



# THE HISTORY OF ARTIFICIAL INTELLIGENCE

**Professor Tim Miller**

School of Computing and Information Systems  
Co-Director, Centre for AI & Digital Ethics  
The University of Melbourne  
[tmiller@unimelb.edu.au](mailto:tmiller@unimelb.edu.au)  
[@tmiller\\_unimelb](https://twitter.com/tmiller_unimelb)



## **Warning**

This material has been reproduced and communicated to you by or on behalf of the University of Melbourne pursuant to Part VB of the *Copyright Act 1968 (the Act)*.

The material in this communication may be subject to copyright under the Act.

Any further copying or communication of this material by you may be the subject of copyright protection under the Act.

**Do not remove this notice**

# LEARNING OUTCOMES

1956 => Now → future

Discuss key eras of the development of artificial intelligence

Apply lessons from the past to discussions in the present.

Make new links between the way we currently build AI, and the effects it has on the people that use it



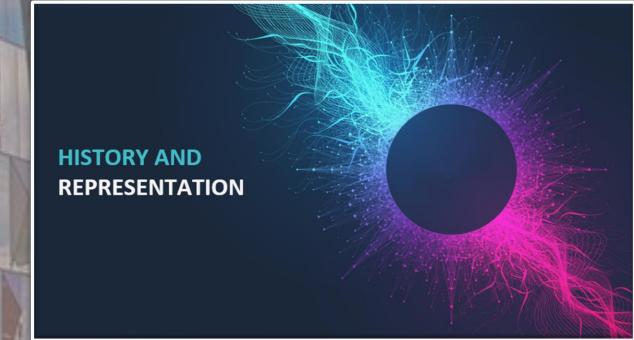
# OVERVIEW



A BRIEF HISTORY OF  
COMPUTING

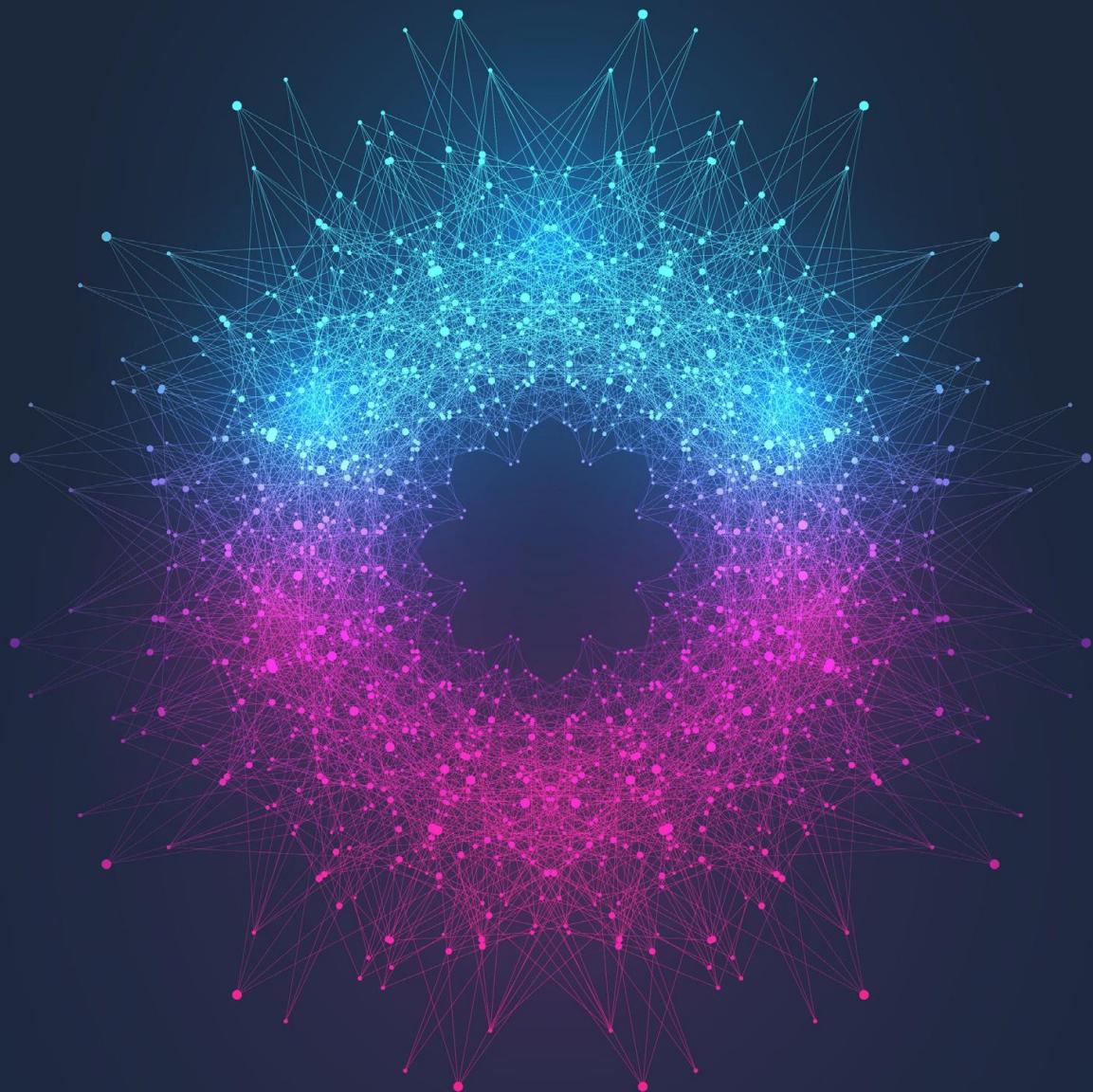


THE ERAS OF  
ARTIFICIAL INTELLIGENCE



HISTORY AND  
REPRESENTATION

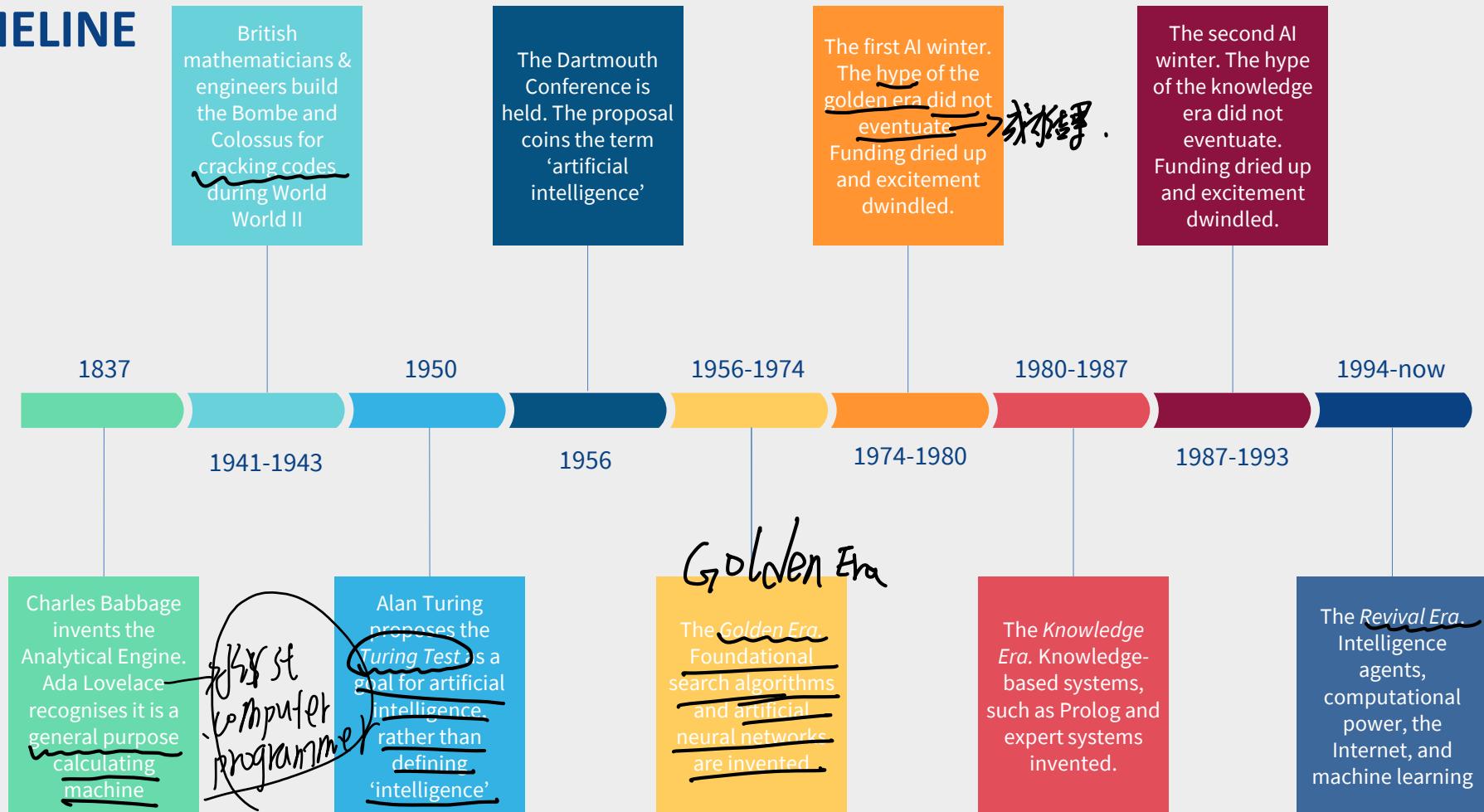
# A BRIEF HISTORY OF COMPUTING



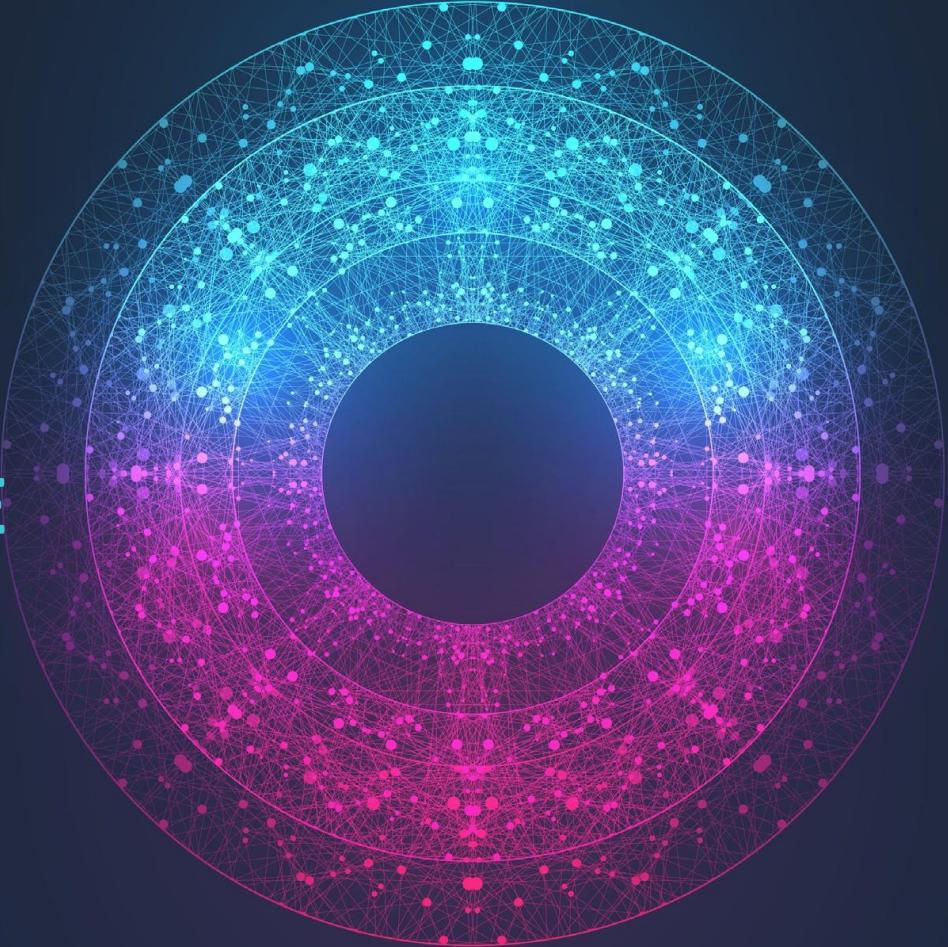
# IN WHAT YEAR WAS THE FIRST PROGRAMMABLE COMPUTER BUILT?

1. 87 BCE Antikythera machine (是第一台可编程的)
2. 1642 只做计算 (X)
3. 1805 (X)
4. 1837 Analytical engine (X)
5. 1943 Colossus (2nd war) (decode German message)
6. 1951 (1st computer to play music) PROGRAMMABLE

# TIMELINE



# THE ERAS OF ARTIFICIAL INTELLIGENCE



# THE TURING TEST (1950)

“The new form of the problem can be described in terms of a game which we call the *‘imitation game’*”

Turing Test

Alan Turing. Computing Machinery and Intelligence, *Mind*, LIX(236):433–460, 1950.

<https://doi.org/10.1093/mind/LIX.236.433>

Does machine ~~can~~ be intelligent?

人  
+  
Screen

→ 通过：机器 —> 与不；真人。电讯



# DARTMOUTH WORKSHOP (1956)

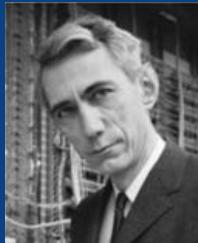
## THE FOUNDERS OF AI



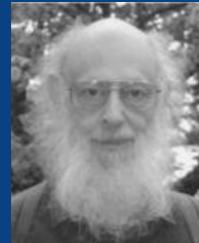
JOHN McCARTHY



MARVIN MINSKY



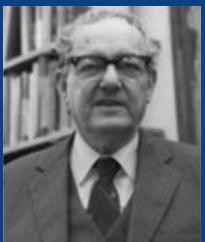
CLAUDE SHANNON



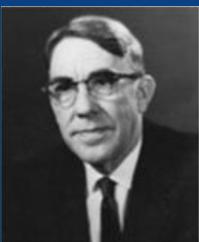
RAY SOLOMONOFF



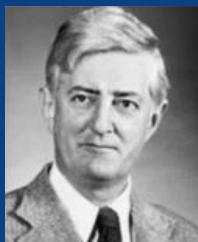
ALAN NEWELL



HERBERT SIMON



ARTHUR SAMUEL



OLIVER SELFRIDGE



NATHANIEL ROTCHESTER

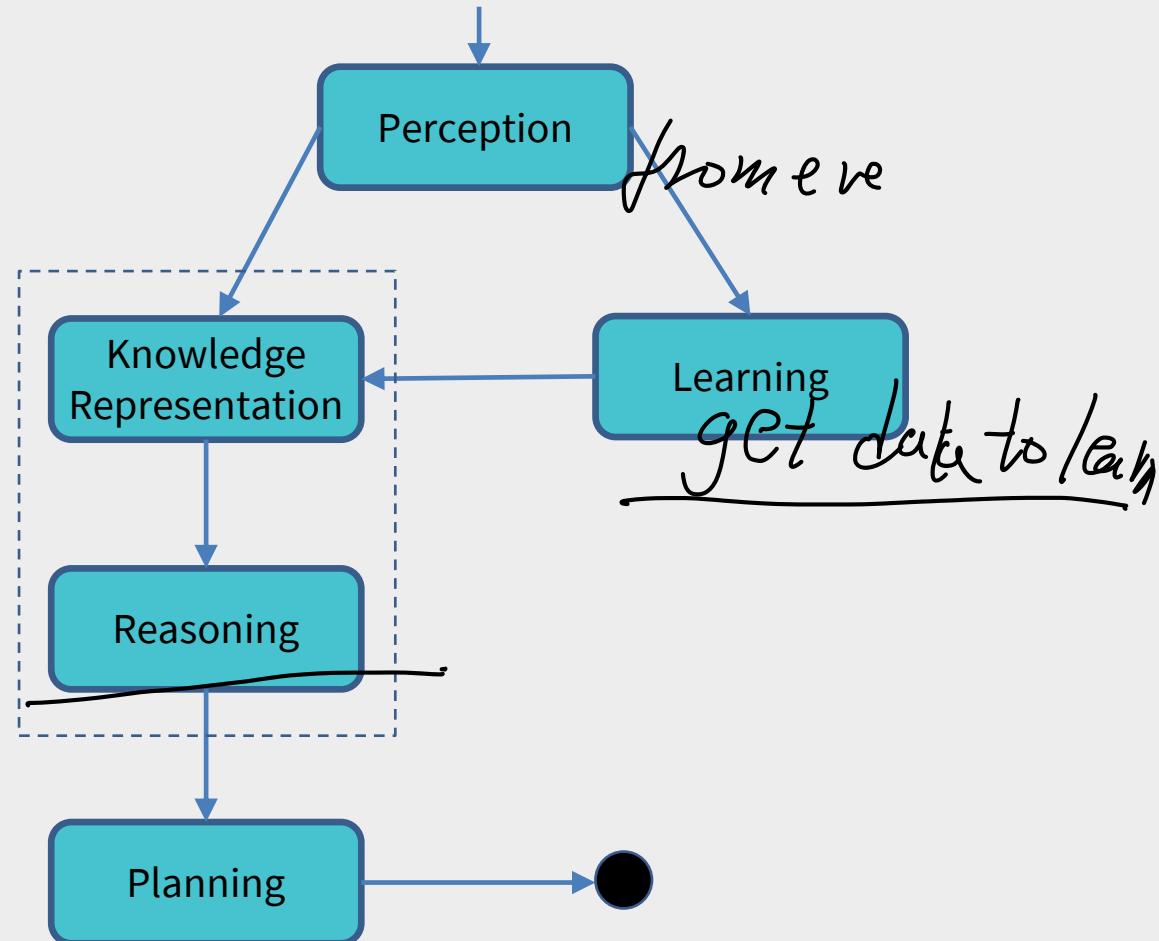


TRENCHARD MOORE

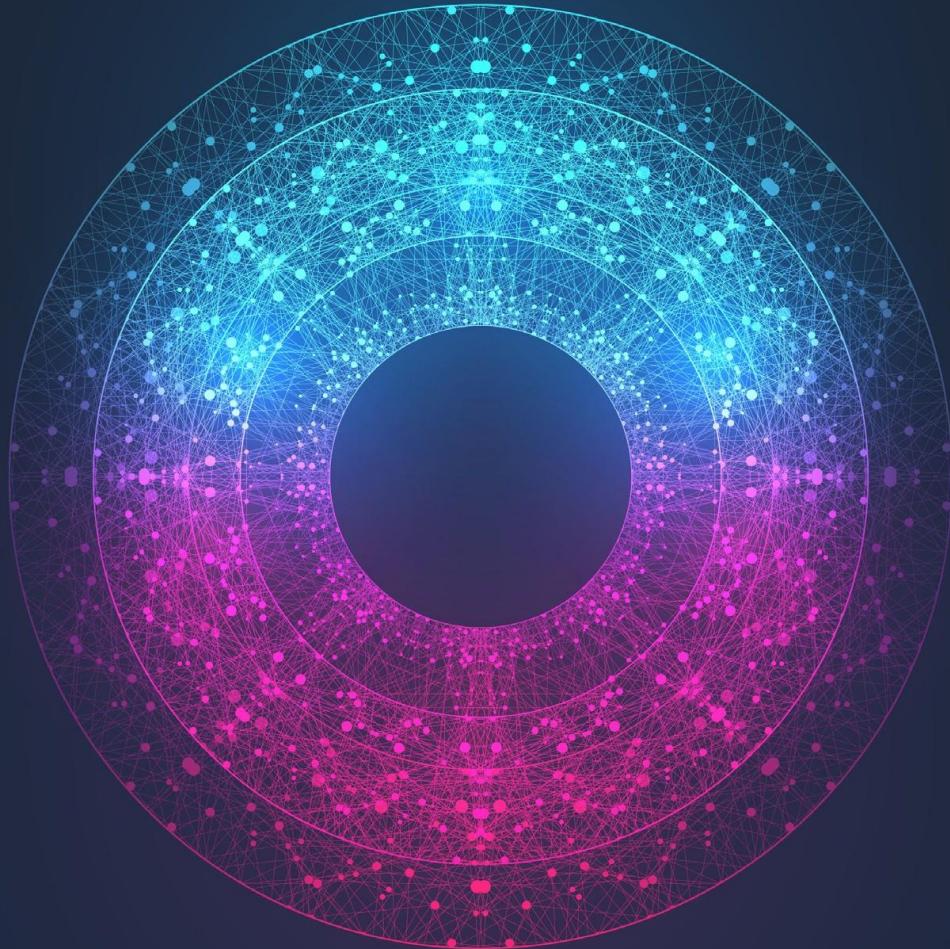
“We propose that a 2-month, 10-man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so 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.” – Dartmouth Summer School on AI Proposal

# DARTMOUTH OUTCOMES

The divide and conquer model of artificial intelligence

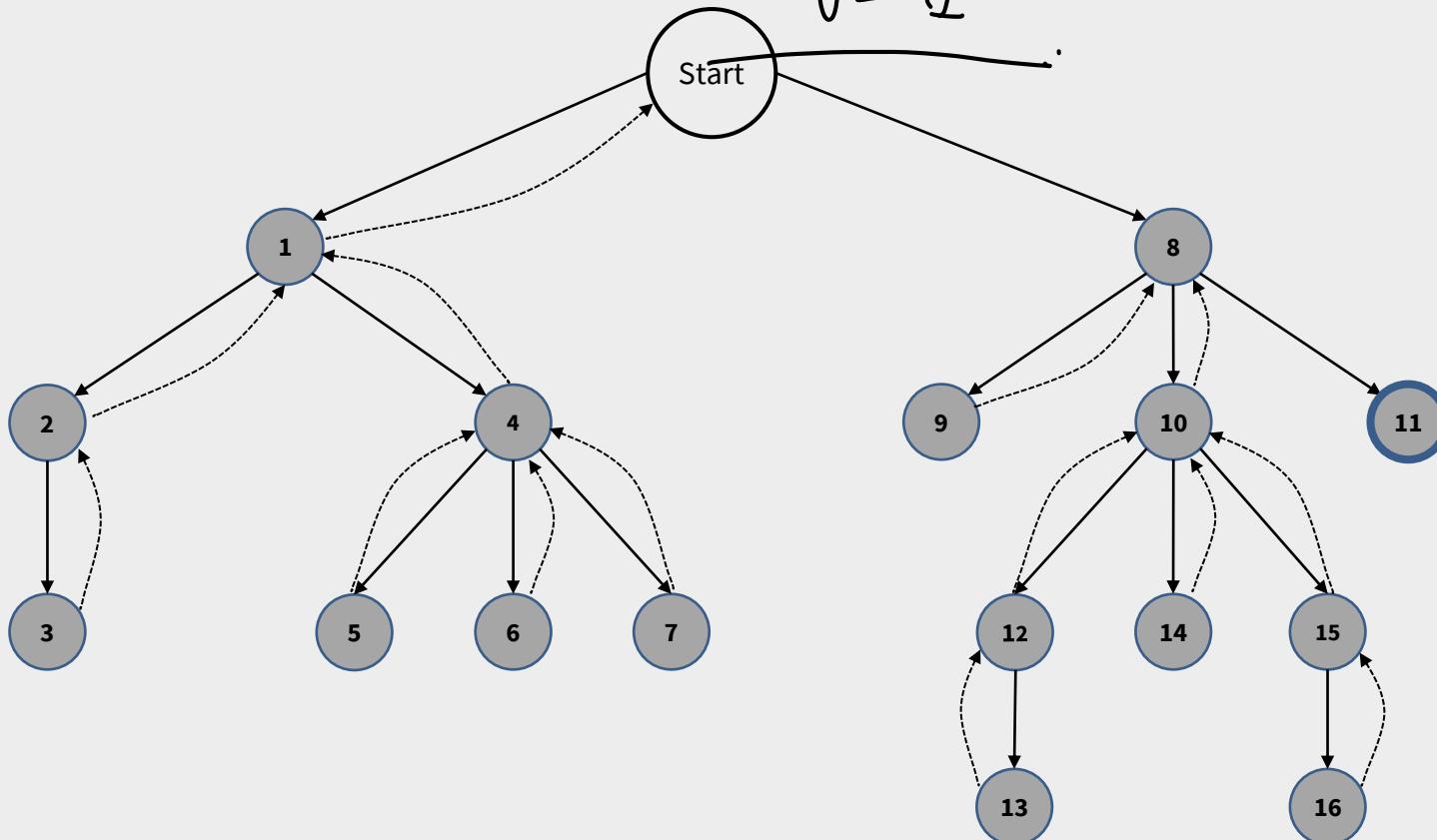


# THE GOLDEN AGE OF AI (1956-1974)



# GOLDEN AGE: REASONING AS SEARCH

推理



# GOLDEN AGE: REASONING AS SEARCH

Initial state: At(A), Level(low), BoxAt(C), BananasAt(B)

Goal state: Have(bananas)

Actions:

// move from X to Y

\_Move(X, Y)\_

Preconditions: At(X), Level(low)

Postconditions: not At(X), At(Y)

// climb up on the box

\_ClimbUp(Location)\_

Preconditions: At(Location), BoxAt(Location), Level(low)

Postconditions: Level(high), not Level(low)

// climb down from the box

\_ClimbDown(Location)\_

Preconditions: At(Location), BoxAt(Location), Level(high)

Postconditions: Level(low), not Level(high)

// move monkey and box from X to Y

\_MoveBox(X, Y)\_

Preconditions: At(X), BoxAt(X), Level(low)

Postconditions: BoxAt(Y), not BoxAt(X), At(Y), not At(X)

// take the bananas

\_TakeBananas(Location)\_

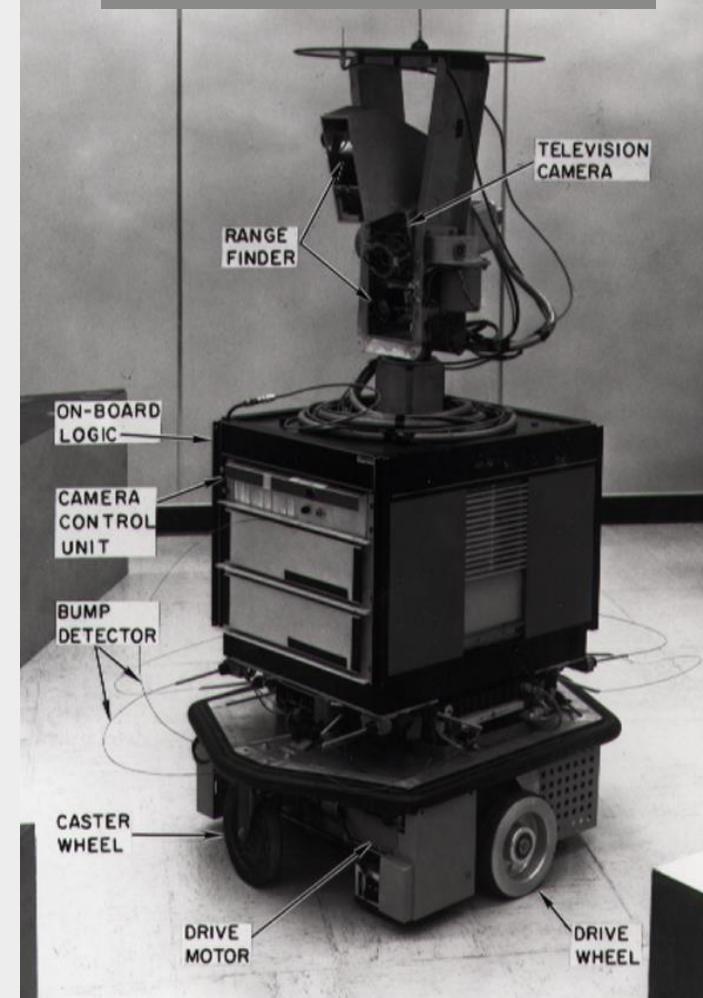
Preconditions: At(Location), BananasAt(Location), Level(high)

Postconditions: Have(bananas)

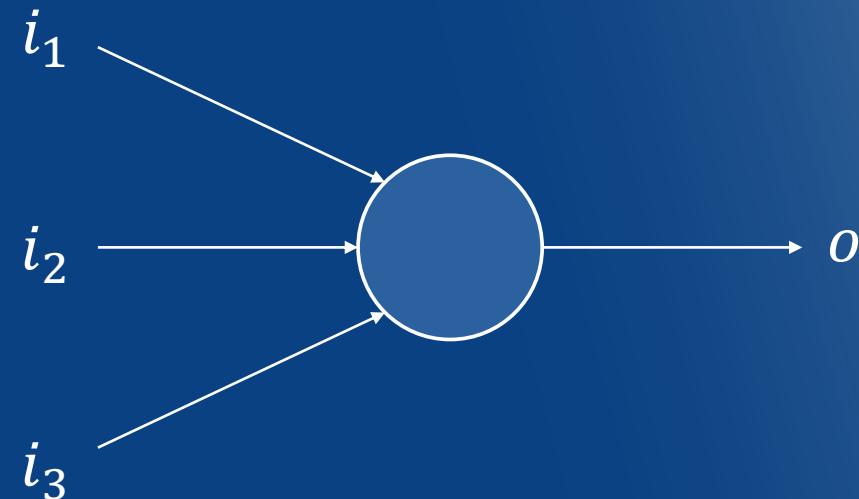
Stanford Research Institute Problem Solver (STRIPS) -- Fikes and Nilsson (1971)

ANTENNA FOR  
RADIO LINK

Shakey the robot and A\* -- Hart, Nilsson,  
and Raphael (1968)

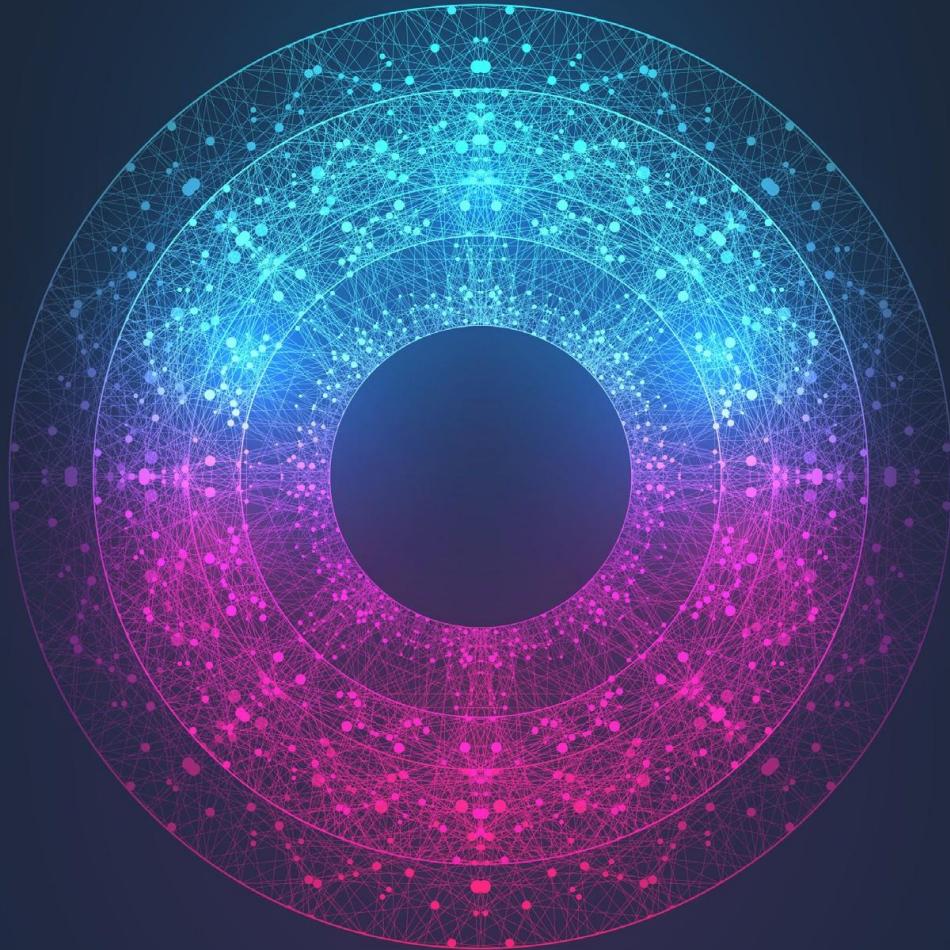


# GOLDEN AGE: PERCEPTRONS AND NEURAL NETWORKS



Single-layer perceptron – Rosenblatt (1958)

# THE FIRST AI WINTER (1974-1980)



# WHAT WENT WRONG?

(X) THE

Outcomes failed to live up to the hype

SCALABILITY

COMMONSENSE KNOWLEDGE

PERCEPTRON LIMITATIONS

MORAVEC'S PARADOX



## WHAT WAS THE RESULT?

Lack of progress meant:

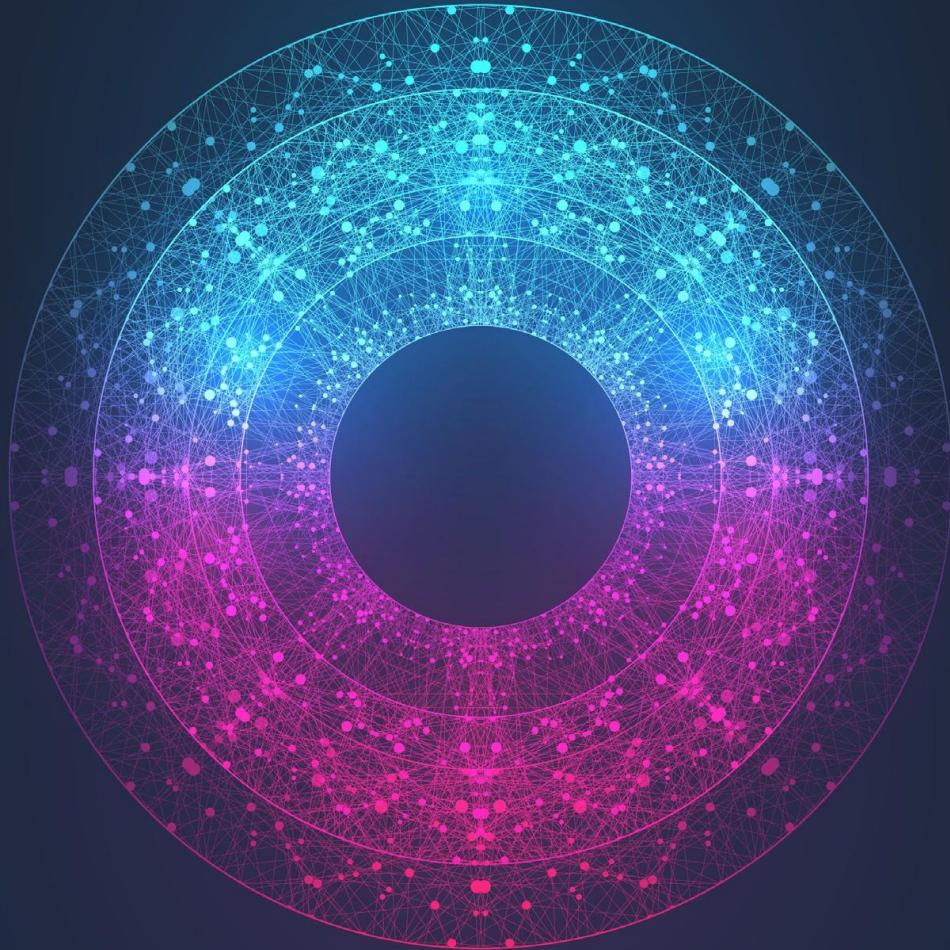
FUNDING DRIED UP

GLOBAL INTEREST IN AI DIED DOWN

~~CRITICISM FROM PHILOSOPHERS AND COGNITIVE SCIENTISTS~~

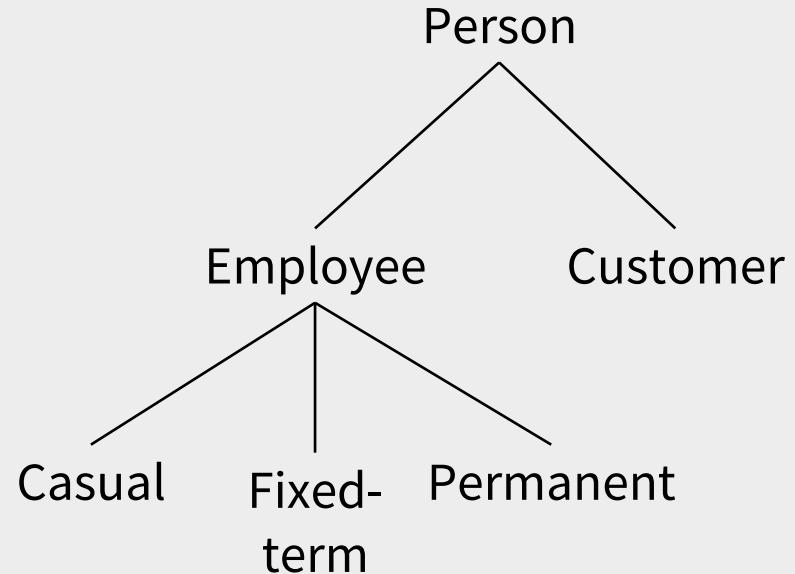


# THE KNOWLEDGE ERA (1980-1987)



# KNOWLEDGE ERA: KNOWLEDGE-BASED SYSTEMS

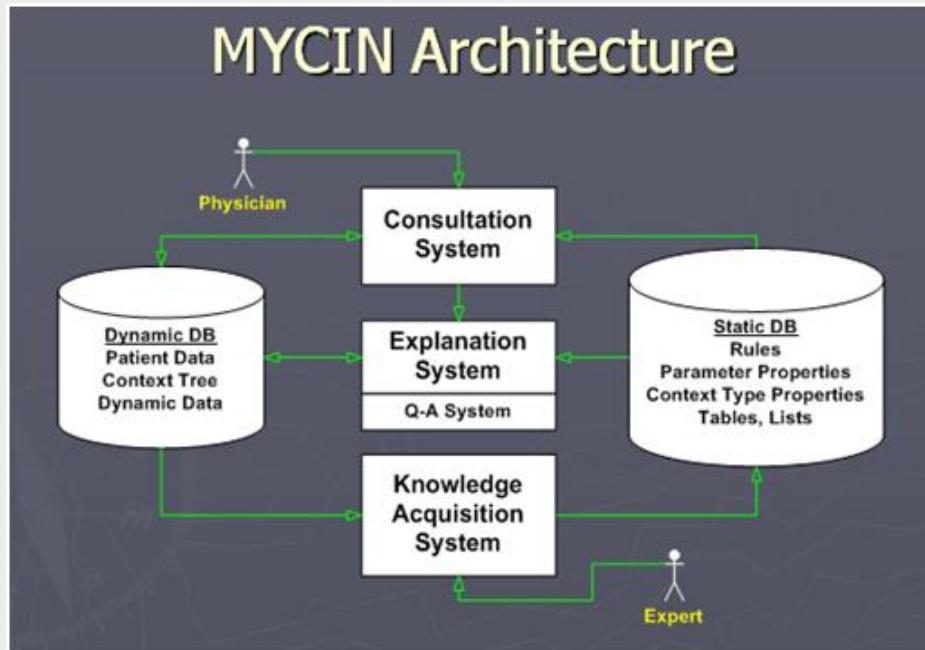
```
mother_child(trude, sally).  
father_child(tom, sally).  
father_child(tom, erica).  
father_child(mike, tom).  
  
sibling(X, Y) :- parent_child(Z, X),  
                 parent_child(Z, Y).  
  
parent_child(X, Y) :- father_child(X, Y).  
parent_child(X, Y) :- mother_child(X, Y).  
  
?- sibling(sally, erica).  
Yes
```



Prolog – Colmerauer and Roussel (1972)

Formal ontologies

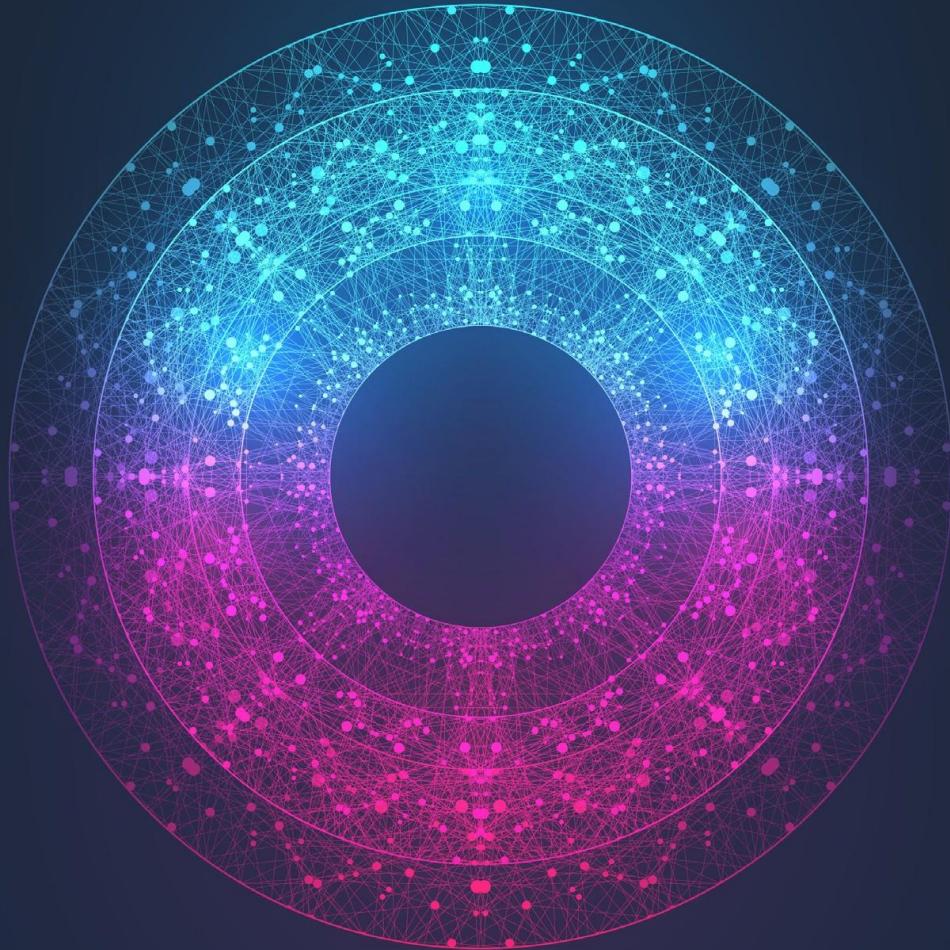
# KNOWLEDGE ERA: EXPERT SYSTEMS



MYCIN expert system for diagnosis of blood diseases –  
Shortcliffe, Buchanan, and Cohen (1970s)



# THE SECOND AI WINTER (1987-1993)



# WHAT WENT WRONG?

Outcomes failed to live up to the hype

SCALABILITY (可擴)

MAINTENANCE (

THE QUALIFICATION PROBLEM

MORAVEC'S PARADOX



# WHAT WAS THE RESULT?

Lack of progress meant:

FUNDING DRIED UP (DARPA DECLARED  
AI WAS 'NOT THE NEXT WAVE')

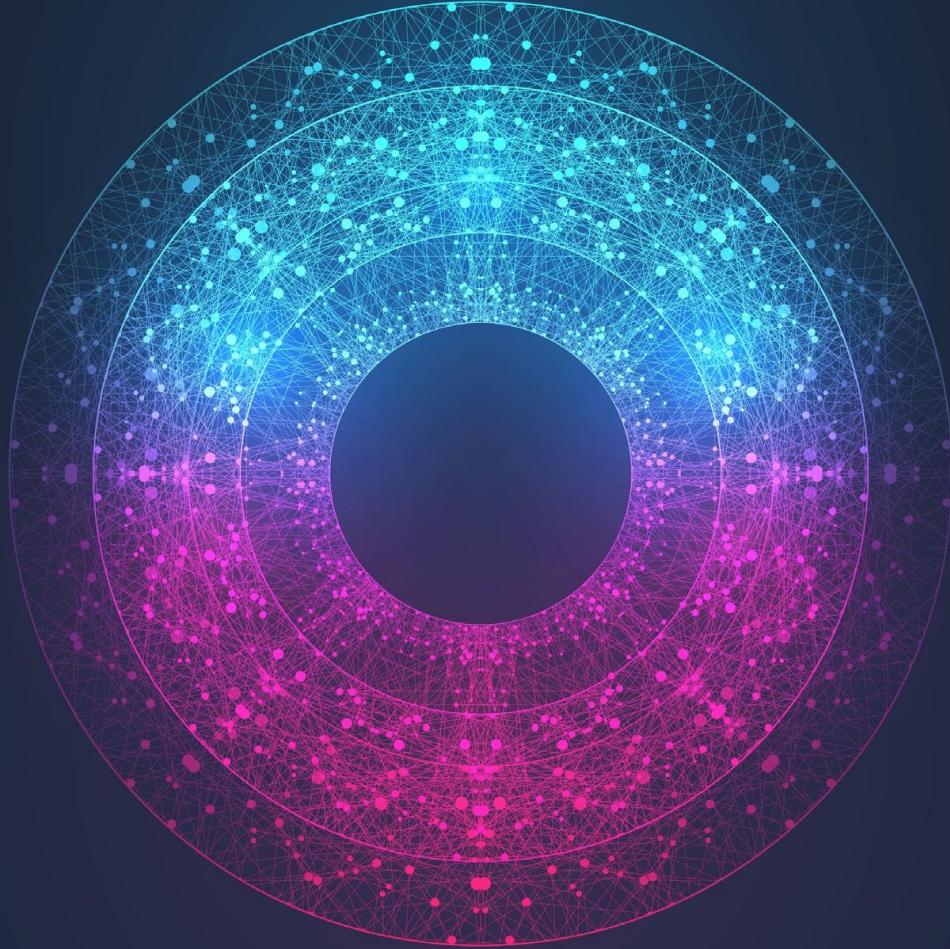
GLOBAL INTEREST IN AI DIED DOWN

AI COMPANIES WENT BANKRUPT

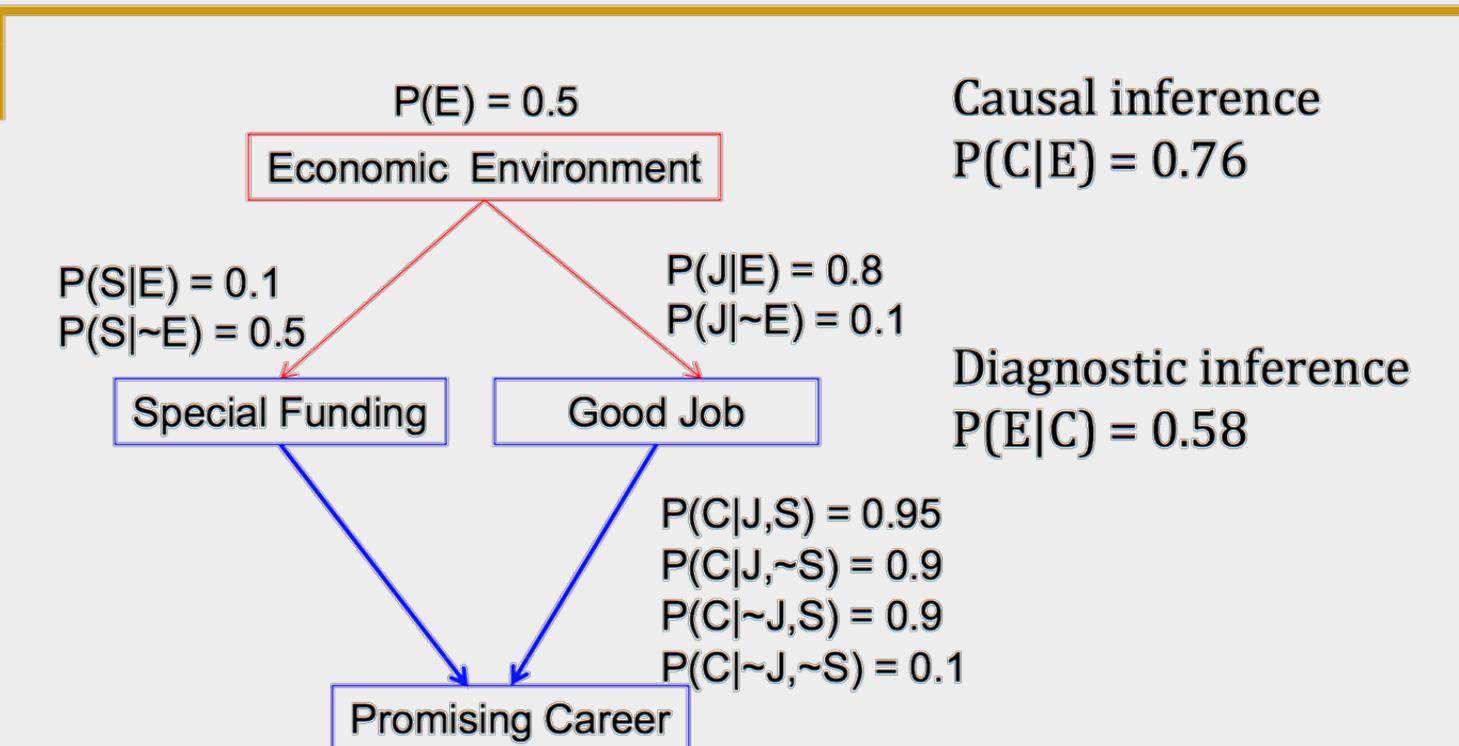
THIS SLIDE IS NOT A COPY FROM THE  
FIRST AI WINTER



# THE AI REVIVAL (1994-present)

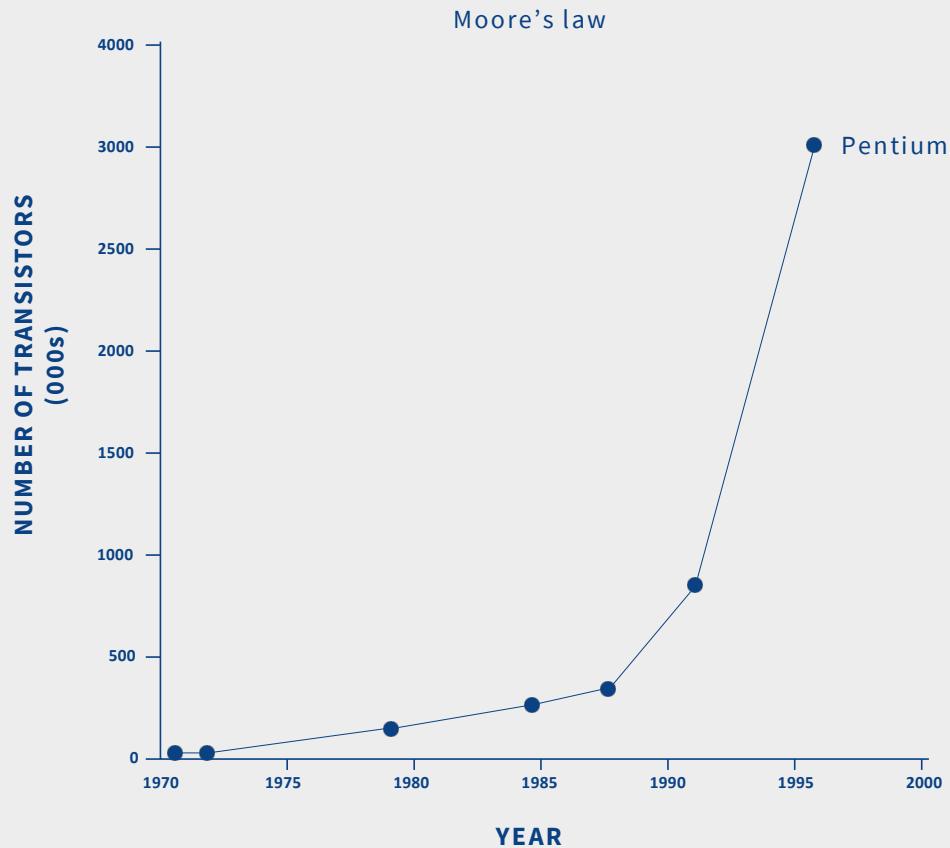


# CURRENT ERA: INTELLIGENT AGENTS AND DECISION THEORY



Bayesian Networks – Pearl (1988)

# CURRENT ERA: COMPUTATIONAL POWER

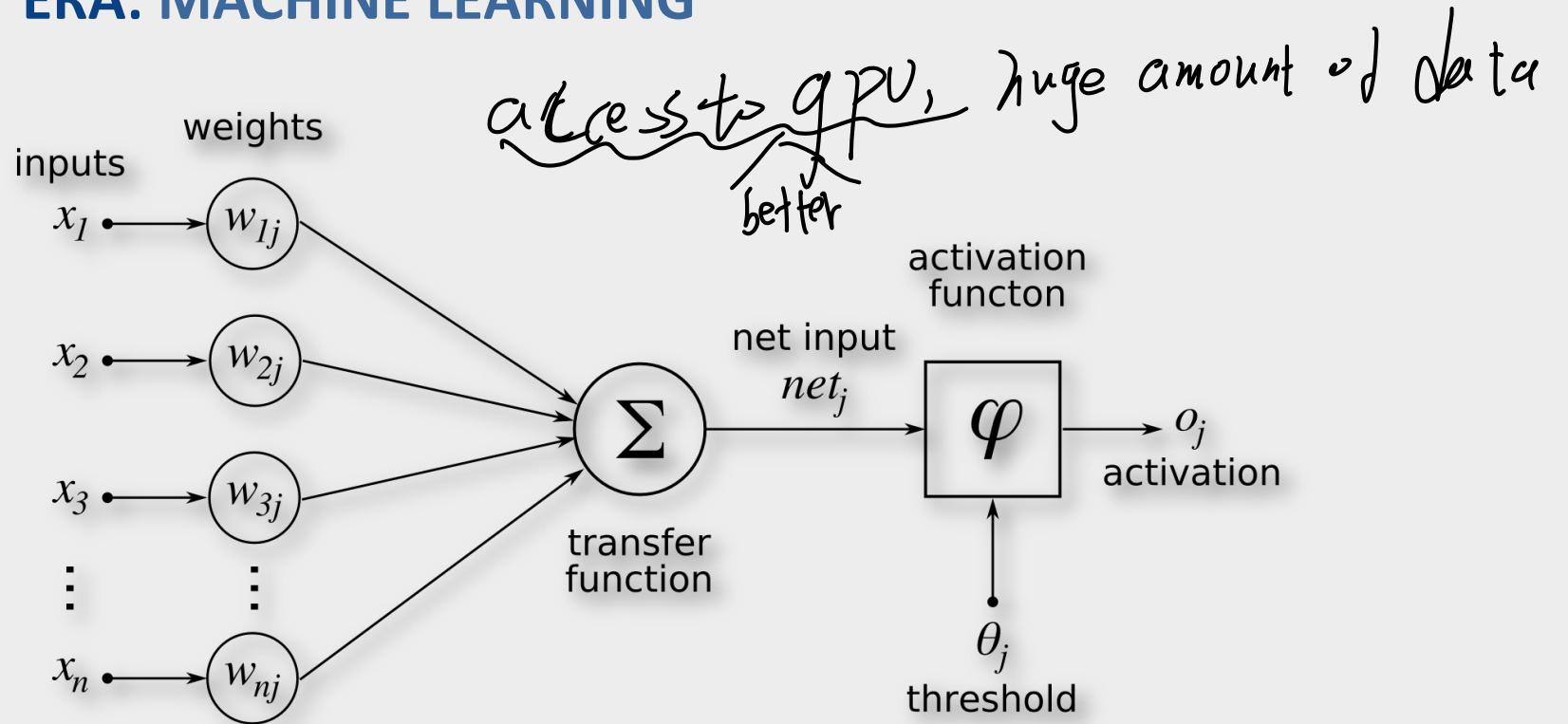


## CURRENT ERA: THE INTERNET AND BIG DATA



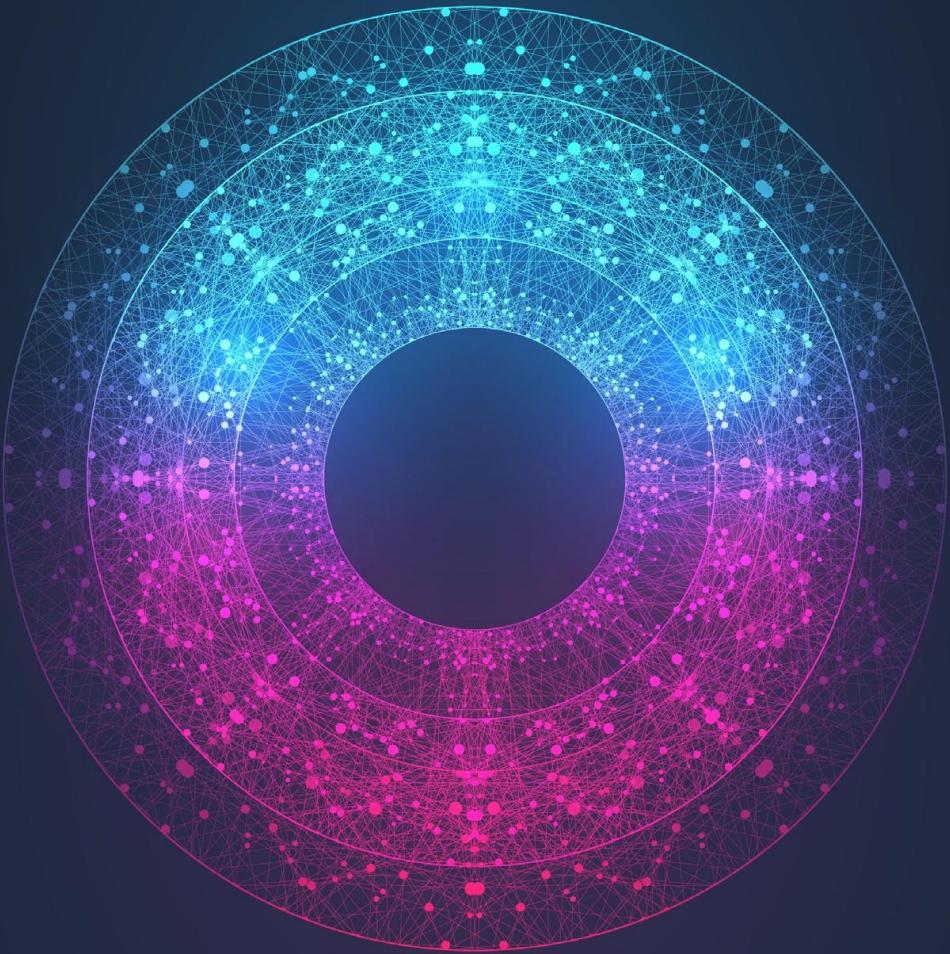
Image source: <https://www.promptcloud.com/blog/want-to-ensure-business-growth-via-big-data-augment-enterprise-data-with-web-data/>

# CURRENT ERA: MACHINE LEARNING



Backpropagation in deep neural networks – Rumelhart, Hinton, and Williams (1986)

# THE THIRD AI WINTER?



# WE'VE BEEN HERE BEFORE

“There is no reason and no way that a human mind can keep up with an artificial intelligence machine by 2035.”

— **Gray Scott** (2017)

“Artificial intelligence will reach human levels by around 2029. Follow that out further to, say, 2045, and we will have multiplied the intelligence – the human biological machine intelligence of our civilization – a billion-fold.”

— **Ray Kurzweil** (1999)

“In from three to eight years we will have a machine with the general intelligence of an average human being.”

— **Marvin Minsky** (1970)

“Machines will be capable, within twenty years, of doing any work a man can do”

— **Herbert Simon** (1956)

“We will have fully self-driving cars on the road by 2017” — **Elon Musk** (2014)

“... the embryo of an electronic computer that [the Navy] expects will be able to walk, talk, see, write, reproduce itself and be conscious of its existence.”

— **New York Times on the Perceptron** (1958)

We tend to overestimate the effect of a technology in the short run and underestimate the effect in the long run.

**Amara's Law**

~~A progress didn't match hypo => winter~~

## WHAT ARE SOME OF THE RISKS?

“By far the greatest danger of Artificial Intelligence is that people conclude too early that they understand it.”

— Eliezer Yudkowsky

“People worry that computers will get too smart and take over the world, but the real problem is that they're too stupid and they've already taken over the world.”

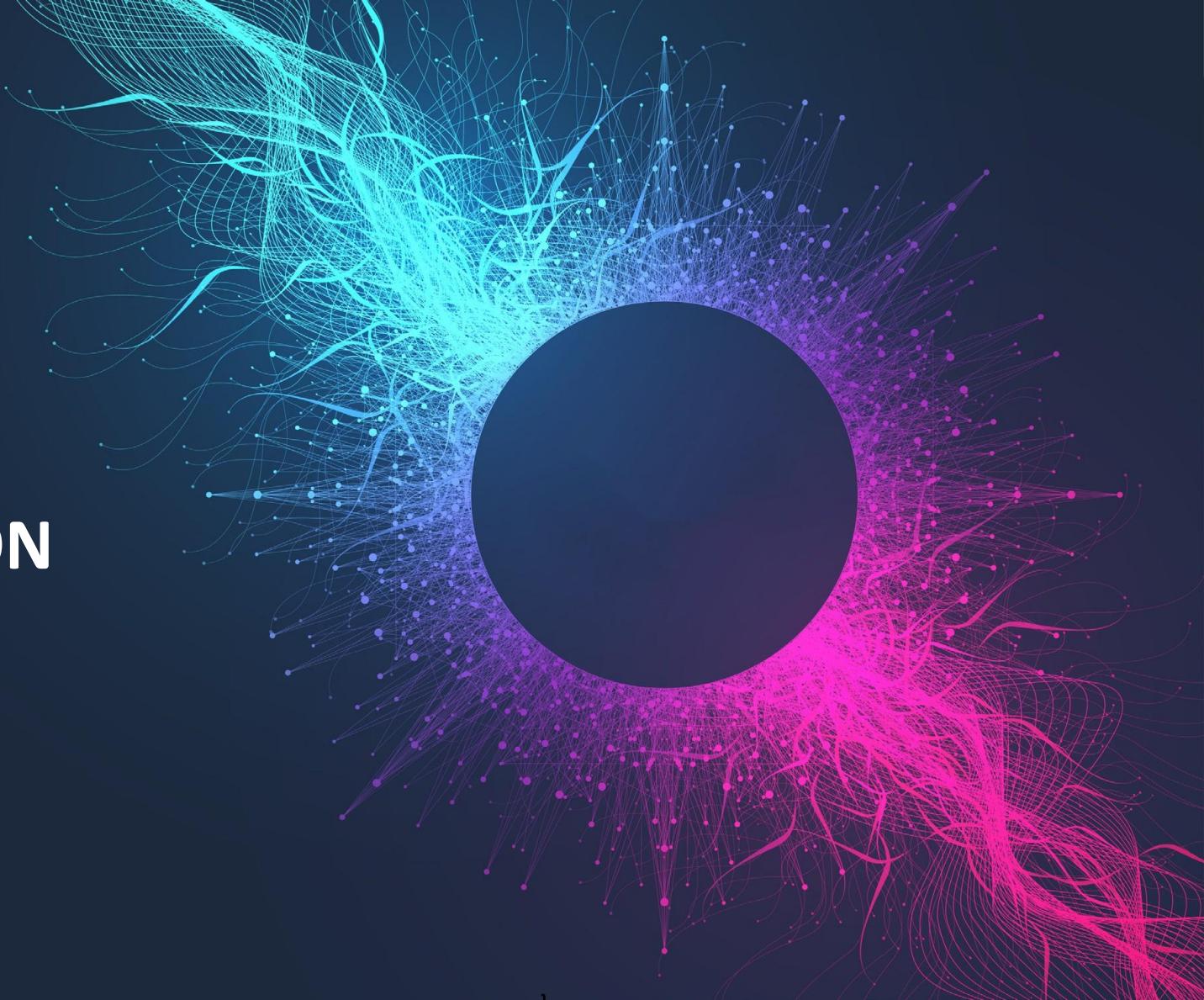
— Pedro Domingos

The beauty of #AI and what we can do with it is currently being overshadowed by reckless #hype, monotheistic techniques, and absurd deification. Yet, it's not the first time in 65+ years nor it'll be the last one. We haven't learnt a thing.

— Dagmar Monett



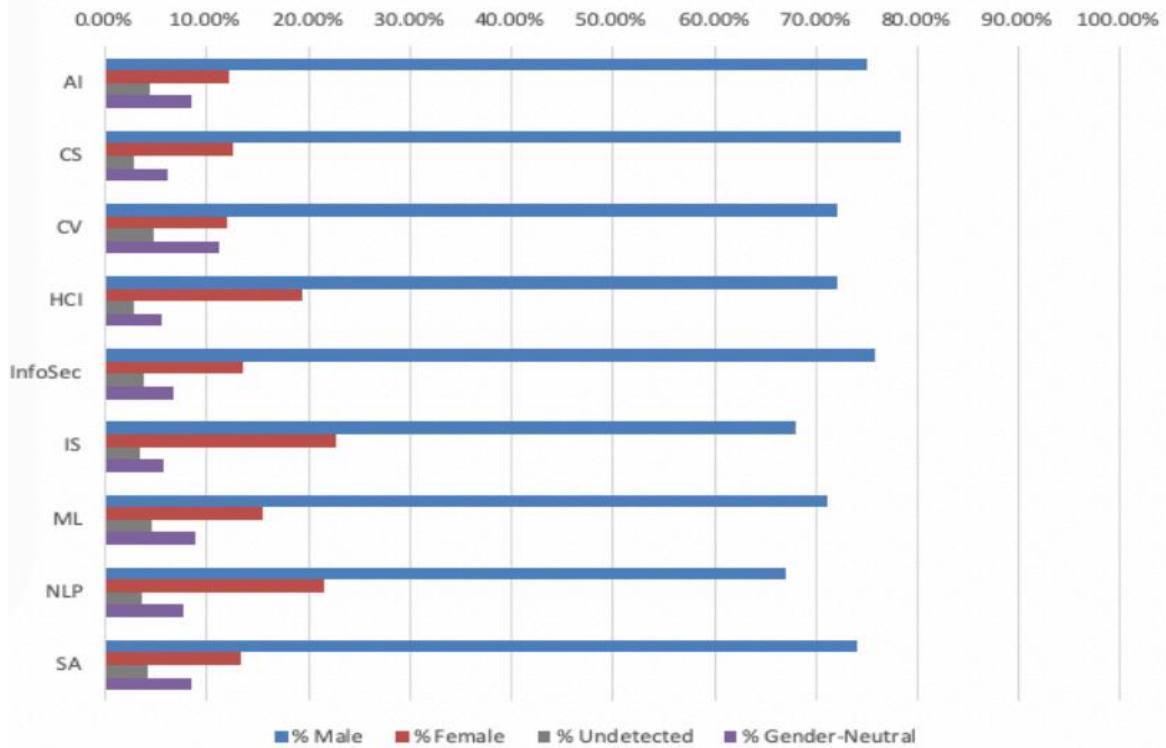
# HISTORY AND REPRESENTATION



histby

“Computer Science Communities: Who is Speaking, and Who is Listening to the Women? Using an Ethics of Care to Promote Diverse Voices”. Marc Cheong, Kobi Leins, Simon Coghlan. In ACM Conference on Fairness, Accountability, and Transparency (FAccT), 2021

# Contribution from diverse community & gender, UNDER-REPRESENTATION IN AI



# EFFECTS OF UNDER-REPRESENTATION

Lack of diversity means lack of:

---

PRIVACY

SAFETY

FAIRNESS

TRANSPARENCY

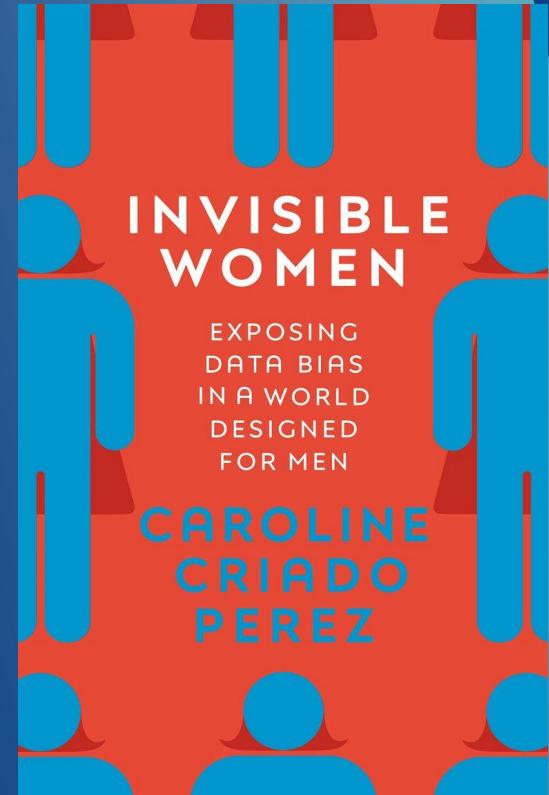
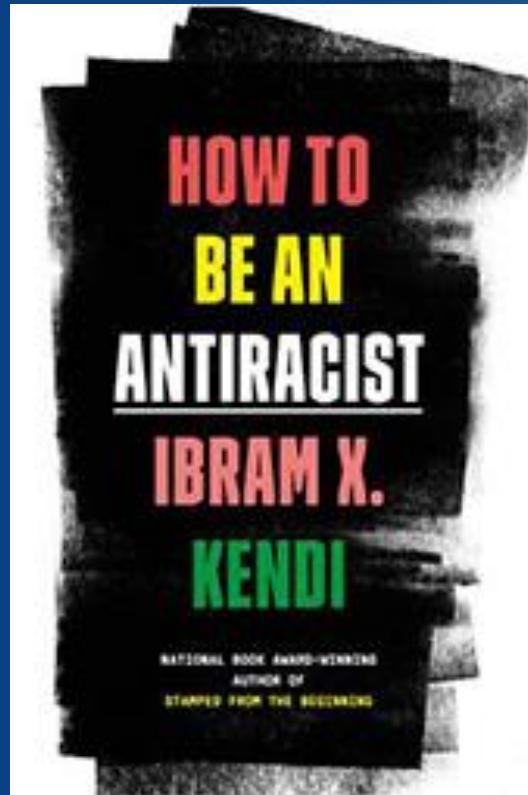
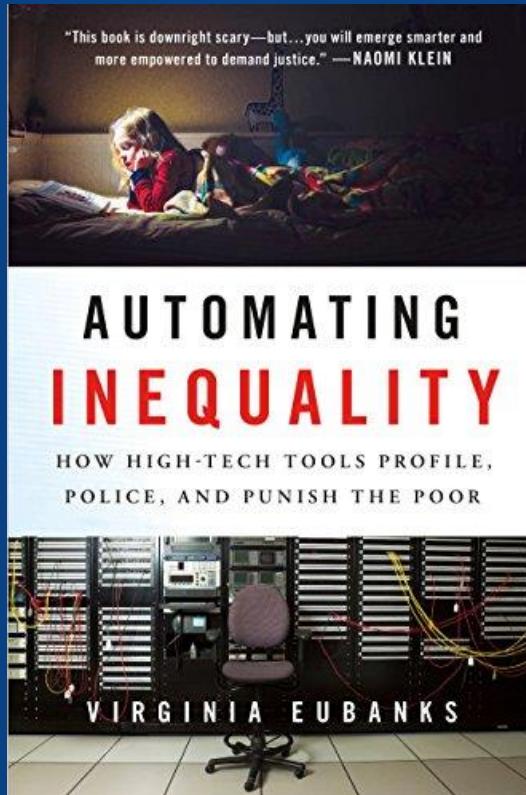
ACCESSIBILITY &  
INCLUSION

FUNCTIONALITY

Design decisions,  
data, attitudes,  
etc., are all  
influenced by who  
we are as  
individuals, as  
teams, and as  
societies

The History of AI ⊑ The History of Culture and Society

## FURTHER READING



白人皮肤  
[白人皮膚]  
能够洗滌只很外

设计为用户  
生育身形设计。

## DIVERSITY

### DIVERSITY IN TEAMS

- Gender
- Culture
- Ethnicity
- Sexual orientation
- Disability

- Family status
- Age
- Class
- Education
- Etc.

### DIVERSITY OF INPUTS

- Work with users
- Get out of the building!

This is not just good for the soul:  
it is good for business!

# HISTORY OF AI: SUMMARY

## HISTORY

Dartmouth workshop is the “birth” of artificial intelligence

AI winters caused by hyped expectations not being met

Eras of artificial intelligence

GOLDEN ERA

KNOWLEDGE ERA

REVIVAL ERA

Will we have another AI winter/autumn?

## HISTORY AND REPRESENTATION

Artificial intelligence (including ethics of AI) has been driven mostly by male, western culture

Important contributions from non-male, non-Western culture, but still marginalised

Culture (and therefore its history) influences design decision of software systems

Diversity

DIVERSE TEAMS

DIVERSE INPUTS



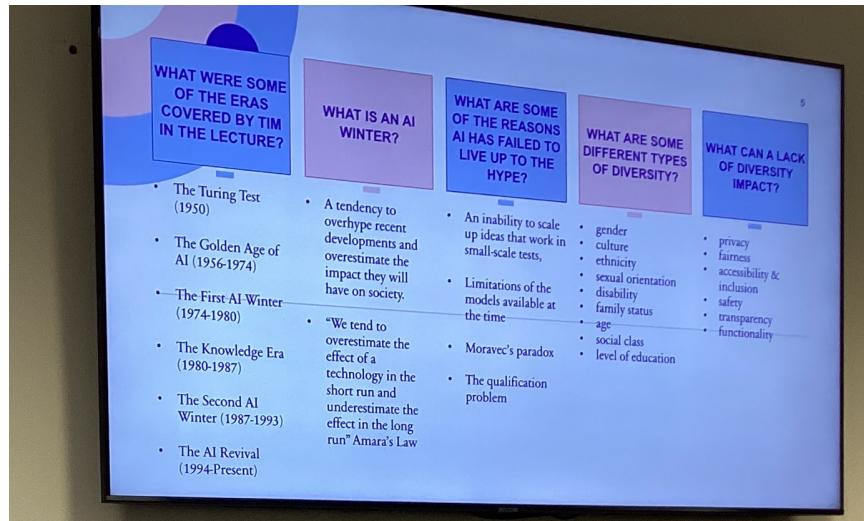
THE UNIVERSITY OF  
MELBOURNE

# THANK YOU

**Professor Tim Miller**

School of Computing and Information Systems  
Co-Director, Centre for AI & Digital Ethics  
The University of Melbourne  
[@tmiller\\_unimelb](https://twitter.com/tmiller_unimelb)





/,