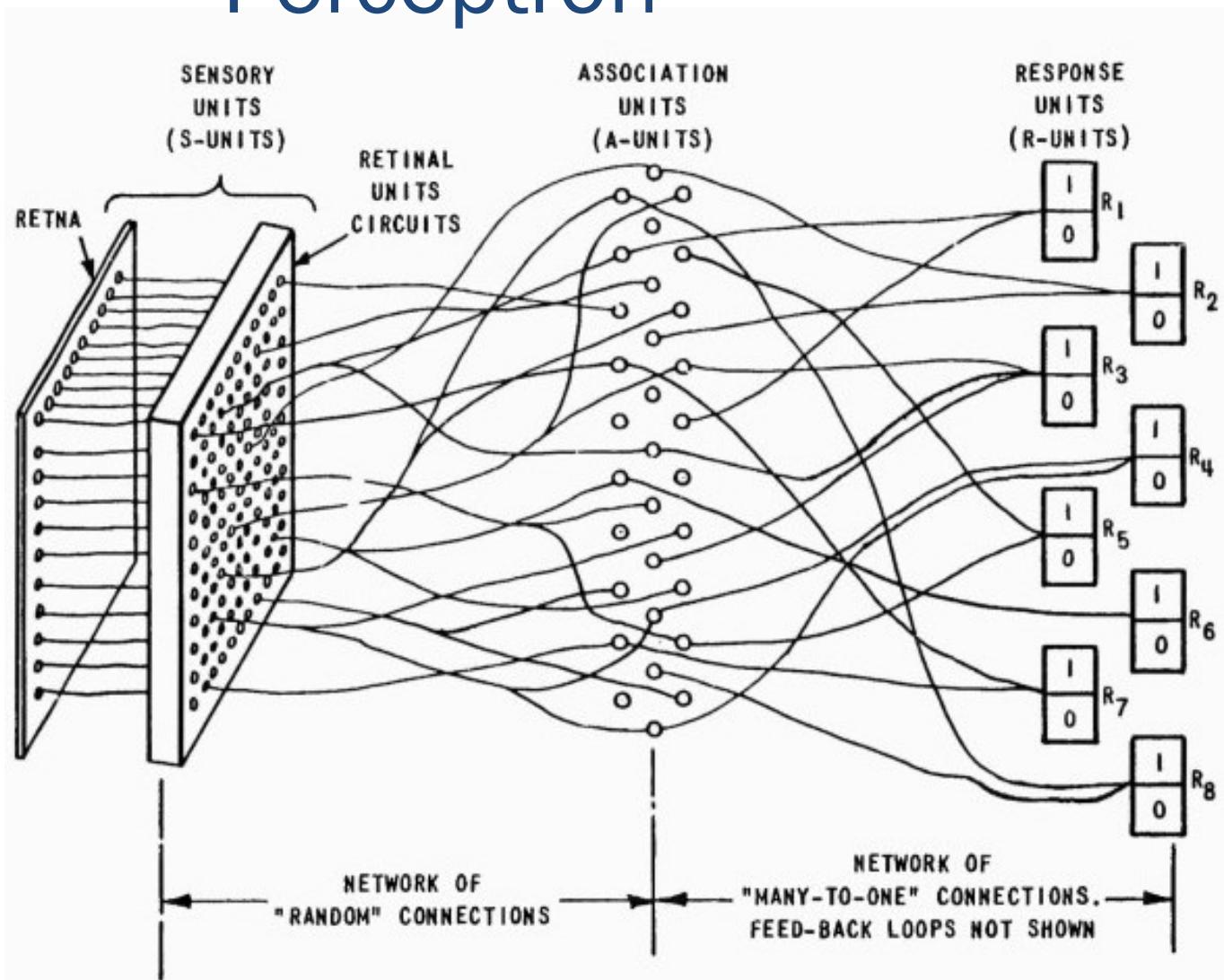
Figure 1 ORGANIZATION OF THE MARK I PERCEPTRON

Perceptron

- 1958 Frank Rosenblatt development and hardware construction of the “**Mark I Perceptron**”
 - the first computer that could learn new skills by trial and error, using a type of neural network that simulates human thought processes.
 - at the Cornell Aeronautical Laboratory funded by the United States Office of Naval Research.

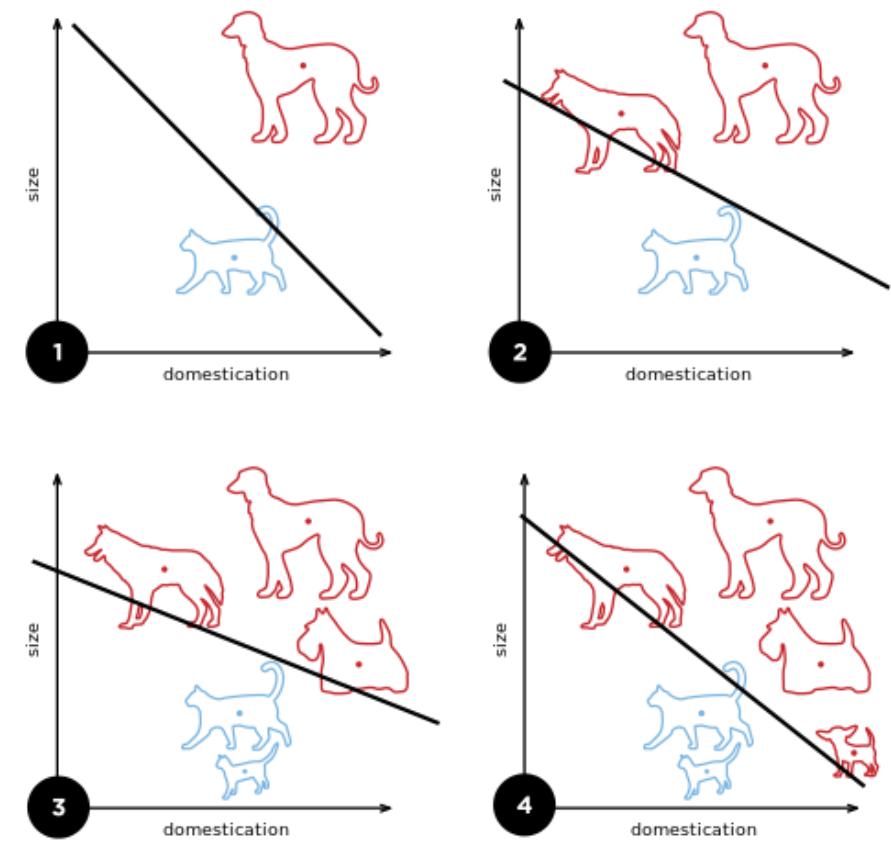
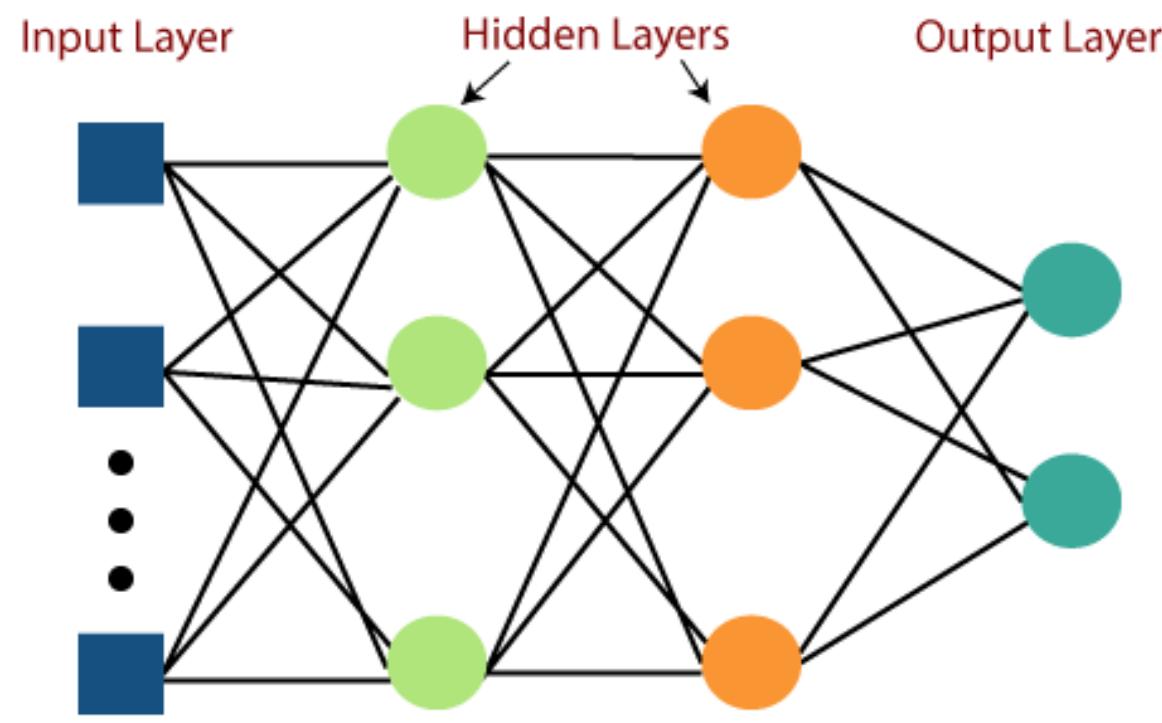


Perceptron



Multi-layer Perceptron

- 1960s Limitation of Perceptron: Failed to classify basic problems
- Ivakhnenko et. al. introduced Multilayer Perceptrons

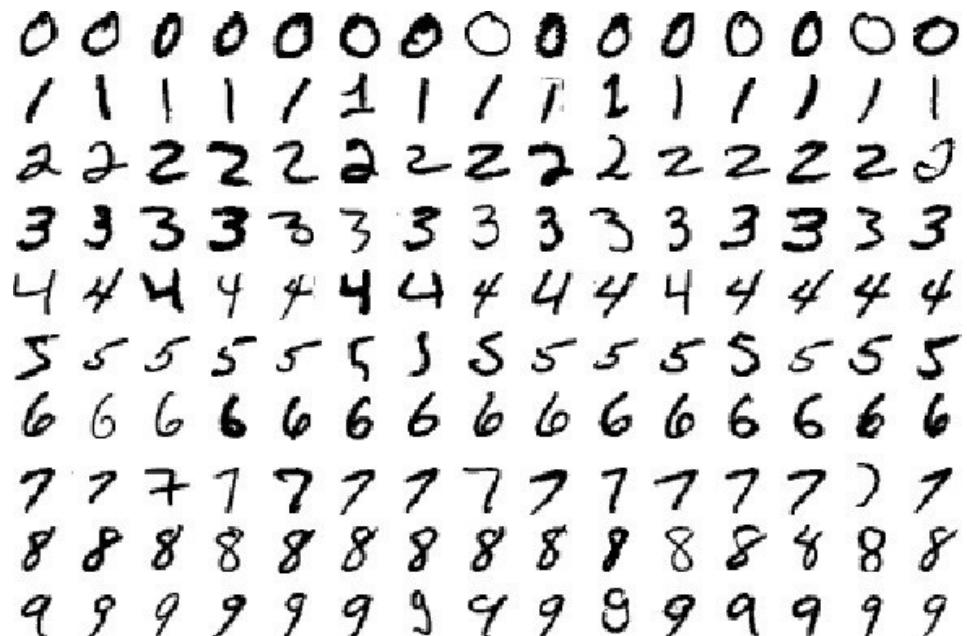


Artificial Neural Network and Backpropagation

- 1982: Werbos introduced Backpropagation first used it in the context of Artificial Neural Networks
- 1986: Rumelhart et. al. popularized the Backpropagation
- Use of Cauchy discovered theories of Gradient Descent on Convex function for Optimization
- 1989: Universal Approximation Theorem “A multilayered network of neurons with a single hidden layer can be used to approximate any continuous function to any desired precision”

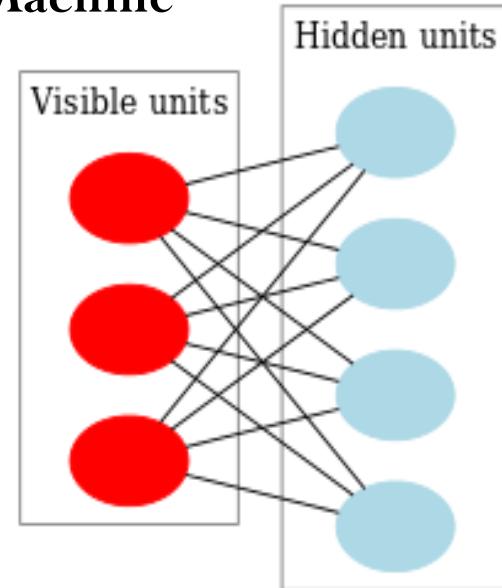
Deep Learning

- Unsupervised Pre-Training
 - 1991-1993: J. Schmidhuber “Very Deep Learner”
 - 2006: Hinton and Salakhutdinov
- 2009: Handwriting Recognition
 - MNIST dataset
- 2010: Speech Recognition
 - Dahl et. al. achieved error reduction of 16.0% to 23.2% over previous works
- Traffic Sign Recognition Competition
 - D. C. Ciresan et. al. got only 0.56% error rate

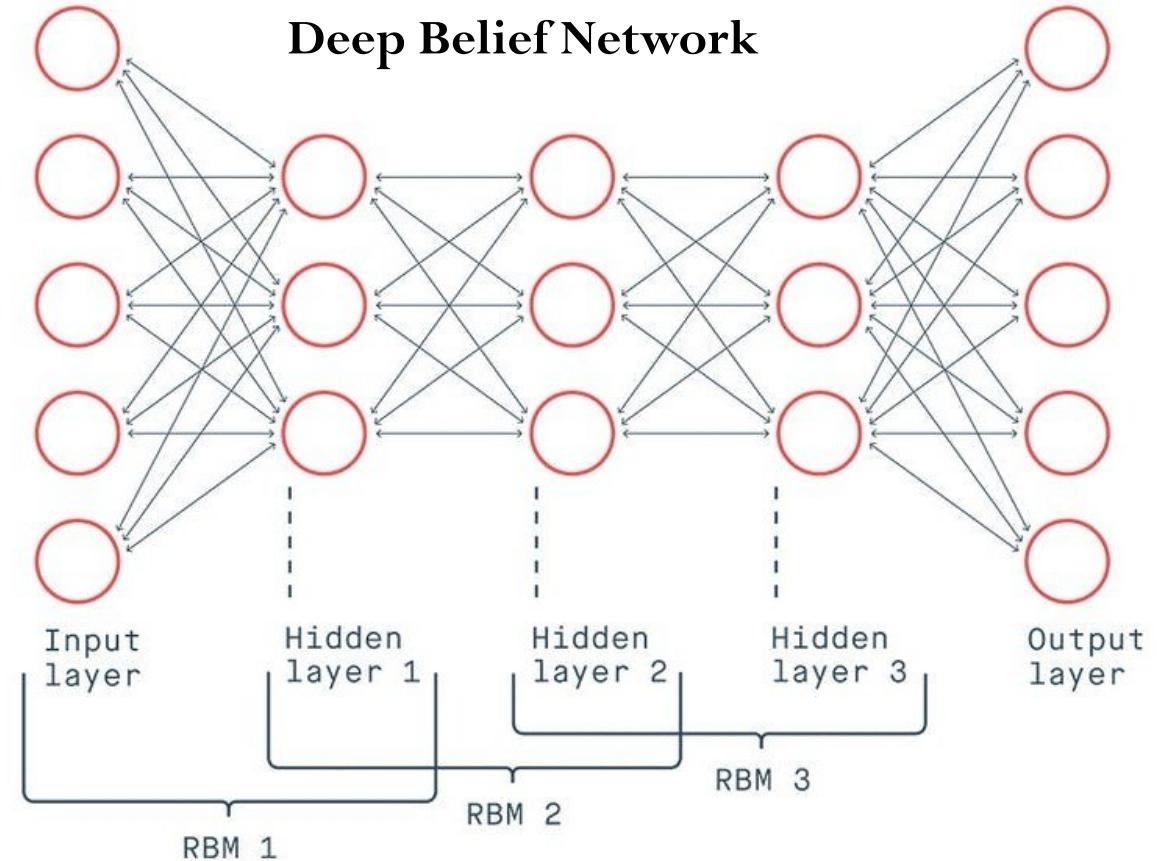


Restricted Boltzmann Machine and Deep Belief Network

**Restricted Boltzmann
Machine**



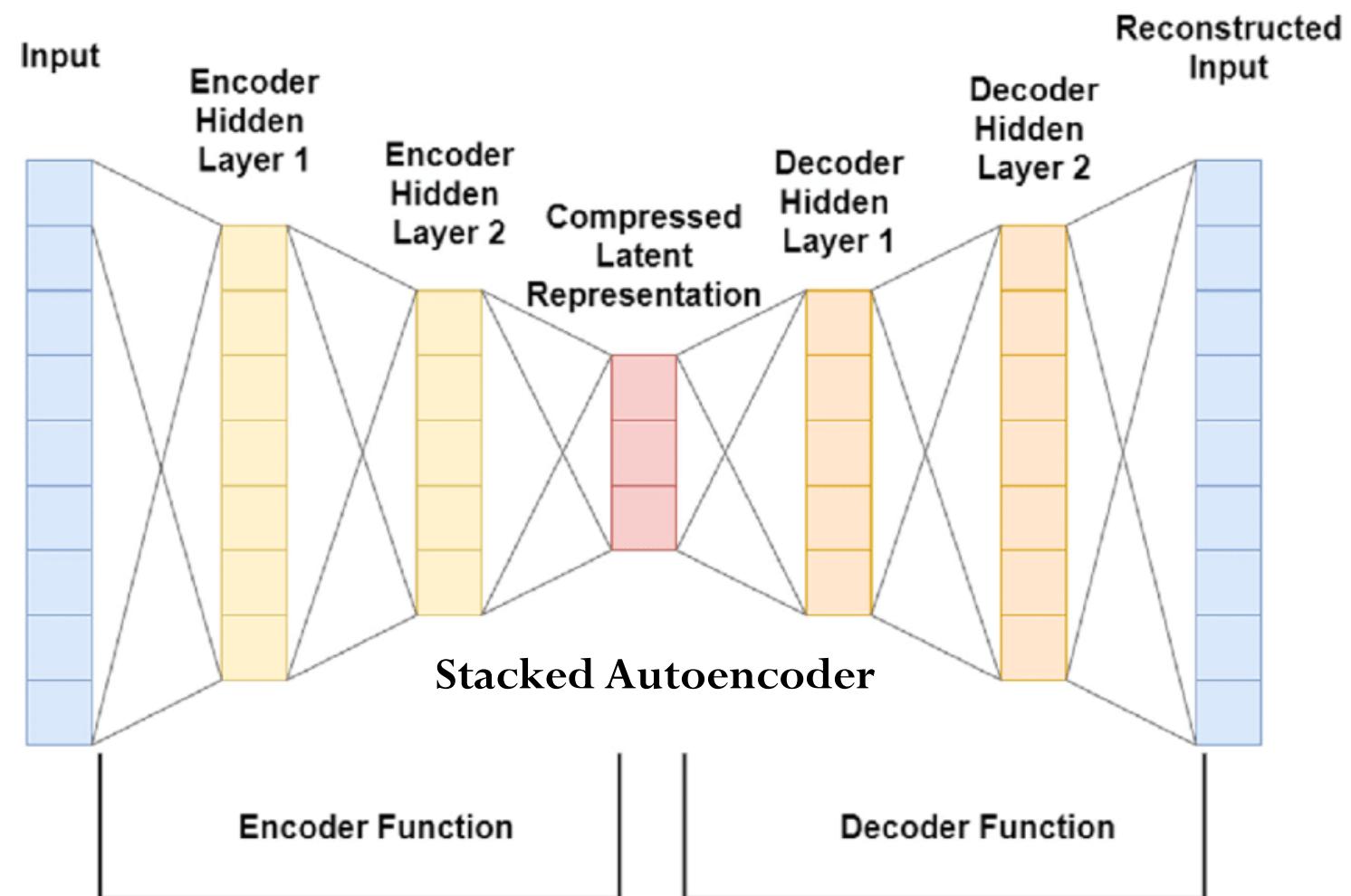
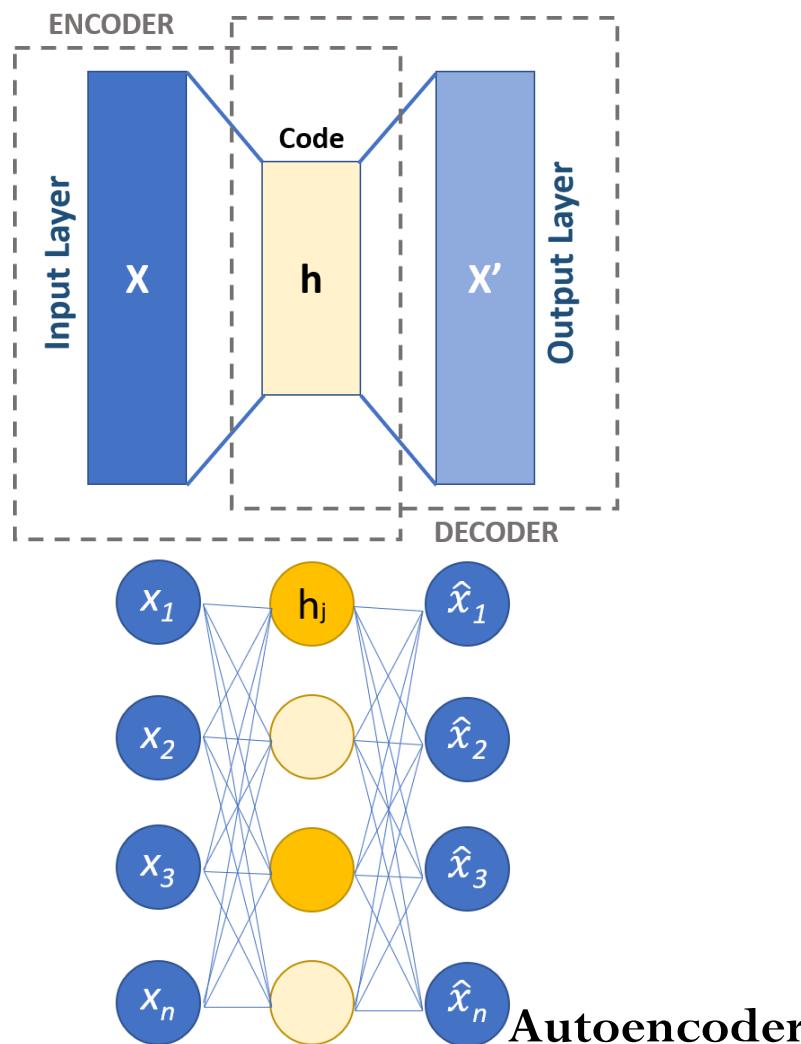
Deep Belief Network



Ruslan Salakhutdinov, Andriy Mnih, and Geoffrey Hinton. "Restricted Boltzmann machines for collaborative filtering." Proceedings of the 24th International Conference on Machine learning. 2007.

Geoffrey E. Hinton, Simon Osindero, and Yee-Whye Teh. "A fast learning algorithm for deep belief nets." Neural computation 18.7 (2006): 1527-1554.

Autoencoder and Stacked Autoencoder

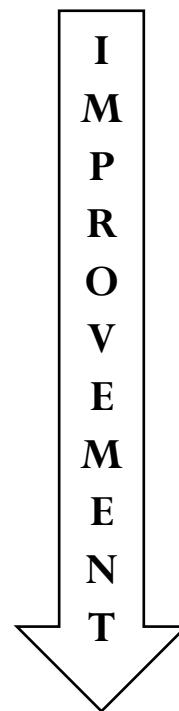


Pascal Vincent, et al. "Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion." Journal of Machine Learning Research 11.12 (2010).

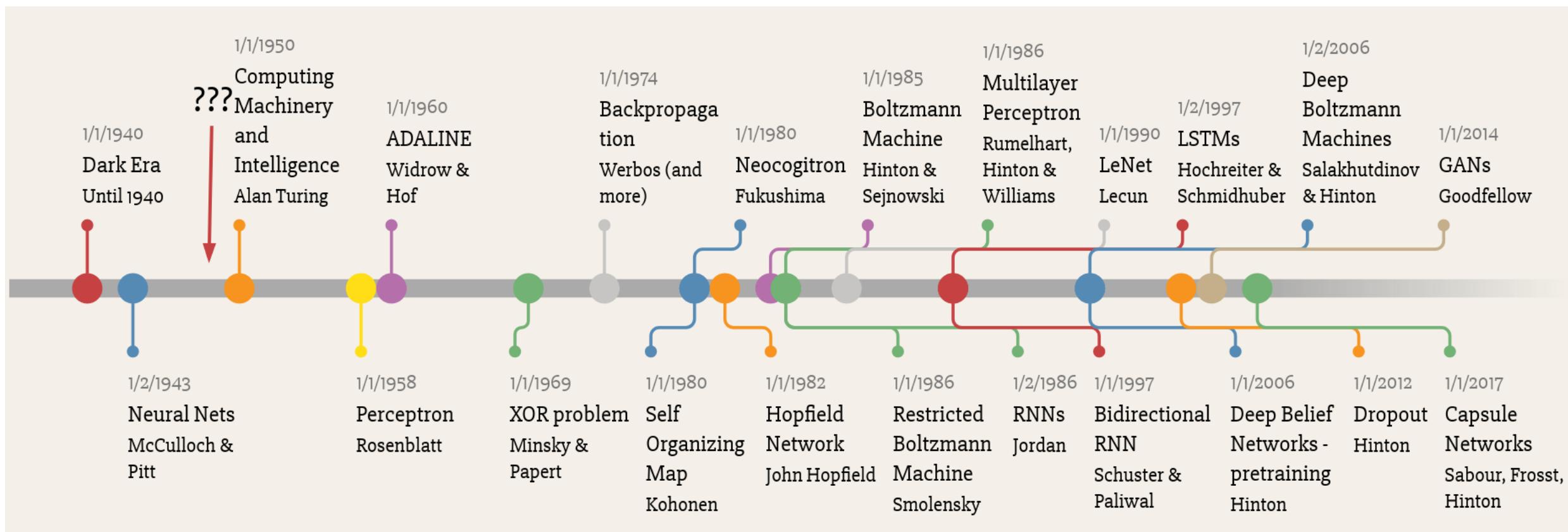
Visual Recognition Challenges

- ImageNet challenge successful 2012-2016

Network	Error	Layers
• AlexNet	16.0%	8
• ZFNet	11.2%	8
• VGGNet	7.3%	19
• GoogLeNet	6.7%	22
• MS ResNet	3.6%	152!!



Artificial Neural Network (ANN) Timeline



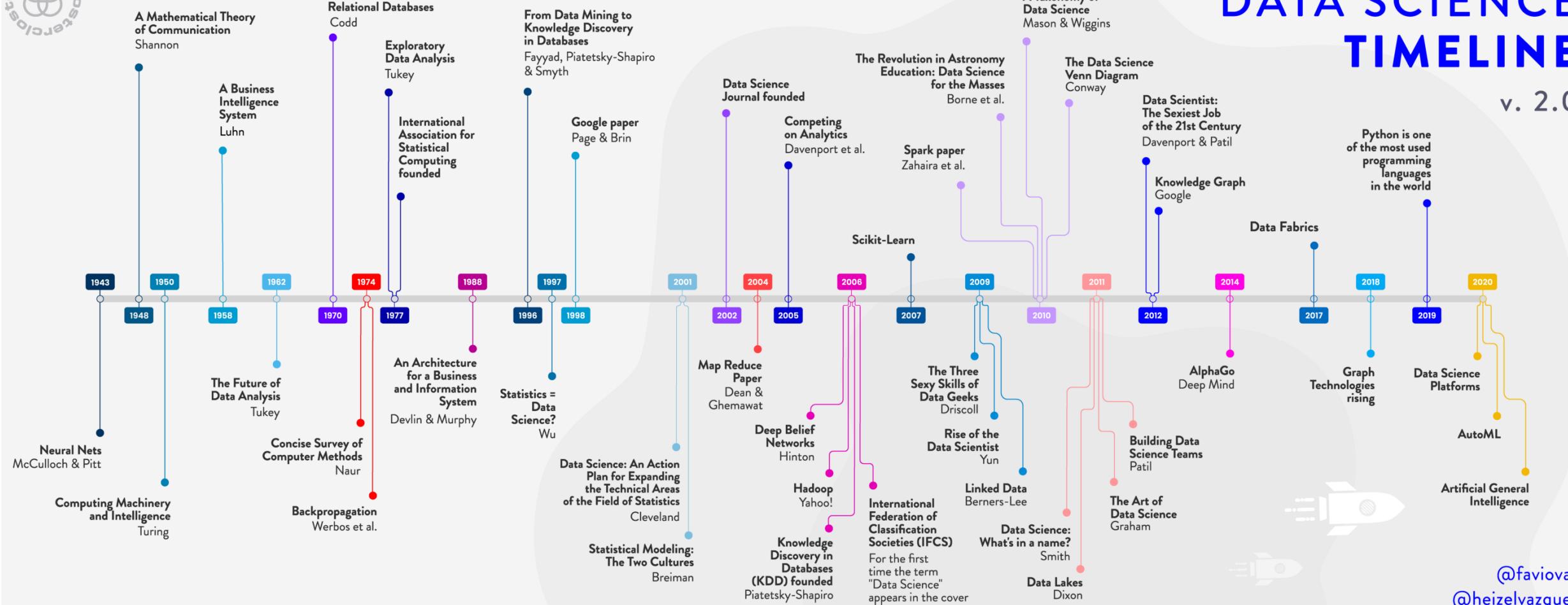
Deep Learning for the Masses (... and The Semantic Layer), Deep Learning timeline by Favio Vazquez
<https://www.kdnuggets.com/2018/11/deep-learning-masses-semantic-layer.html>

History of Data Science

Data Science

DATA SCIENCE TIMELINE

v. 2.0



The Roots of Data Science by [Favio Vázquez](#)

<https://towardsdatascience.com/the-roots-of-data-science-77c71115229>

@faviovaz
@heizelvazquez



Data Science

A Mathematical Theory
of Communication
Shannon

A Business
Intelligence
System
Luhn

Relational Databases

Codd

Exploratory
Data Analysis
Tukey

International
Association for
Statistical
Computing
founded

From Data Mining to
Knowledge Discovery
in Databases

Fayyad, Piatetsky-Shapiro
& Smyth

8

Data Science
Journal founded

Com
on AI
Dave

1943

1950

1948

1958

1962

1970

1974

1988

1996

1997

1998

2001

2002

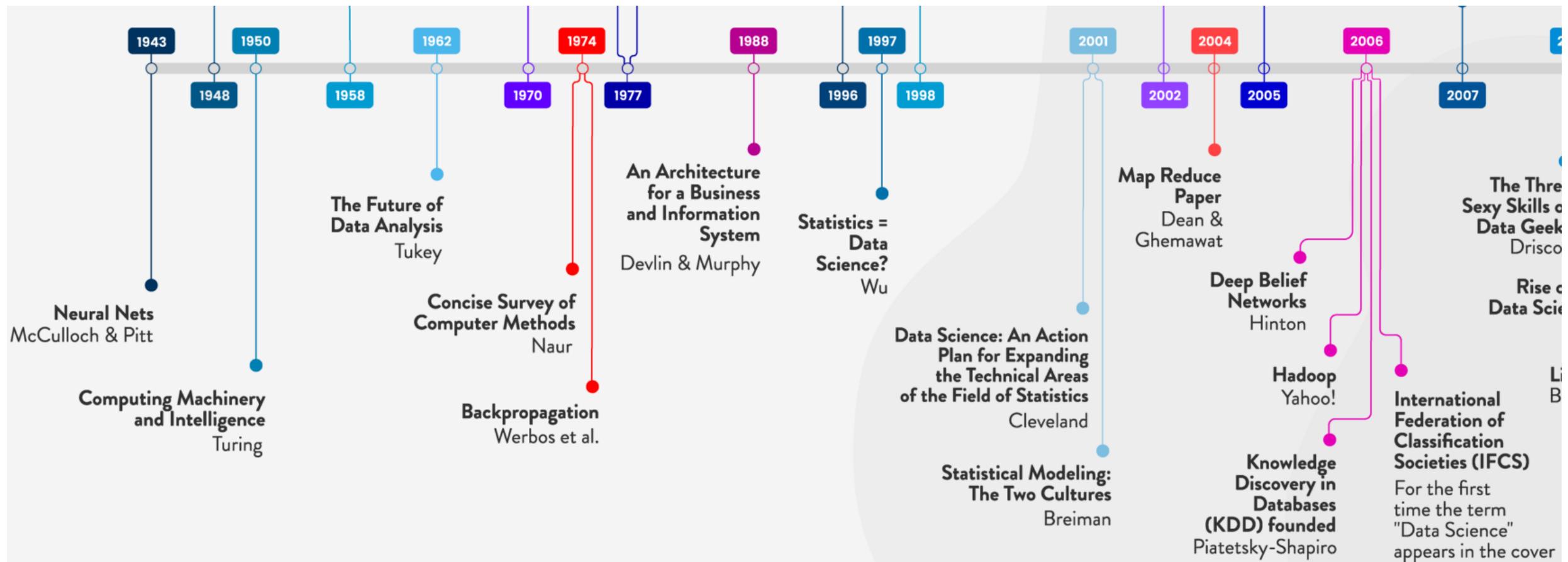
2004

2005

The Roots of Data Science by [Favio Vázquez](#)

<https://towardsdatascience.com/the-roots-of-data-science-77c71115229>

Data Science



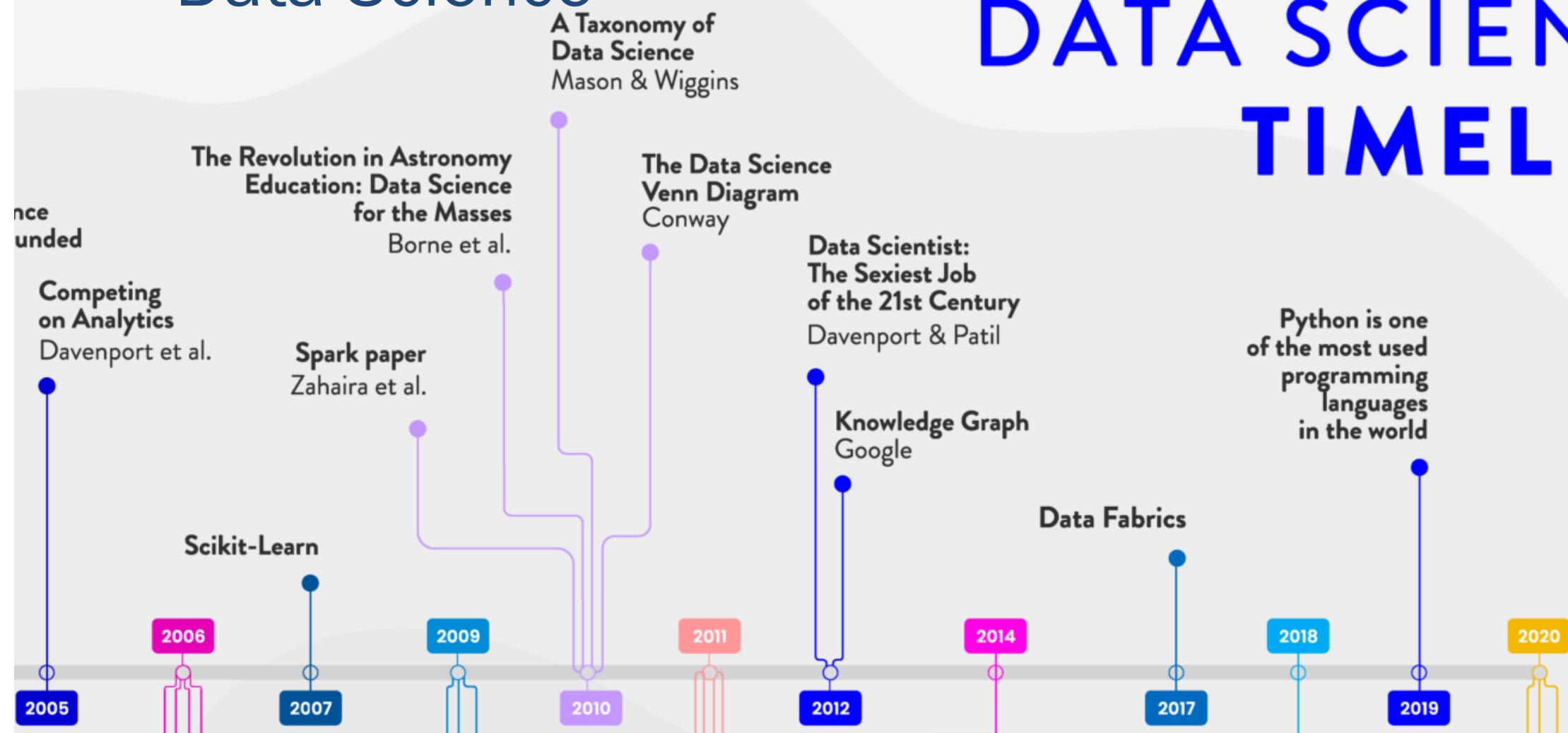
The Roots of Data Science by [Favio Vázquez](#)

<https://towardsdatascience.com/the-roots-of-data-science-77c71115229>

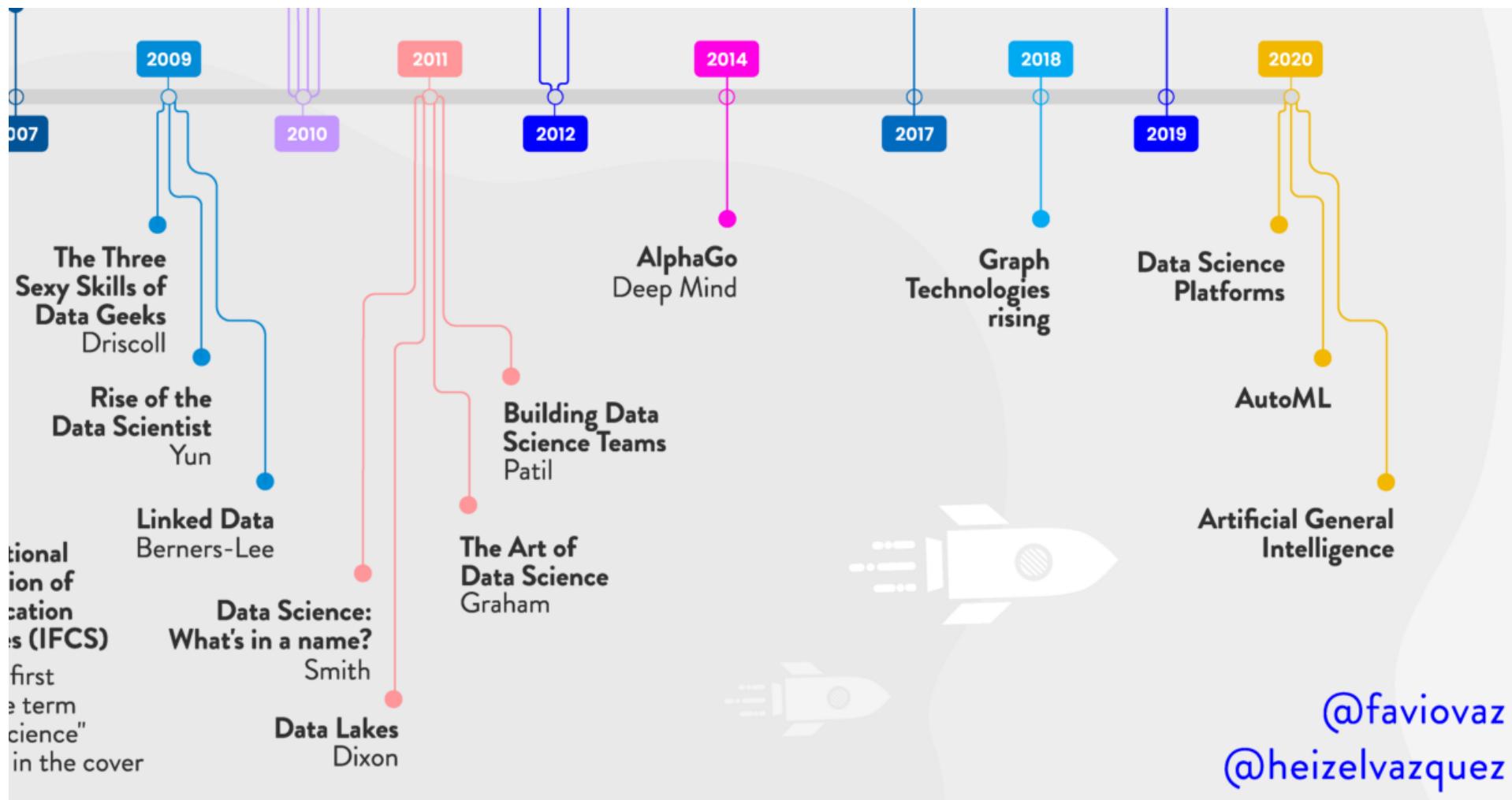
Data Science

DATA SCIENCE TIMELINE

v. 2.0



Data Science



The Roots of Data Science by [Favio Vázquez](#)

<https://towardsdatascience.com/the-roots-of-data-science-77c71115229>

@faviovaz

@heizelvazquez

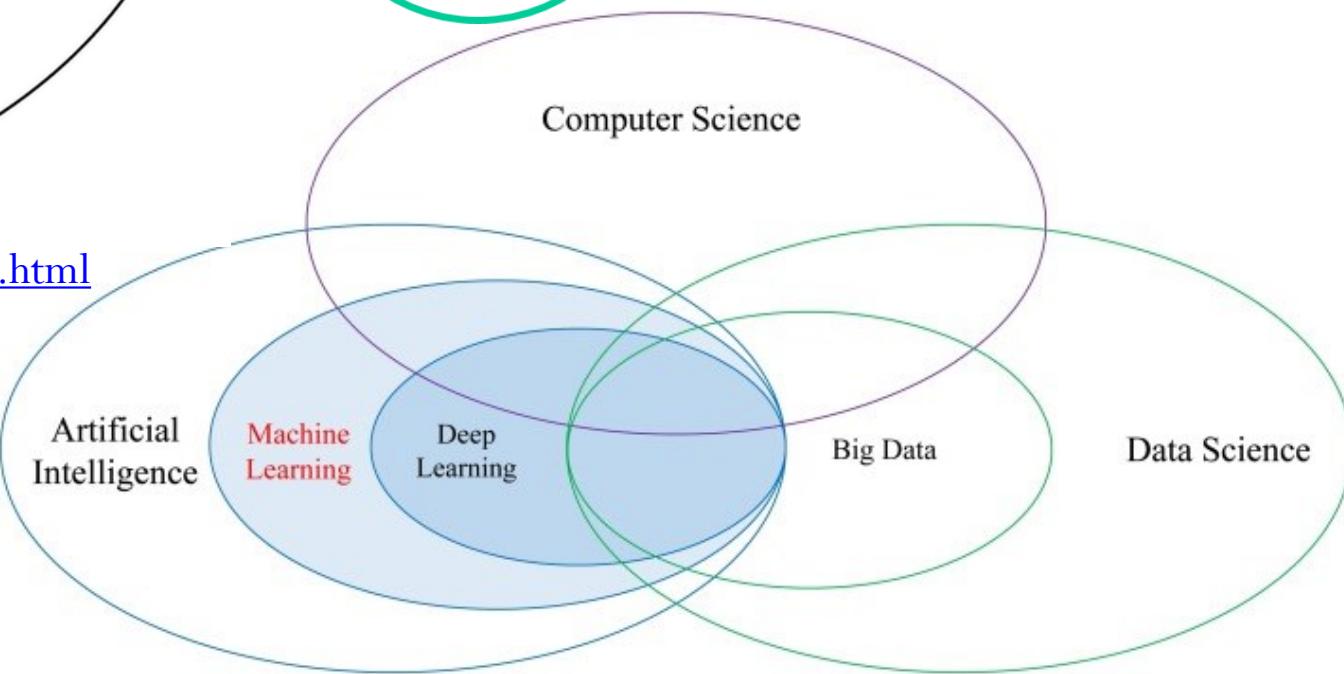
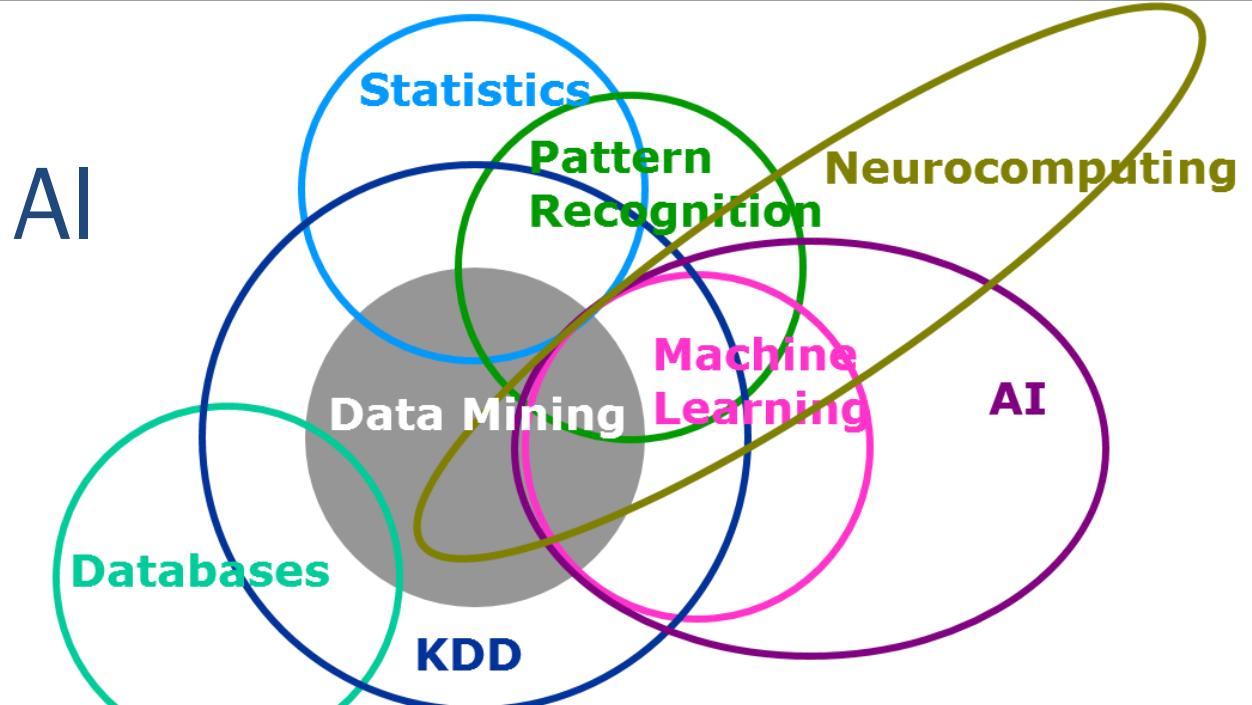
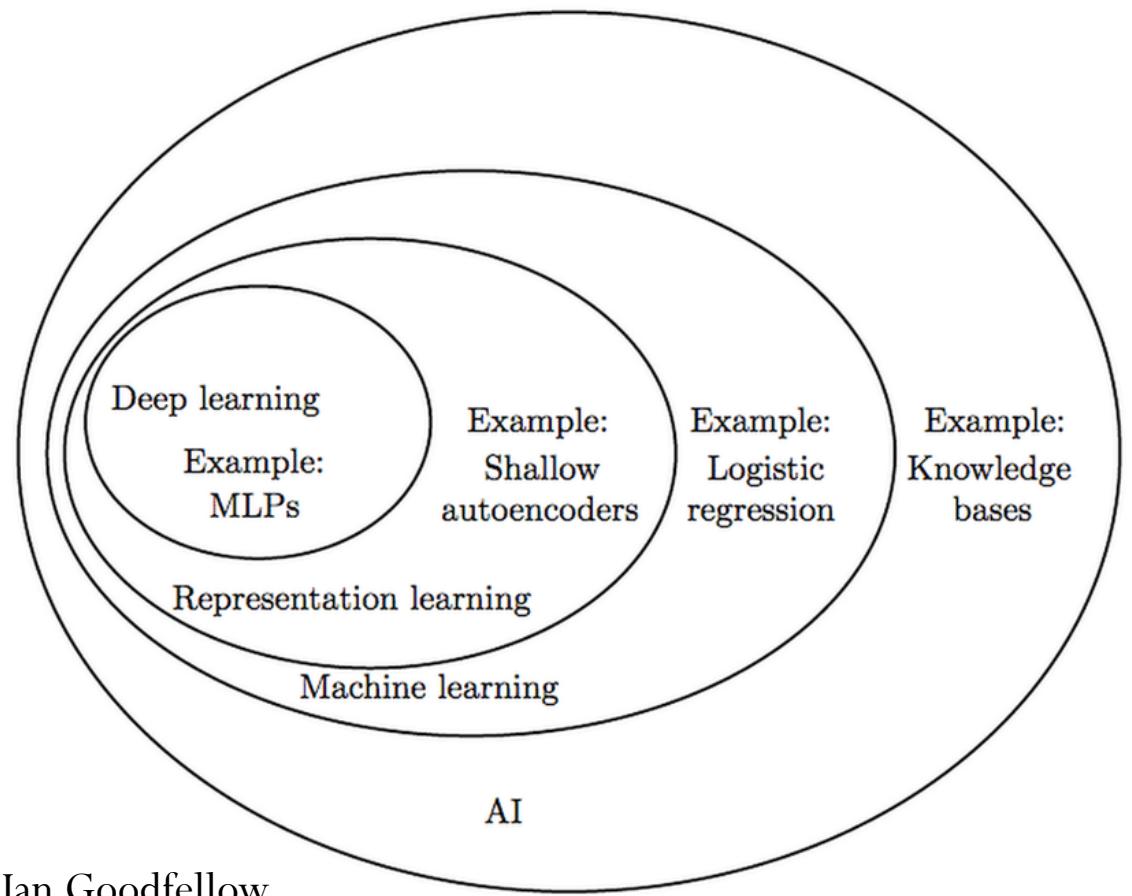
Summary

Artificial Intelligence

- Founded in 1956 (Dartmouth Research Project on Artificial Intelligence).
- Ancient History shows that AI was an active field of studies since ancient times.
- Simulating the brain, modeling human problem solving, formal logic, large databases of knowledge and imitating human behavior
- Search, Optimization, Formal logic,
- Artificial Neural Networks, Deep learning,
- Classifiers, Reasoning, Knowledge Representation,
- Planning, Learning, Robotics,
- Natural Language Processing, Computational Social Scienc, etc.

Artificial Intelligence

- (1837) Analytical Engine, (1925) Houdina Radio Control,
- (1946-1960) ENIAC heralds the dawn of Computing,
- (1946-1960) Turing Test, (1959) Machine Learning,
- (1964) Eliza Chatbot, (1966) Shakey general purpose mobile robot,
- (1960-74) Expert Systems,
- AI Winter (1974-80), AI Spring (1980-87), AI Winter (1987-93),
- (1993) Apriori algorithm: Association Rule Mining,
- (1997) Deep Blue ends Human Supremacy in Chess,
- (1998) Deep Space, (2005) Driverless Cars,
- (2011) IBM Watson beaten Human in Jeopardy,
- (2016) Alpha Go



Ian Goodfellow

<https://www.deeplearningbook.org/contents/intro.html>

References

- Stuart Russel, and Peter Norvig. "Artificial intelligence: A modern approach. Third edit." Upper Saddle River, New Jersey 7458 (2015). <http://aima.cs.berkeley.edu/>
- Wikipedia contents <https://www.wikipedia.org/>
- Images are from several sources e.g. movies, TV serials, internet, miscellaneous links, slides, blogs, etc.
- Vedas Tantra , The Ancient Indian Secret Of Artificial Intelligence, <https://www.youtube.com/watch?v=RAPK-RstuWI>
- Raja Bhoja – The Polymath Monarch, D.K.Hari & D.K.Hema Hari <https://bharathgyanblog.files.wordpress.com/2017/12/raja-bhoja.pdf>
- Subbarao Kambhampati, Arizona State University, <https://rakaposhi.eas.asu.edu/cse471/>
- Artificial Intelligence, Prof. Mausam <https://www.cse.iitd.ac.in/~mausam/courses/col333/autumn2019/>

תודה רבה

Hebrew

Danke

German

Merci

French

Grazie

Italian

Gracias

Spanish

Obrigado

Portuguese

Ευχαριστώ

Greek

Спасибо

Russian

ধন্যবাদ

Bangla

ಧನ್ಯವಾದಗಳು

Kannada

ధన్యవాదాలు

Telugu

ਧੰਨਵਾਦ

Punjabi

धन्यवादः

Sanskrit

Thank You

English

நன்றி

Tamil

മന്ത്രി

Malayalam

આમાર

Gujarati

شُكْرًا

Arabic

多謝

Traditional Chinese

多谢

Simplified Chinese

ありがとうございました

Japanese

ຂອບຄຸມ

Thai

감사합니다

Korean

<https://sites.google.com/site/animeshchaturvedi07>