Data & risks

Prof. dr. Marielle Stoelinga
University of Twente | Radboud University Nijmegen

The first algorithm?

BC

1843

1956

1997

The first algorithm!

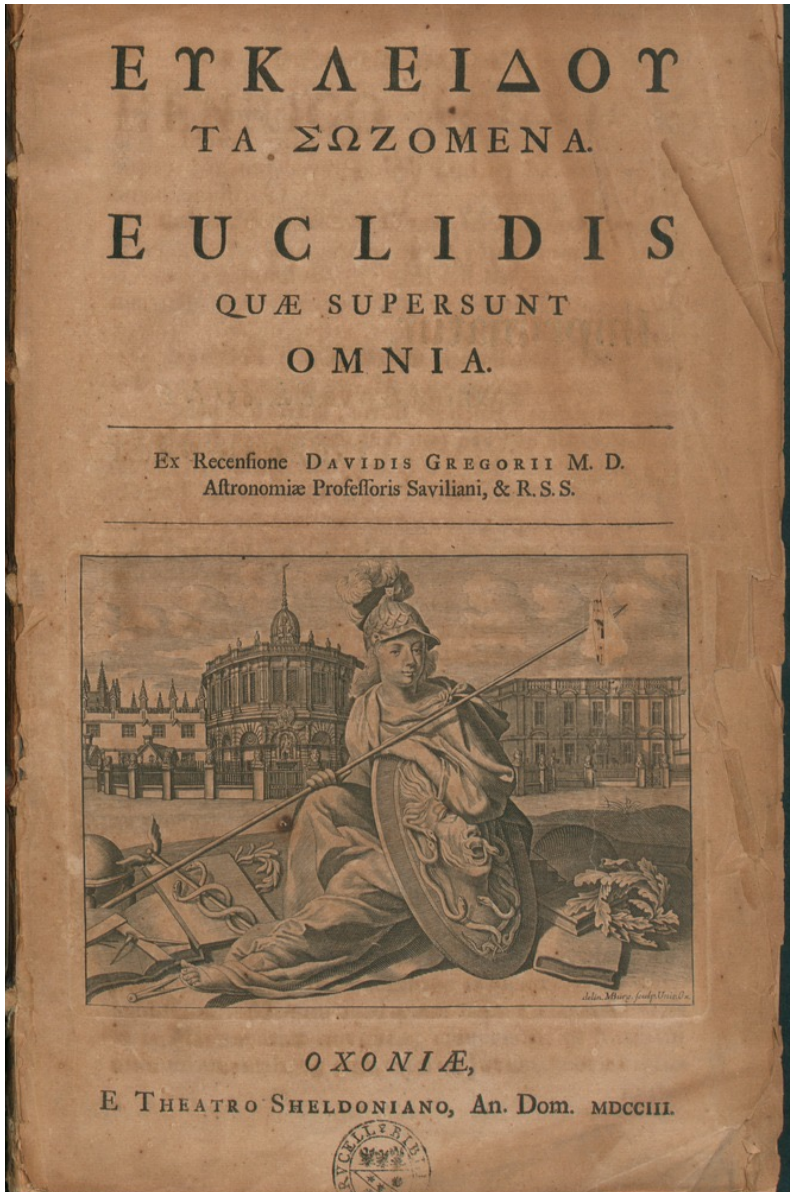
300 BC

- Euclid / Euclides
- Ca. 325 – 270 BC
- Alexandria, Greece
- Mathematician

Algorithm still used today!


Raphael
School of Athens
ca 1500

Claim to fame: The elements

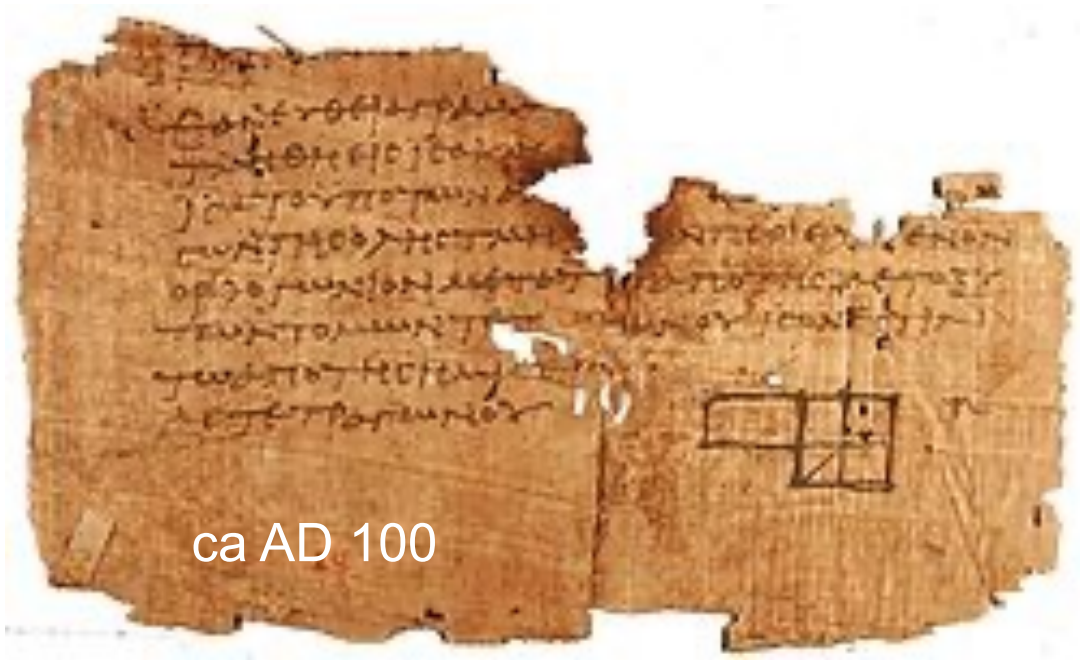

ΕΥΚΛΕΙΔΟΥ
ΤΑ ΣΩΖΟΜΕΝΑ.

EUCLIDIS
QUÆ SUPERSUNT
OMNIA.

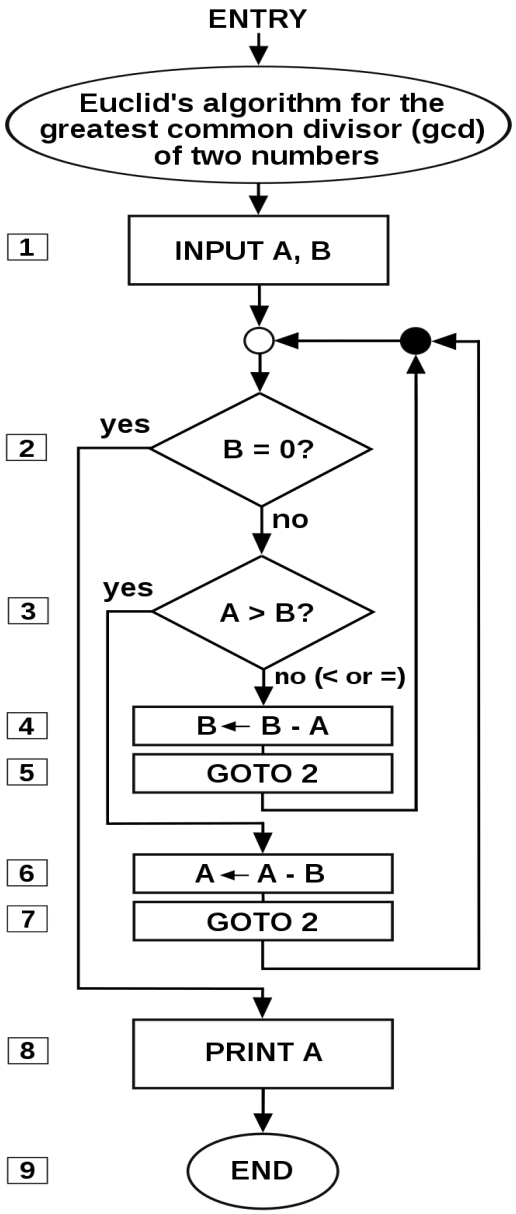
Ex Recenſione DAVIDIS GREGORII M. D.
Aſtronomiæ Profeſſoris Saviliani, & R. S. S.

OXONIÆ,
E THEATRO SHELDONIANO, An. Dom. MDCCIII.

- Numbers & geometry (triangles, cones)
- Euclid: first come up with procedure
- Algorithm
  - step wise procedure =
  - list of instructions =
  - recipe



For nerds only: Euclid's algorithm



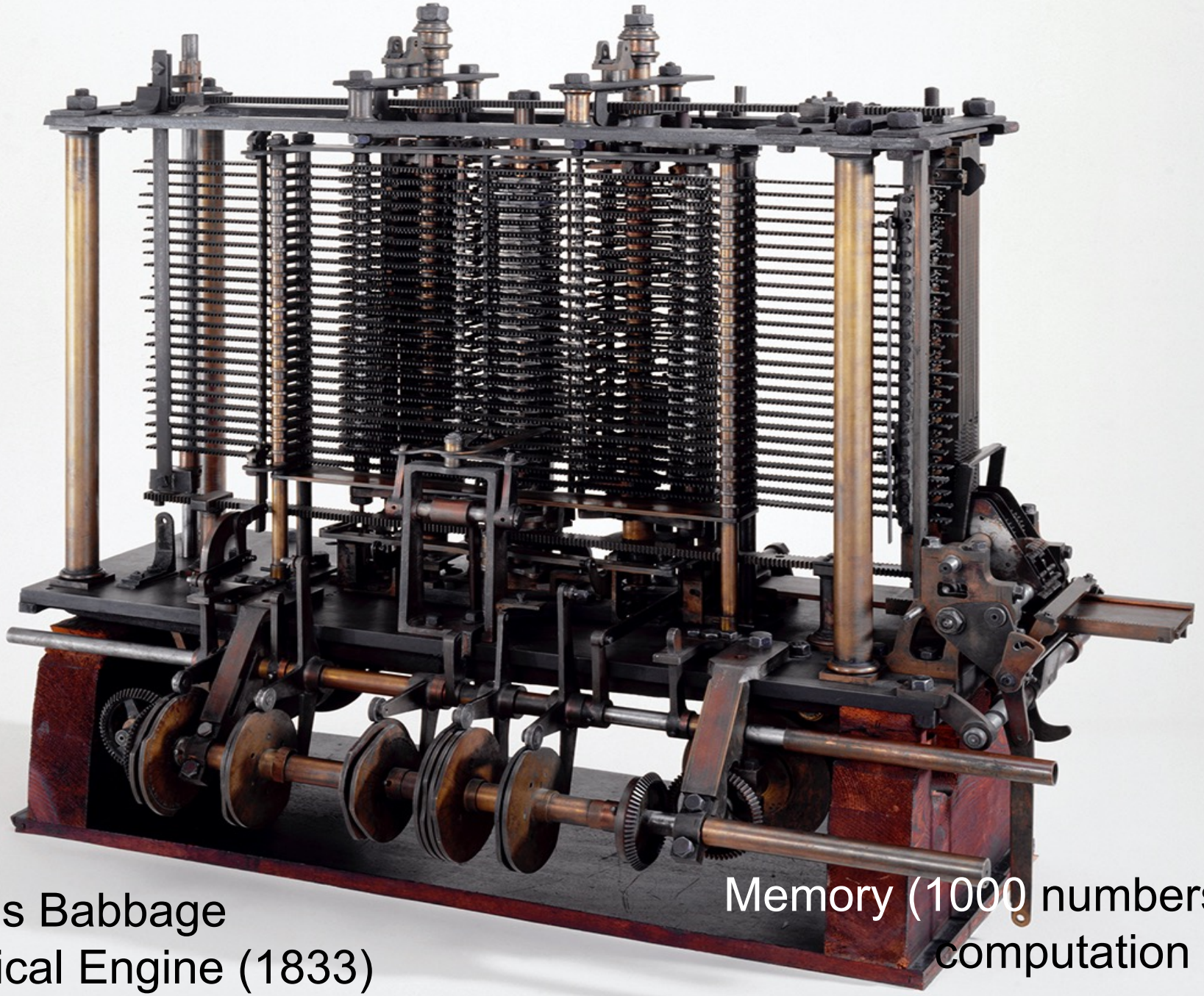
- computes greatest common divisor
- gcd(30,18) = 6

Idea / trick / magic

gcd(30,18) =
= gcd(30-18,18) = gcd(12,18)
= gcd(12,18-12) = gcd(12,6)
= gcd(12-6,6) = gcd(6,6) =
= gcd(6,6-6) = gcd(6,0) = 6

Algorithms: no computer needed

The first algorithm for computers



Charles Babbage
Analytical Engine (1833)

Memory (1000 numbers) +
computation unit

Ada Lovelace

- Lived 1815 –1852
- Daughter of Lord Byron

Analytical Engine

- Translated + extended description
- Invented calculation method for Bernoulli numbers

→ first computer algorithm

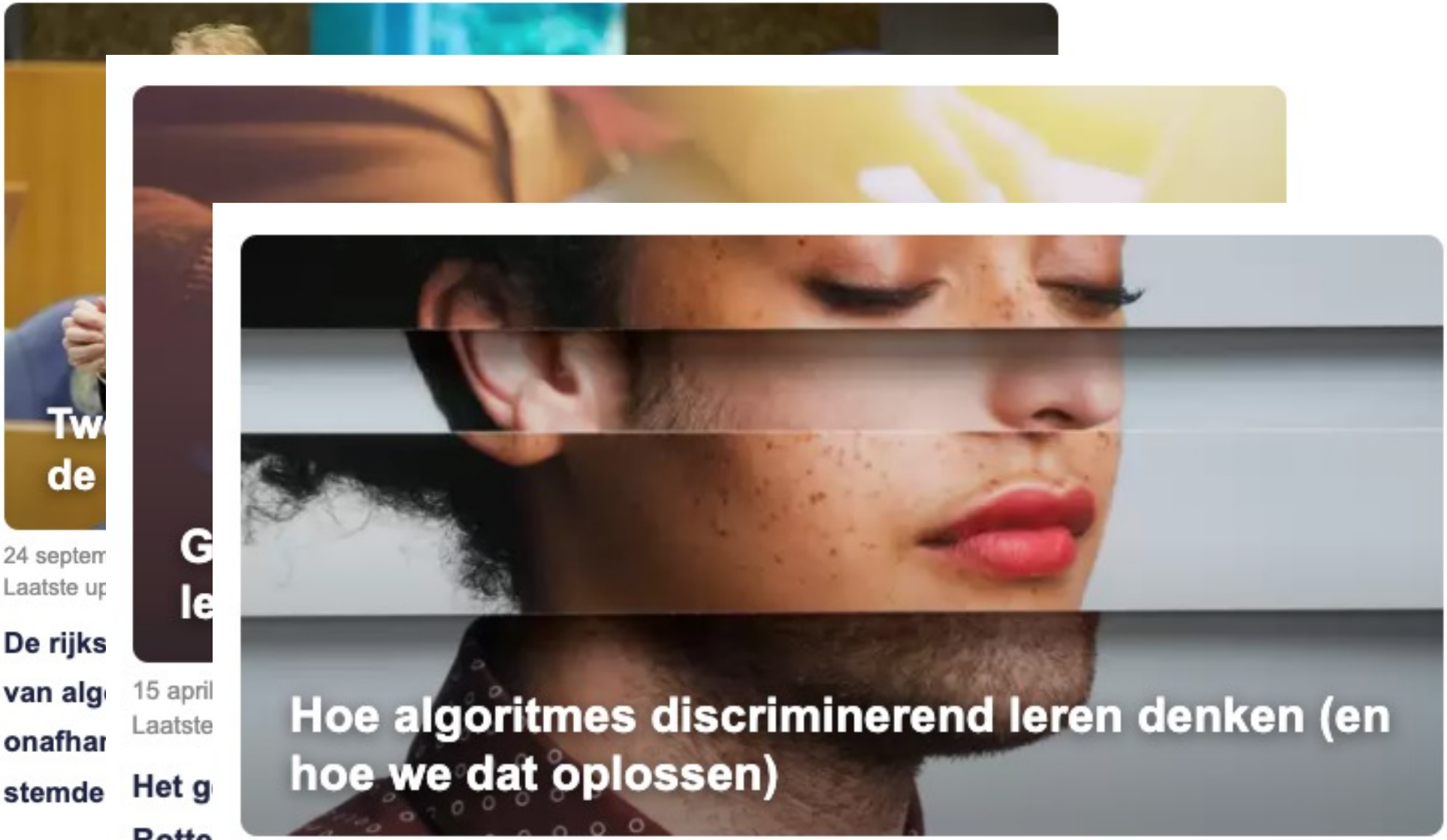
Vision

computers go beyond calculating + number crunching

From Ada's notes

Number of Operations	Nature of Operations	Variables for Data						Working Variables									Variables for Results	
		$^1V_0$	$^1V_1$	$^1V_2$	$^1V_3$	$^1V_4$	$^1V_5$	$^0V_6$	$^0V_7$	$^0V_8$	$^0V_9$	$^0V_{10}$	$^0V_{11}$	$^0V_{12}$	$^0V_{13}$	$^0V_{14}$	$^0V_{15}$	$^0V_{16}$
		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		m	n	d	m'	n'	d'										$\frac{dn'-d'n}{mn'-m'n} = x$	$\frac{d'm-dm}{mn'-m'n} = y$
1	×	m	....	....	....	n'	....	mn'										
2	×	....	n	....	m'	....	....	....	$m'n$									
3	×	....	....	d	....	....	....	....	....	dn'								
4	×	....	0	....	....	....	d'	....	....	....	$d'n$							
5	×	0	....	....	....	....	0	....	....	....	....	$d'm$						
6	×	....	....	0	0	....	....	....	....	....	....	.....	dn'					
7	−	....	....	....	....	....	....	0	0	....	....	.....	.....	$(mn' - m'n)$				
8	−	....	....	....	....	....	....	....	....	0	0	.....	.....	.....	$(dn' - d'n)$			
9	−	....	....	....	....	....	....	....	....	....	....	0	0	.....	.....	$(d'm - dn')$		
10	÷	....	....	....	....	....	....	....	....	....	....	.....	.....	$(mn' - m'n)$	0	.....	$\frac{dn'-d'n}{mn'-m'n} = x$	
11	÷	....	....	....	....	....	....	....	....	....	....	.....	.....	0	.....	0	.....	$\frac{d'm-dm'}{mn'-m'n} = y$

Algorithms: What about the fuss?



The first self-learing algorithms

1956

AI algorithms

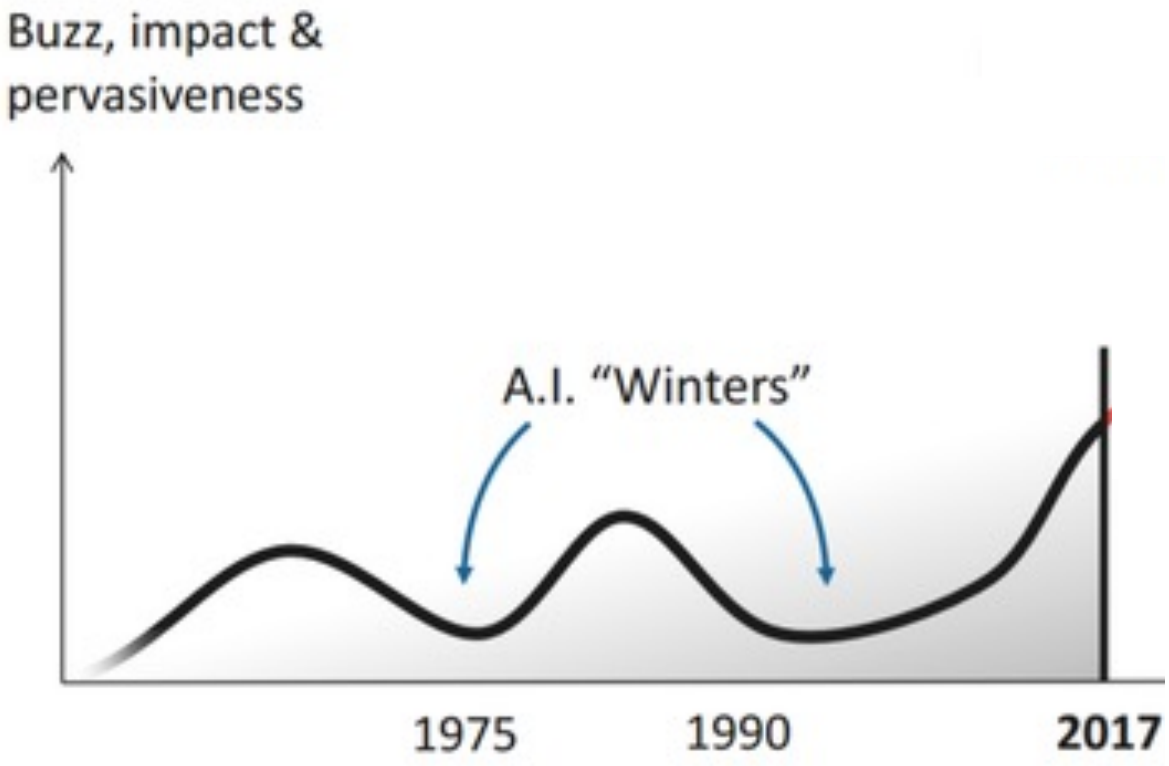
- Not programmed by humans
- But: Self-learning
  - Learn automatically from data / examples
  - Eg: cats vs dogs

Buzz words

- Artificial Neural Networks (“deep learning”)
- Bayesian Networks
- Genetic algorithms
- …..

- 1956 darthmouth workshop
- The problem: *they did not work!*

AI winters



Winters
- AI solutions failed
- No investments

Why these winters?
- Problems are complex
- Computer power + data limited


AI Winter
1969 - 1990

The first working AI algorithms

1997

- DeepBlue beats Kasparov
- Classical algorithms

2013

- AlphaGO defeat2 Lee Sedol
- Neural networks

Why did it suddenly work?

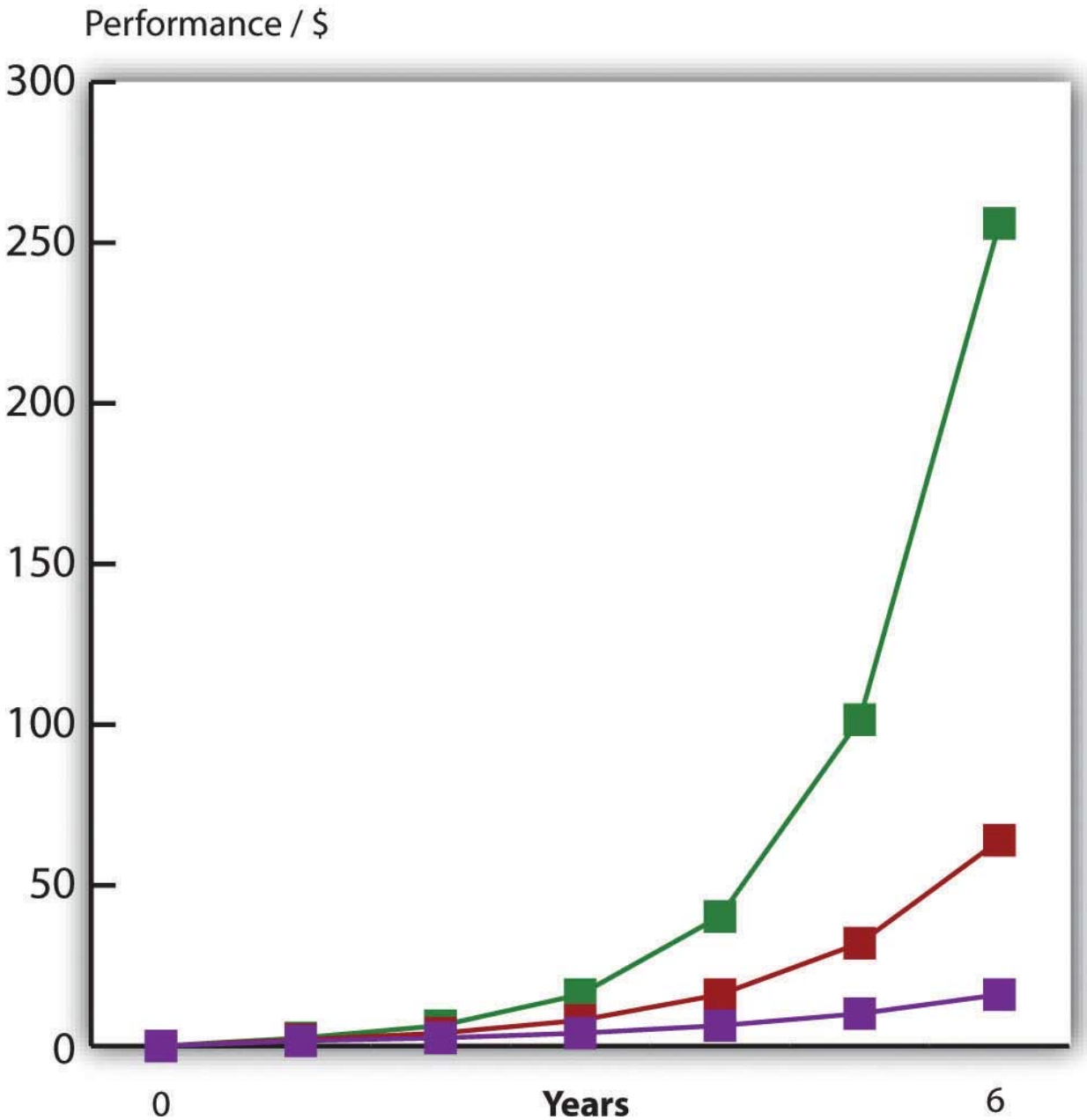
Increased computer power

- Moore's law

Data storage cheaper

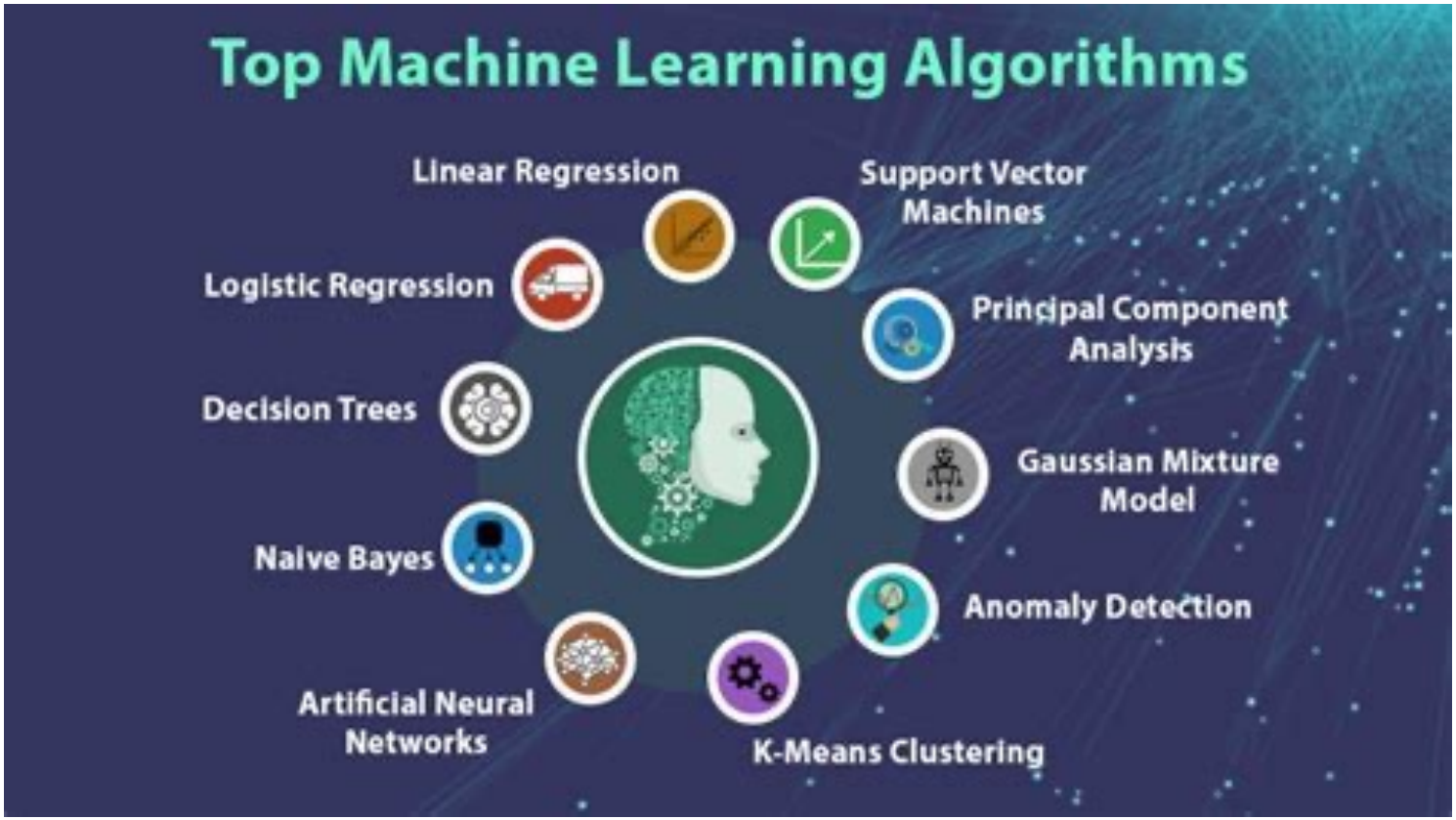
- More data available

Exponential growth: computing power & data storage



- Moore's Law— Doubles 18 months (number of transistors)
- Data Storage— Doubles 12 months (bits per square inch)
- Optical Fiber— Doubles 9 months (bits per second)

See keynote Prof. Vossen


Top Machine Learning Algorithms
Linear Regression
Support Vector Machines
Logistic Regression
Principal Component Analysis
Decision Trees
Gaussian Mixture Model
Naive Bayes
Anomaly Detection
Artificial Neural Networks
K-Means Clustering

Summary

300 BC

- First classical algorithm: *stepwise procedure*

1843

- First computer algorithm: *on analytical machine*

1956

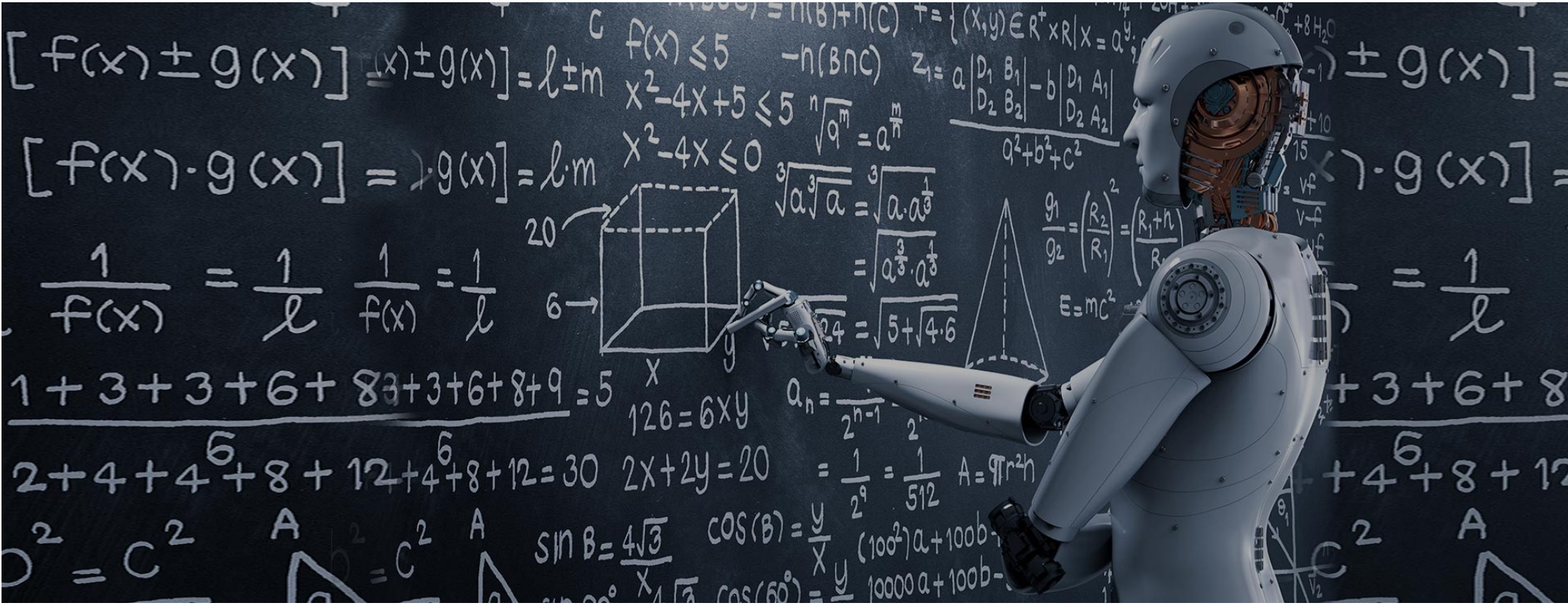
- First AI algorithms: *self-learning*

1997

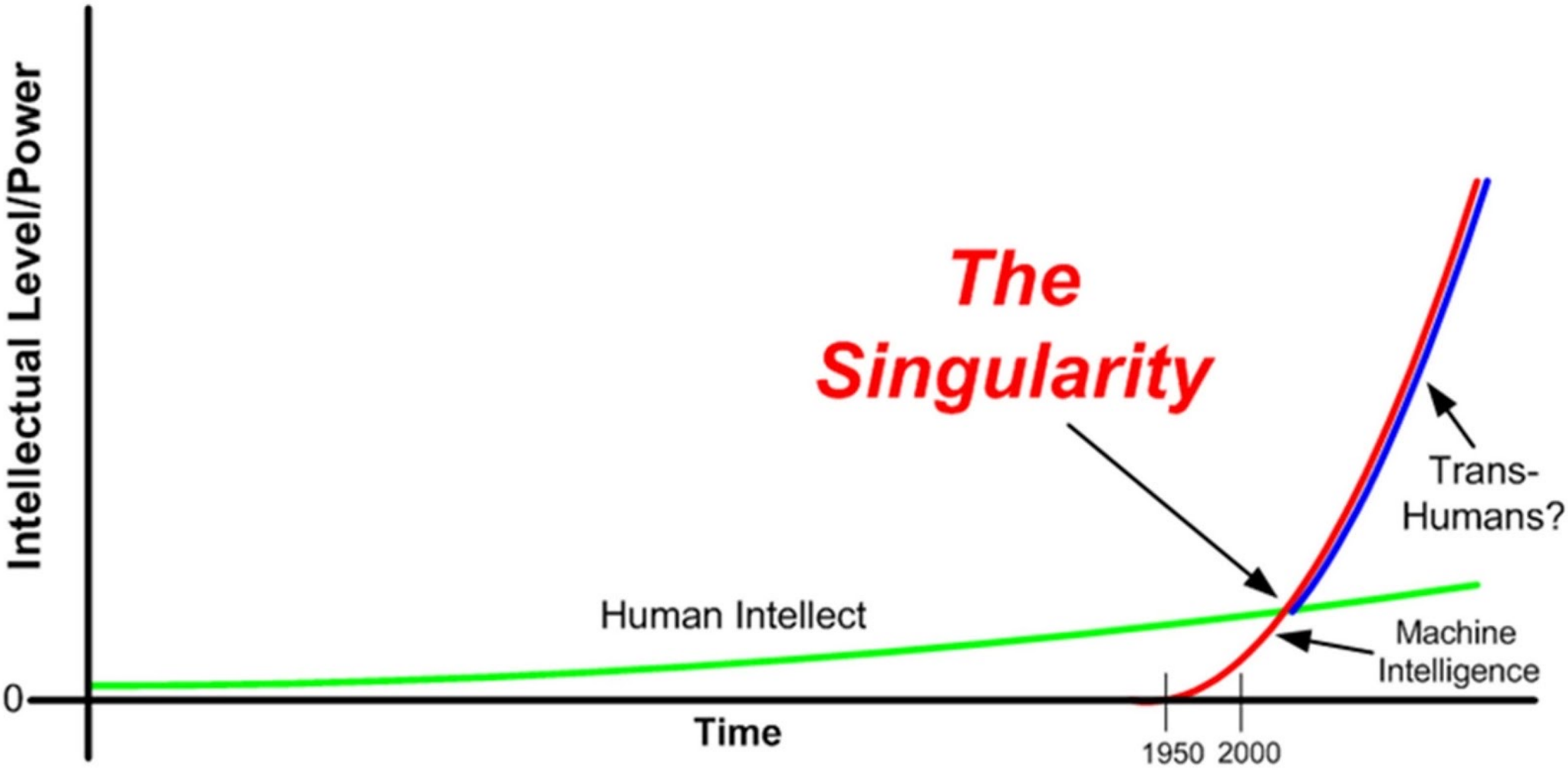
- First AI algorithms: *now effective*

If you hear the word “algorithm”, is it human-written or self-learning?

Where will we end up?



Where will we end?

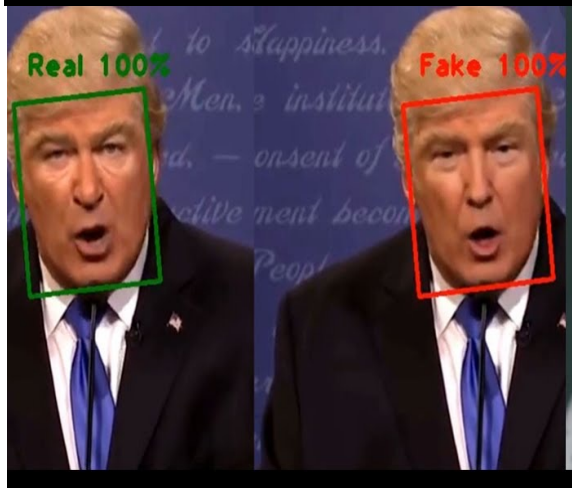


How does it all work?

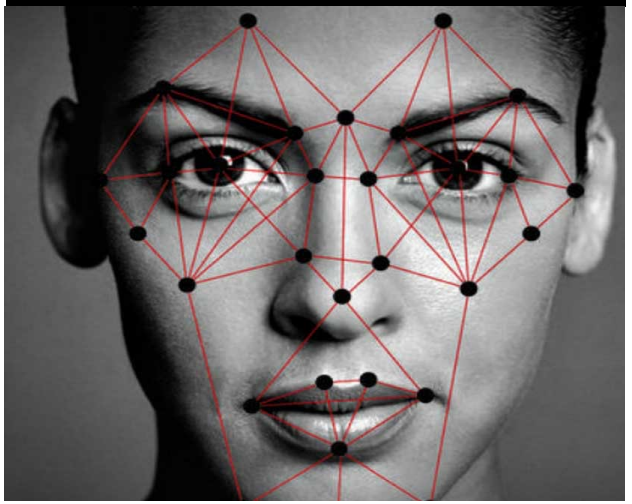
Applications

Self-driving cars

Deep fake



Face recognition



Sales forecasting

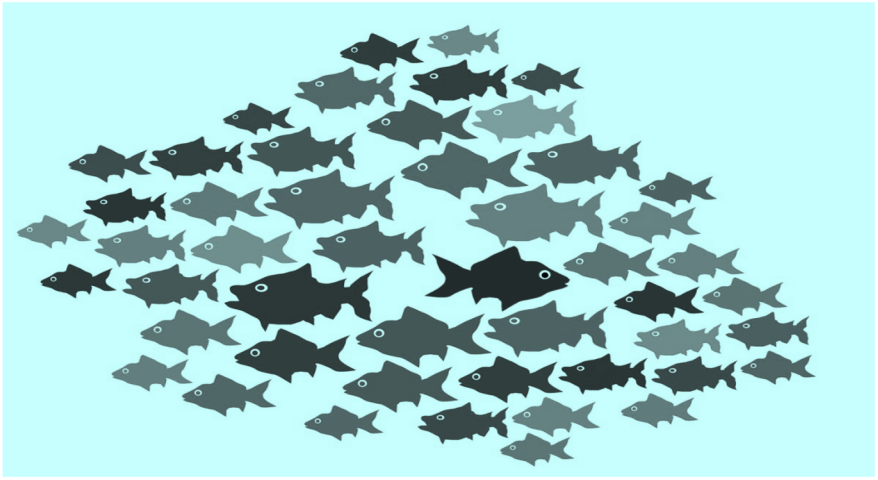
profiling

Applications in Risk Analysis

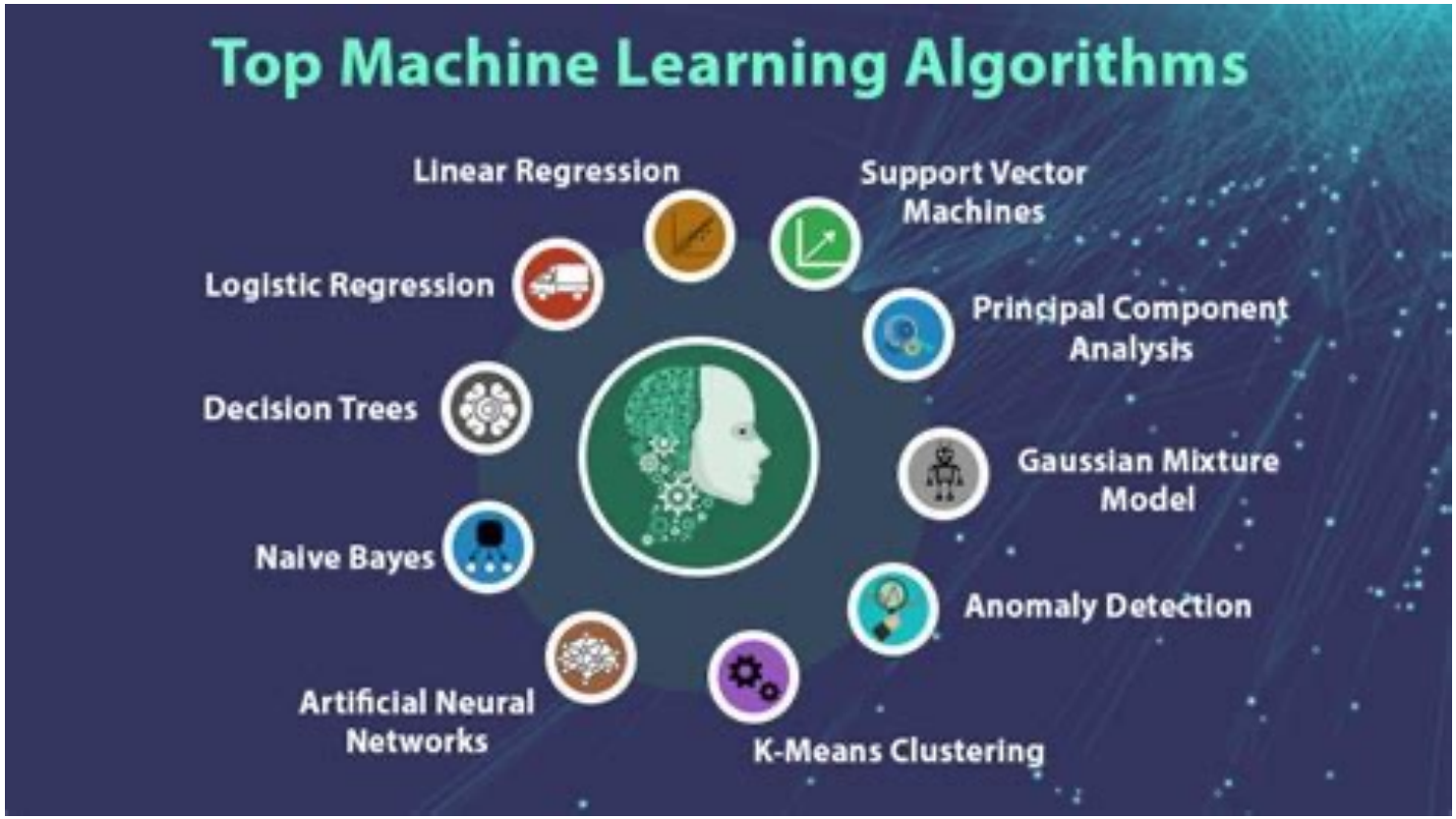
Spam filtering

“Spoorlopers”

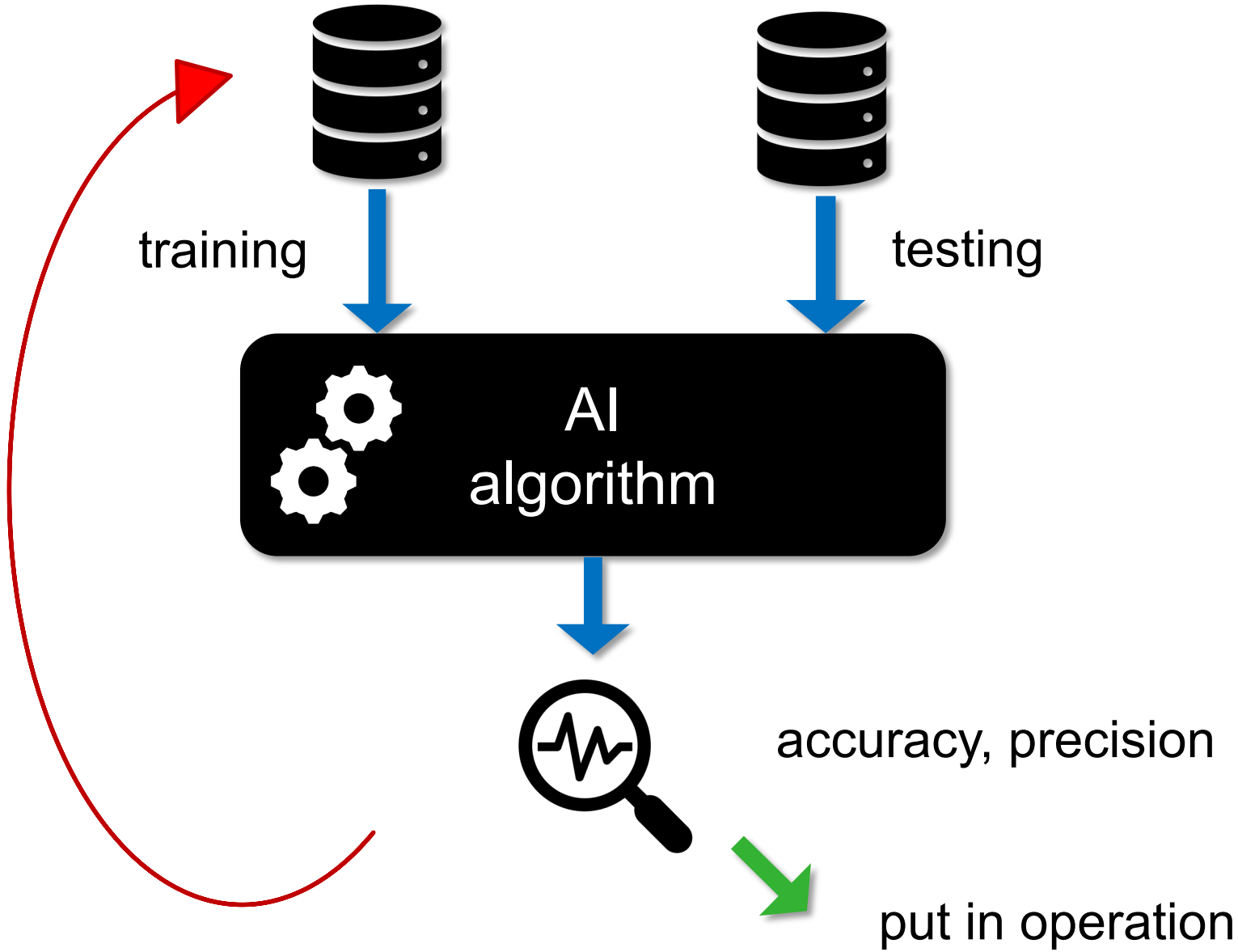
Anomaly detection



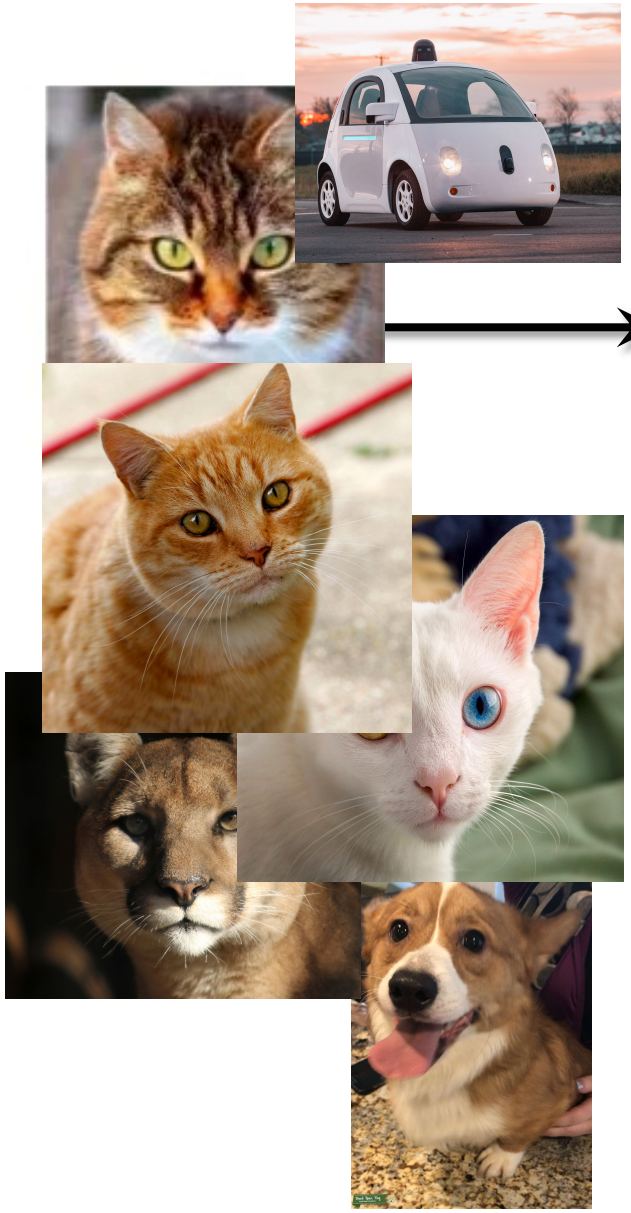
Predictive maintenance


Top Machine Learning Algorithms
Linear Regression
Support Vector Machines
Logistic Regression
Principal Component Analysis
Decision Trees
Gaussian Mixture Model
Naive Bayes
Anomaly Detection
Artificial Neural Networks
K-Means Clustering

How it works: Self-learning algorithms



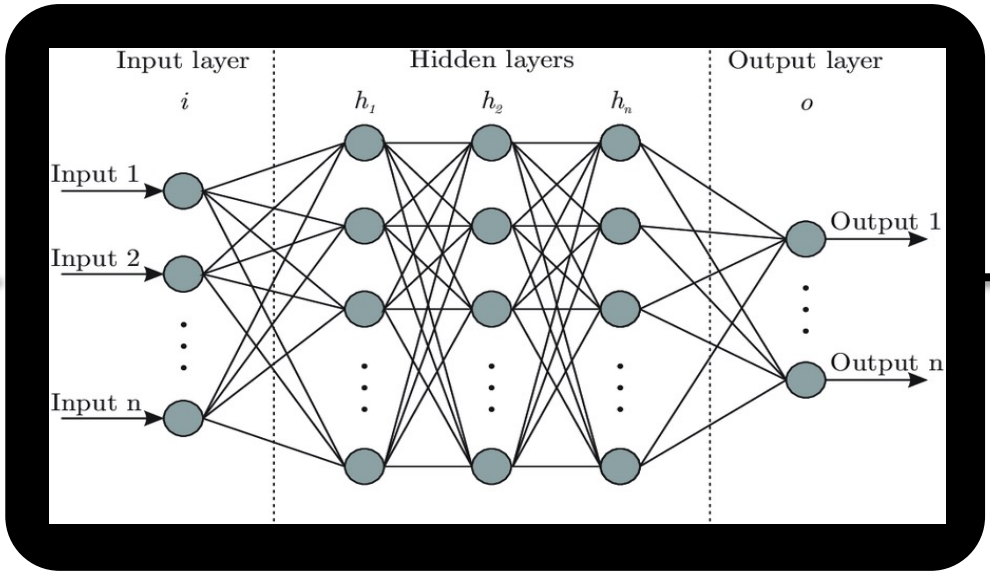
example: neural networks



Black box algorithm

cat:
yes / no

neural networks: inside

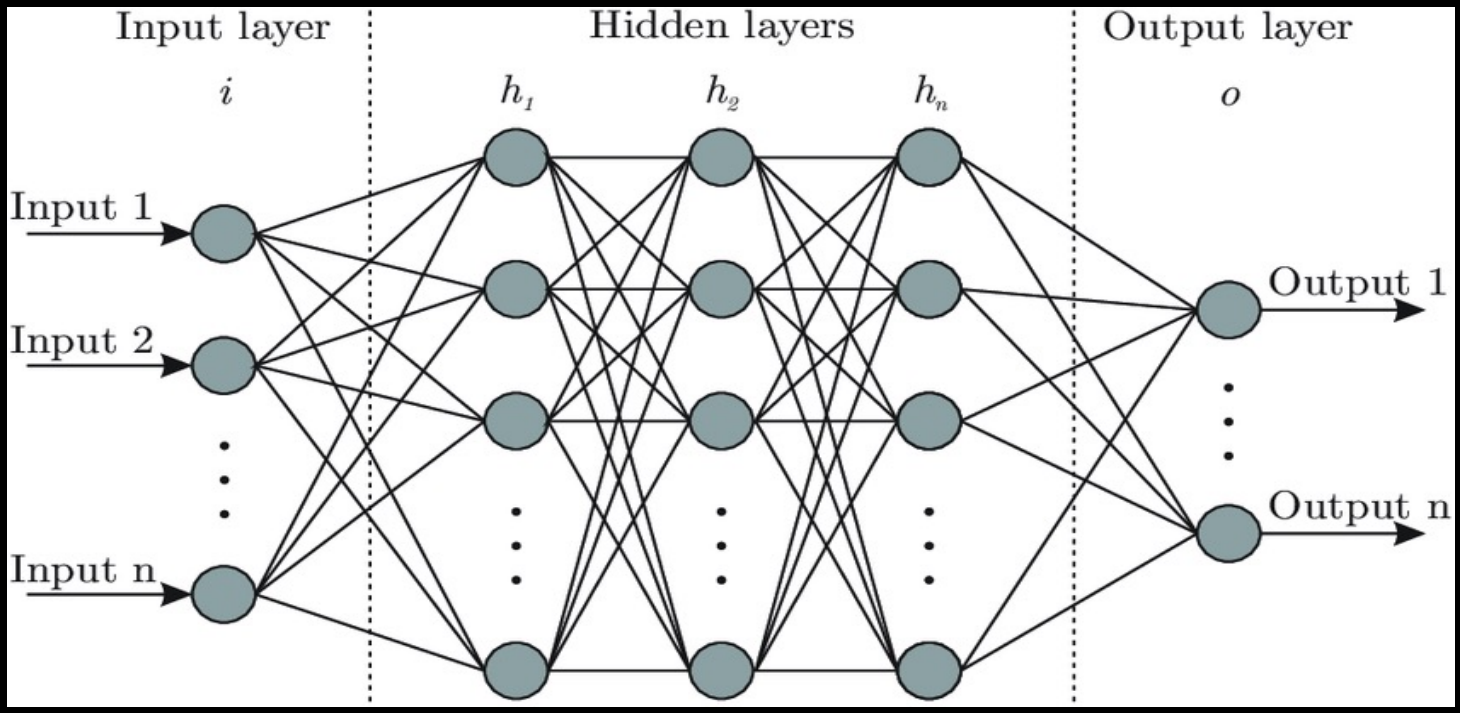


cat:
yes / no

	algo: cat	algo: no cat
picture: cat	true positive	false negative
picture: no cat	false positive	true negative

Accuracy = % correctly classified pictures

neural networks: inside



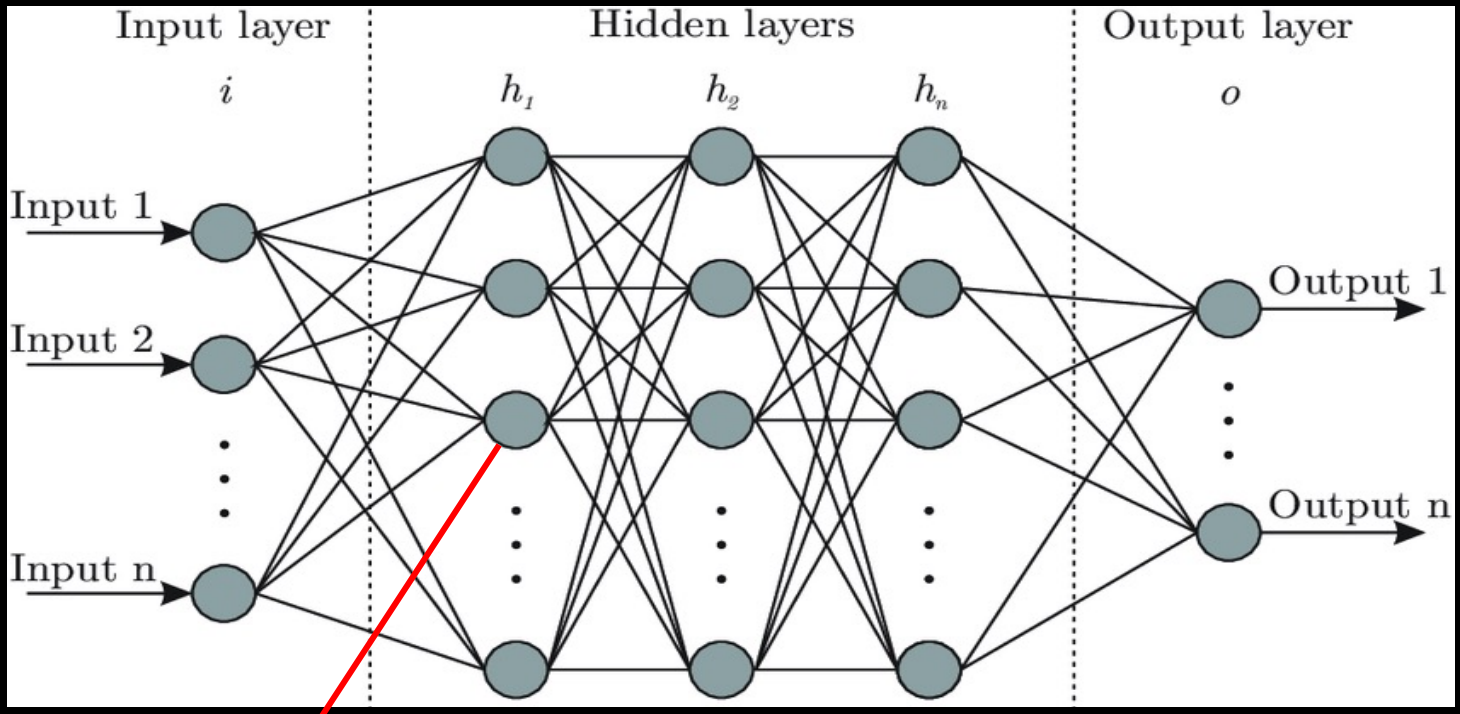
inputs

- (cat) picture: 64 x 64 pixels
- grey scale in [0,1]: 0=white, 1= black, 0.8 rather black

Output

- binary: 1 (yes, cat) or 0 (no cat)

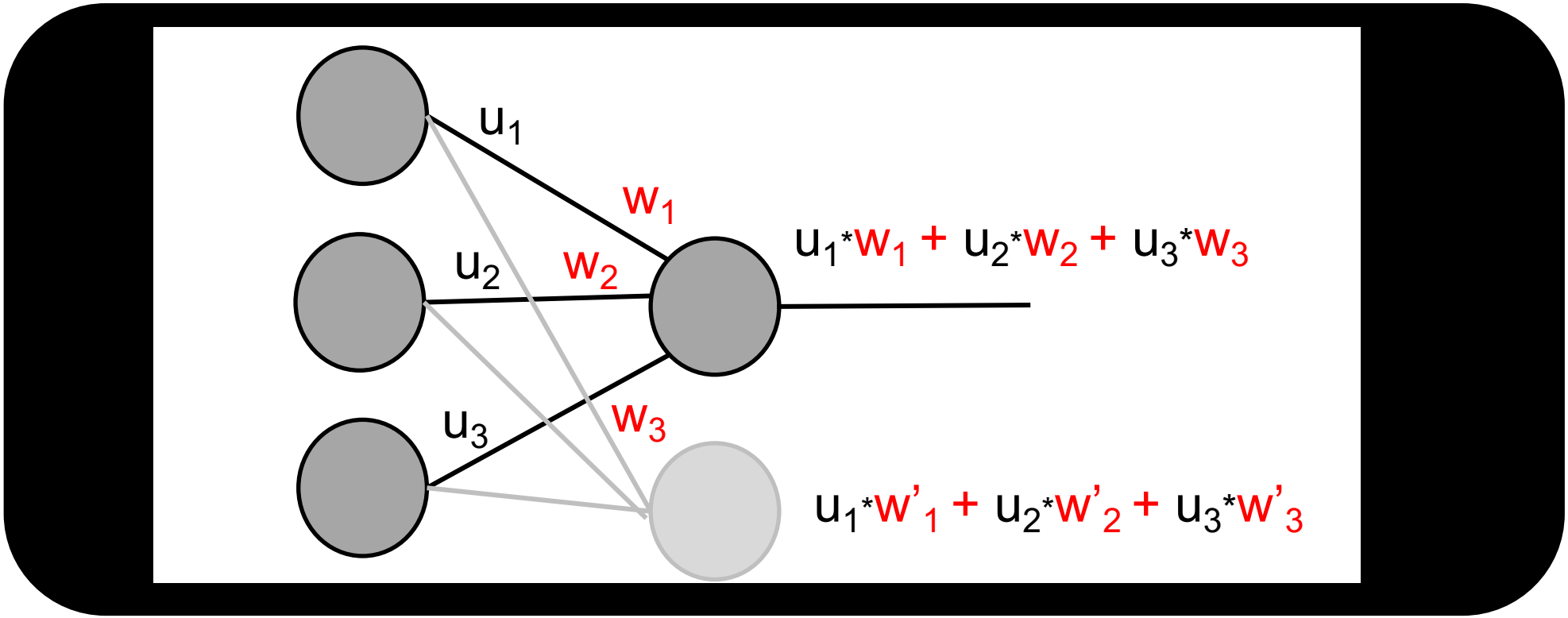
neural networks: inside



firing neurons

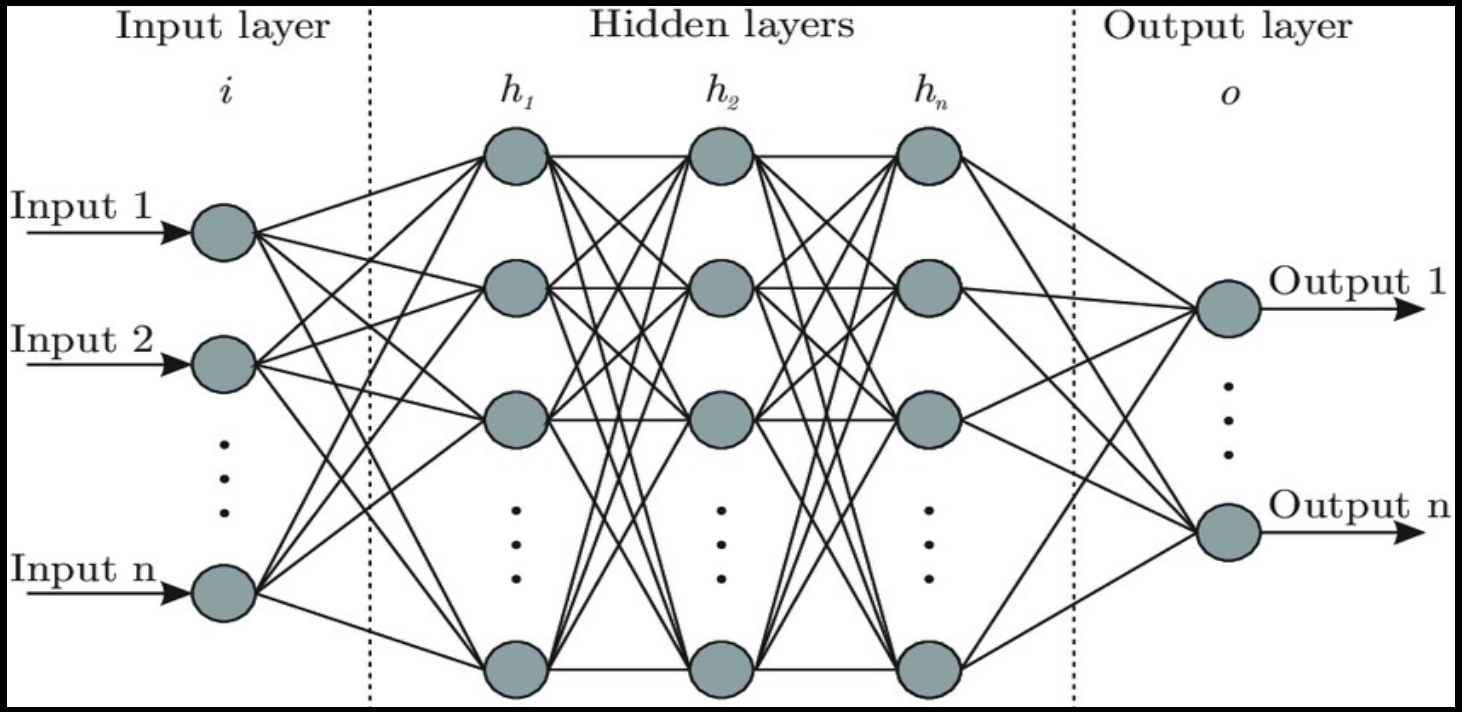
- a neuro fires iff their weighted sum of its inputs > T
- training = finding the weights
- many variants: recurrent, convolutional, long short term memory

neural networks: inside



- a neuro fires iff their weighted sum of > T
- (renormalize to get numbers in [0,1])
- Learning = finding good weights

neural networks: black box

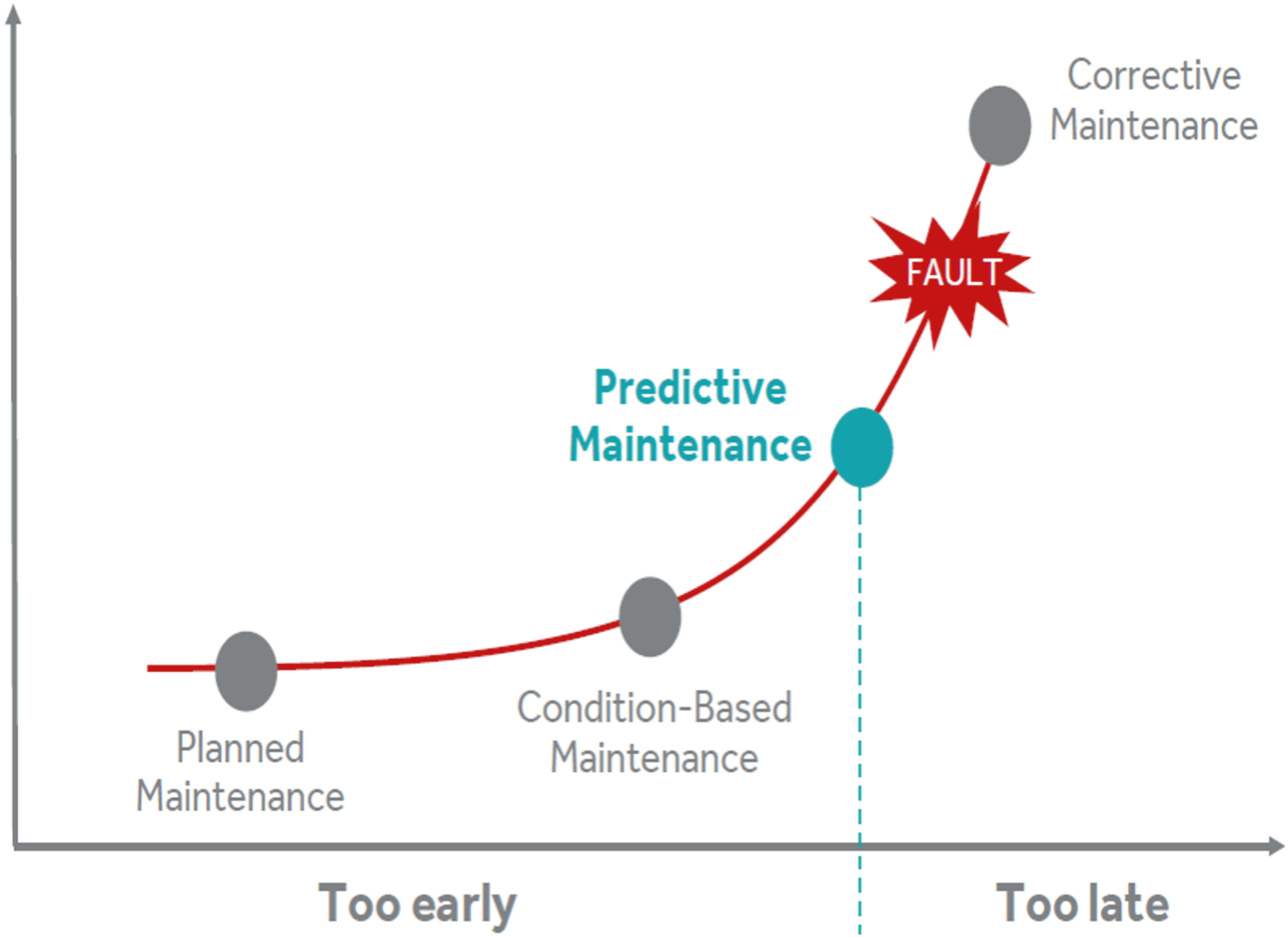


Propblem: not understandable

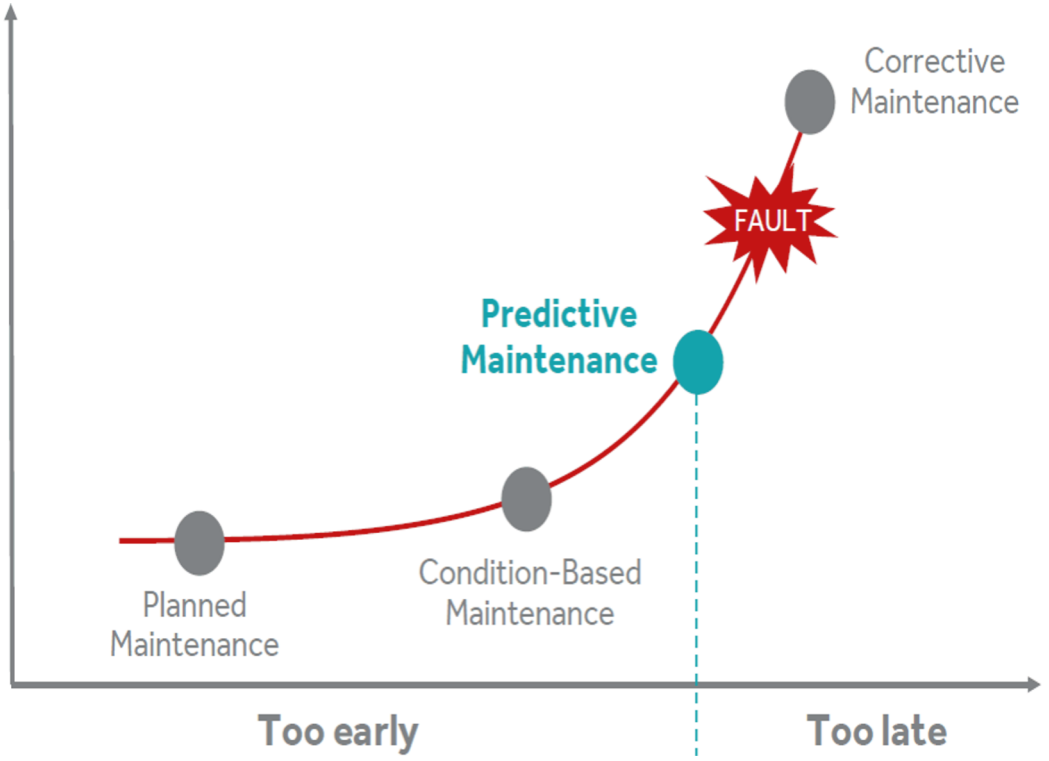
- GIGO: garbage in, garbage out
- No patterns in NN (eg eyes, ears)
- Acceptable? Responsible? Liable?

PART 2: Predictive maintenance

Predictive maintenance



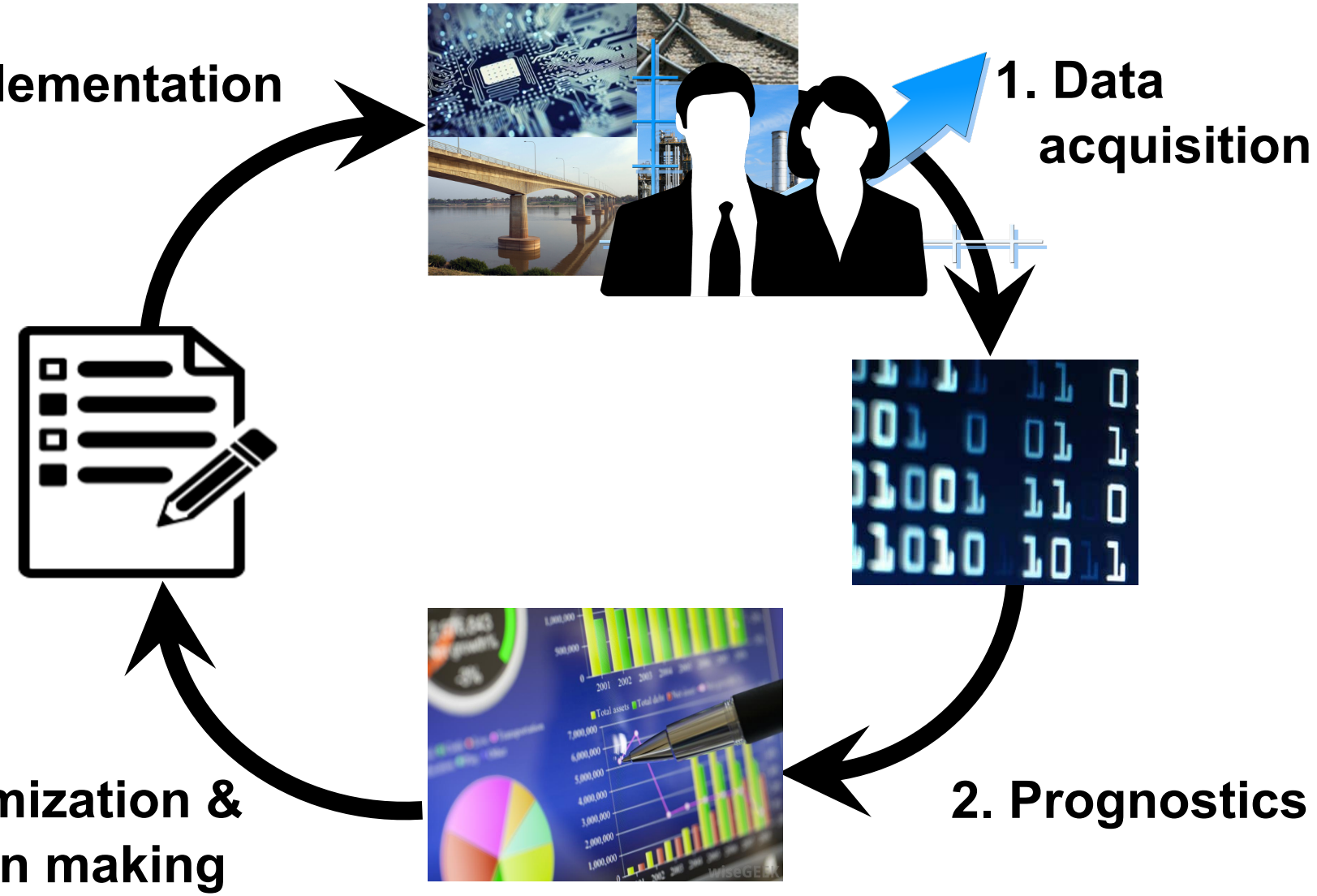
Predictive maintenance



McKenzie

- PdM is most promising IoT application
- 63 billion / year cost savings
- 50% downtime reductions

workflow: predictive maintenance



Discussion

I am curious to hear

- Which data could help your organization?
- What would you like to predict?