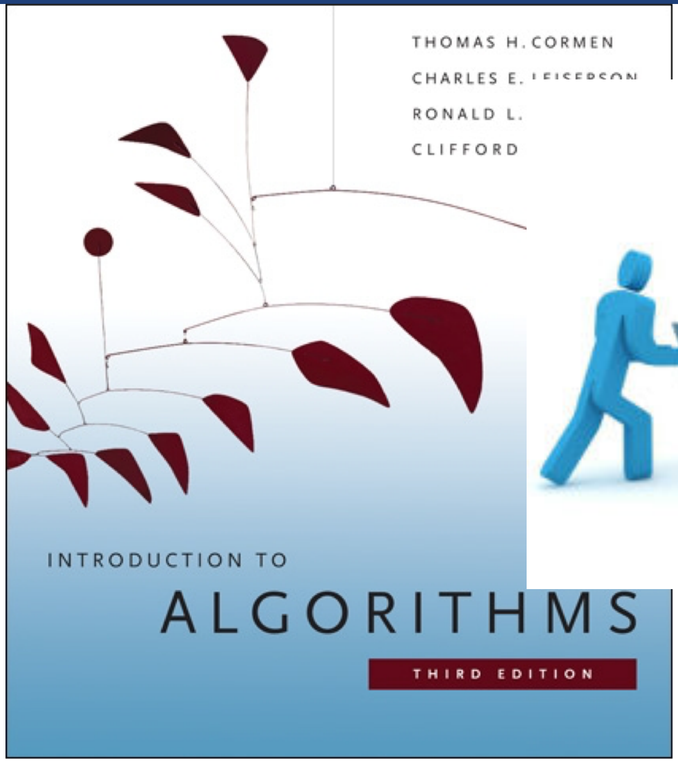
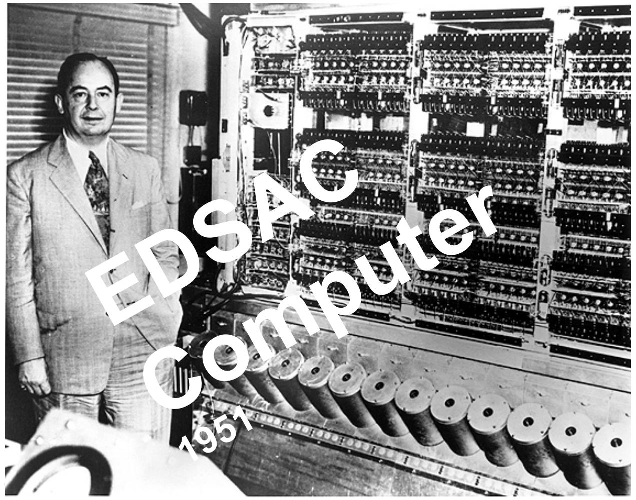
Advanced Machine Learning, Lecture 2019
Philosophical Motivation

Joachim M. Buhmann

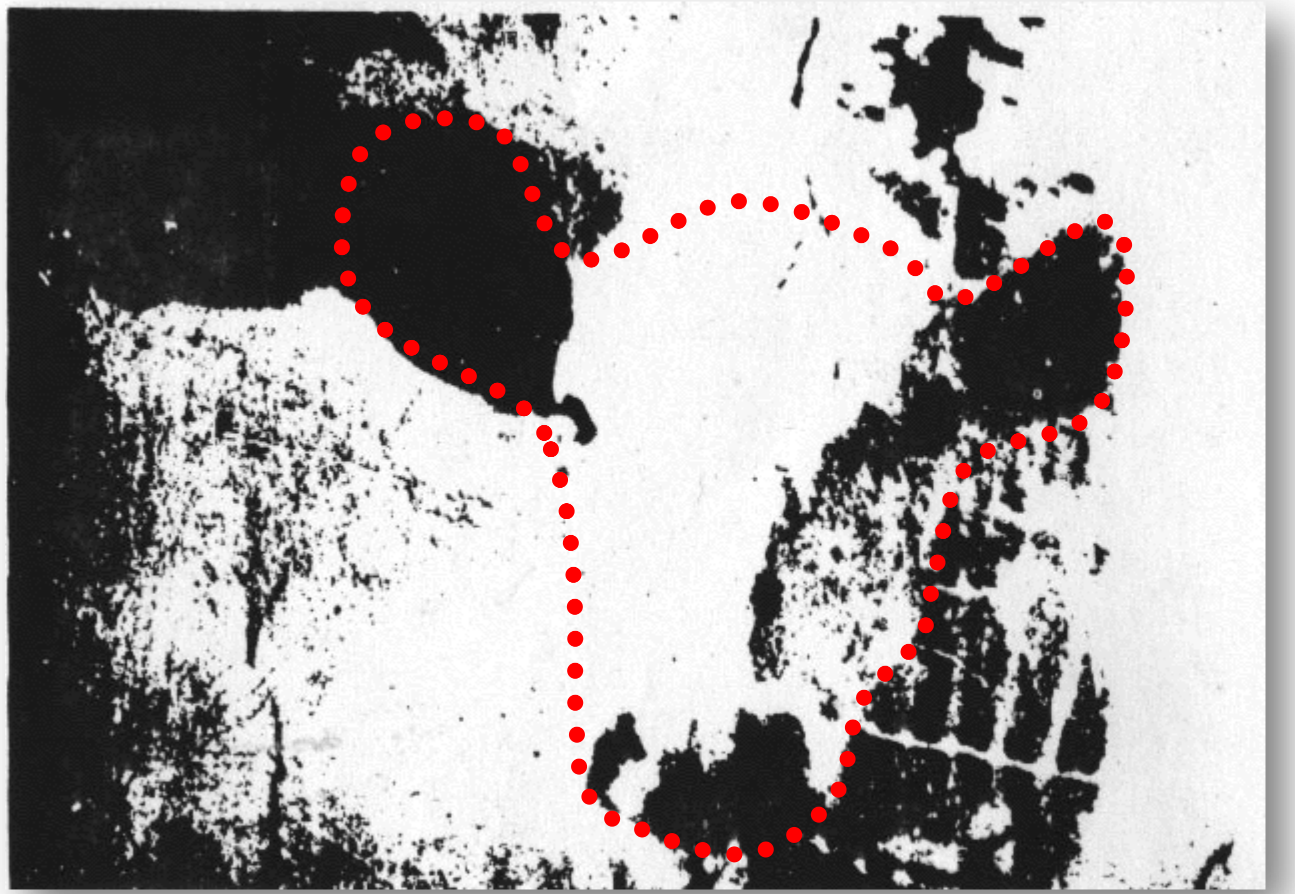
Institute for Machine Learning, D-INFK, ETH Zurich

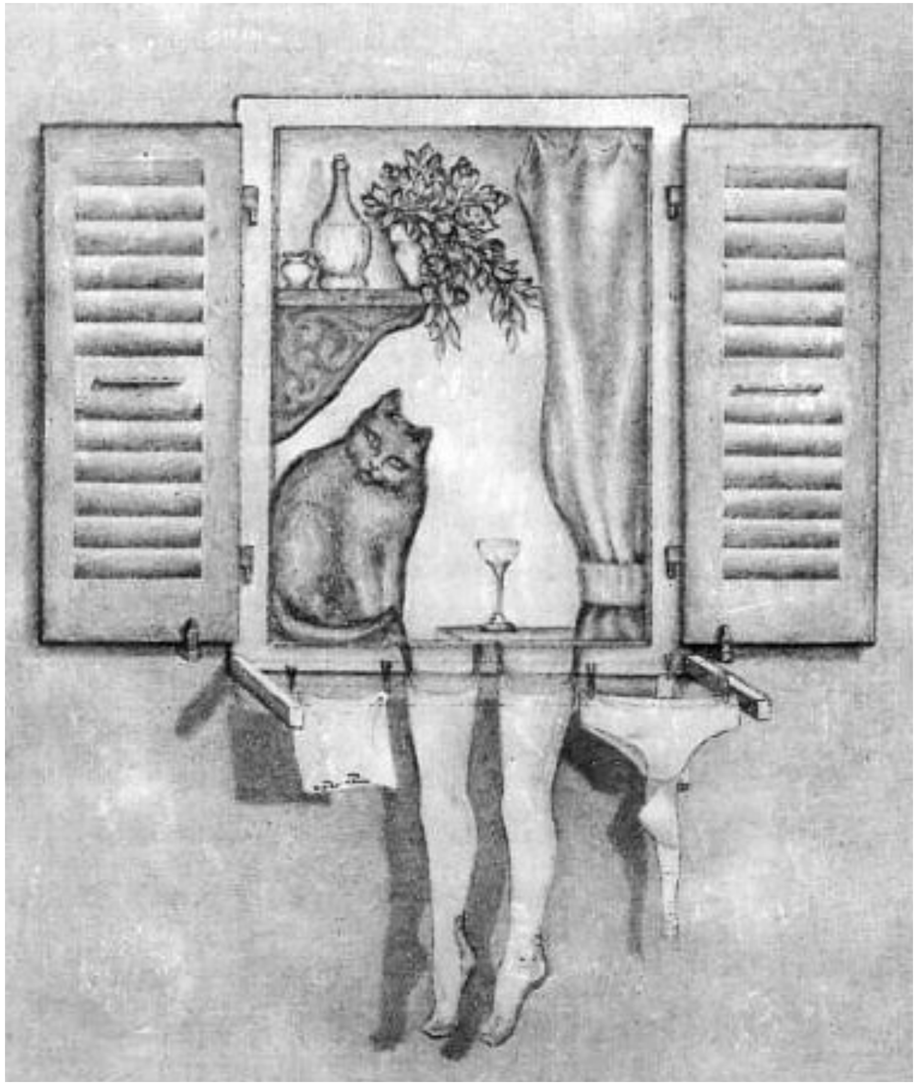
Our world, in which we live!





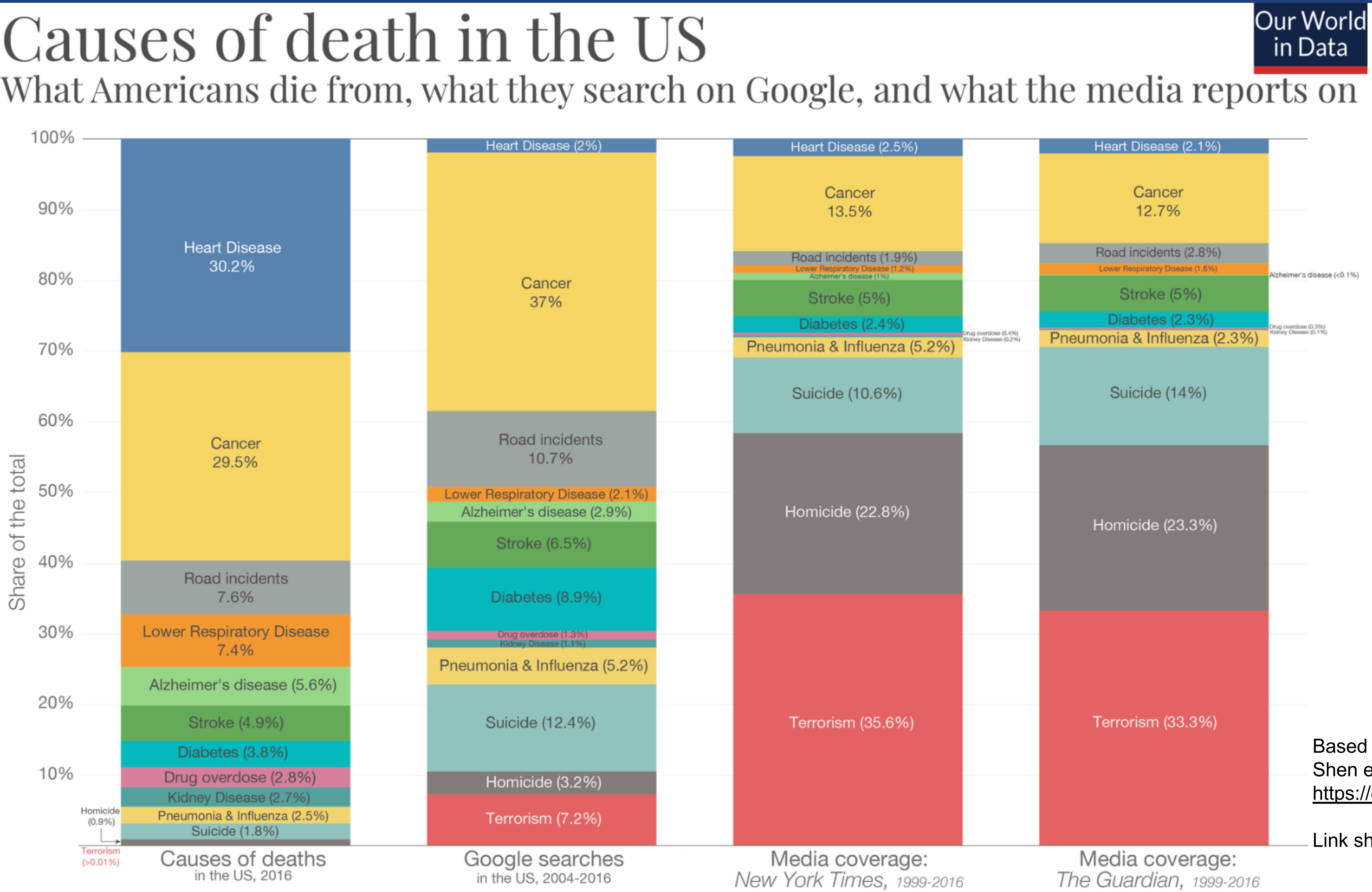
Seeing patterns in data (vision) is difficult!





Causes of death in the US

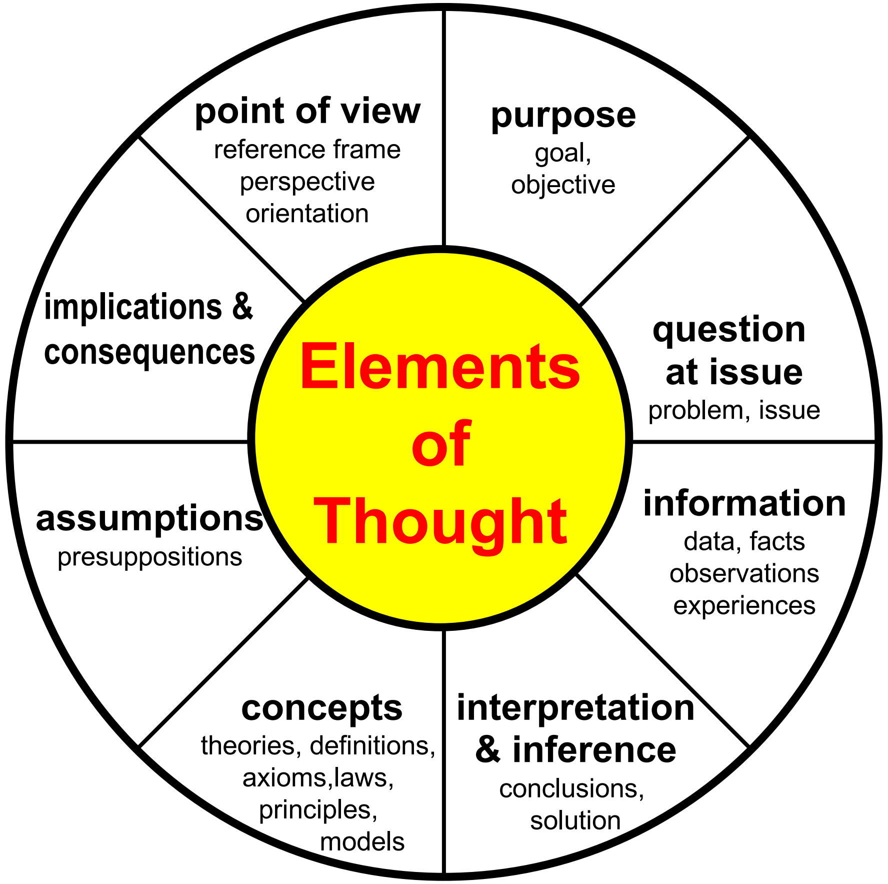
What Americans die from, what they search on Google, and what the media reports on

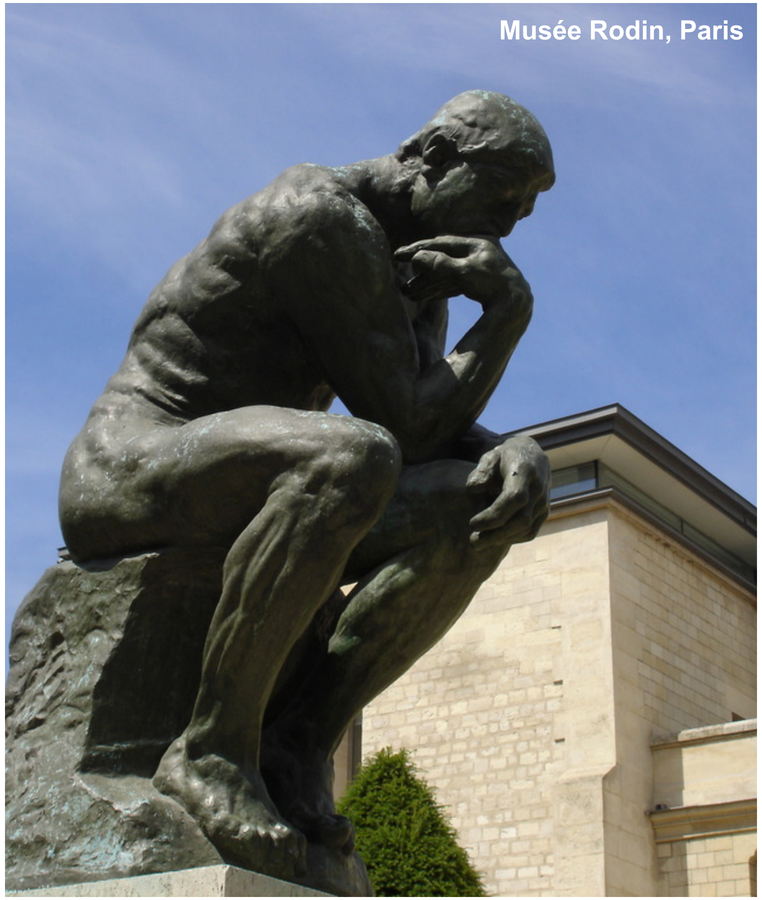


Based on data from
Shen et al. (2018) – Death: reality vs. reported.
https://owenshen24.github.io/charting-death

Link shared by **Alessandro Curioni**, IBM Research

Elements of thinking

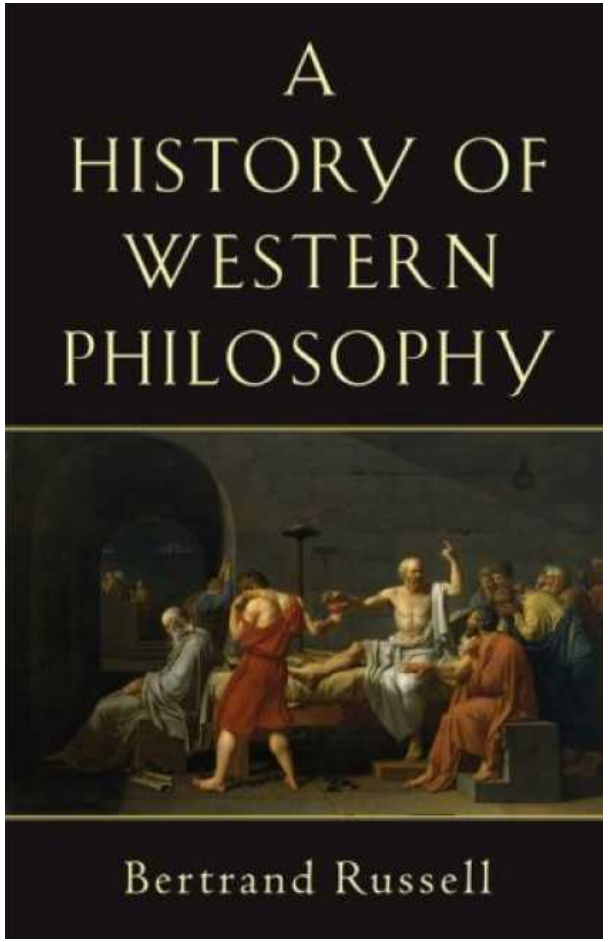




What is learning and understanding?

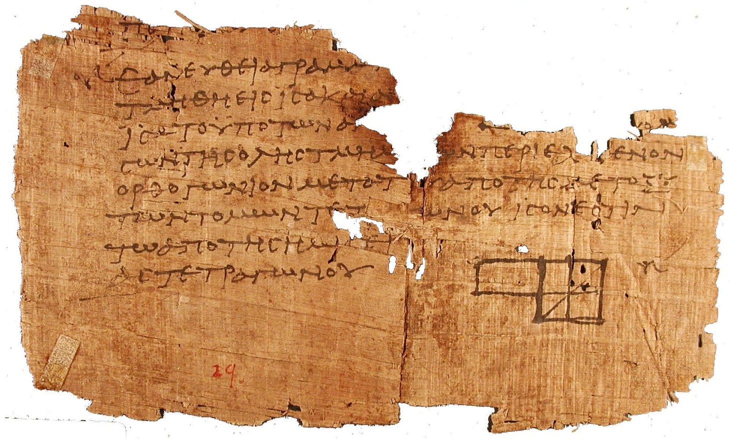
- **Thought processes**
  - **deduction**: draw conclusions from axioms, i.e., derive theorems from obviously true statements
  - **induction**: abstraction, i.e., conclude general laws from special cases
  - … intuition, creativity, causality …
- **Understanding** and **explanation**

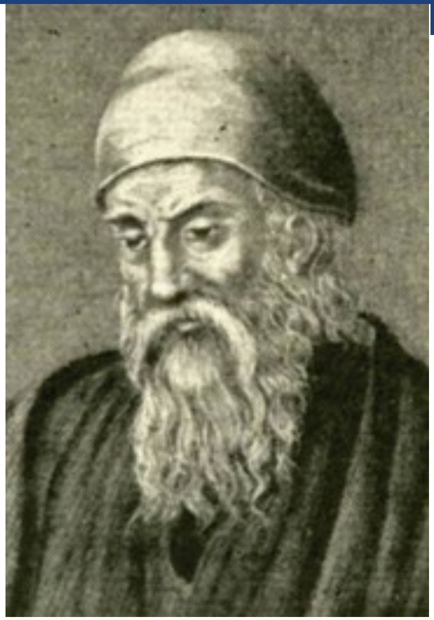
  Why and how does something work?
- **Planning**
  - **conterfactual reasoning**: What would be, if …?
  - **causal thinking**: Why? What cause?

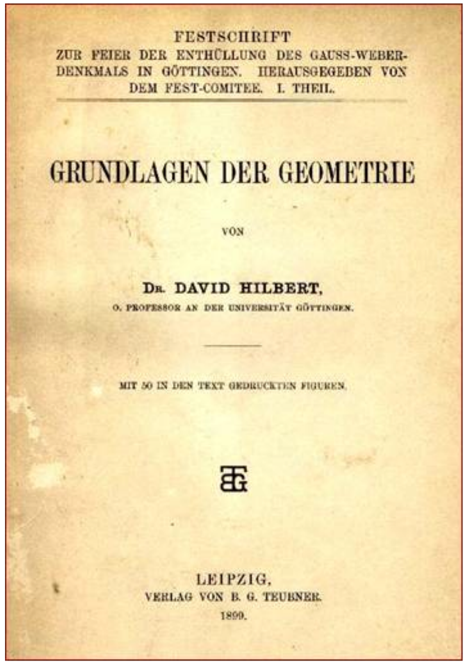


Power of deduction

- **Euclid** (300BC) *Elements*,
  - Axiomatization of geometry with 5 postulates
- **David Hilbert** (1899) *Grundlagen der Geometrie*
  - 20 Axioms are the foundation of Euclid’s geometry!
- **David Mumford** (1937 - …)
  - Algebraic geometry
  - Bayesian reasoning & AI







FESTSCHRIFT
ZUR FEIER DER ENTHÜLLUNG DES GAUSS-WEBER-DENKMALS IN GÖTTINGEN. HERAUSGEGEBEN VON DEM FEST-COMITEE. I. THEIL.

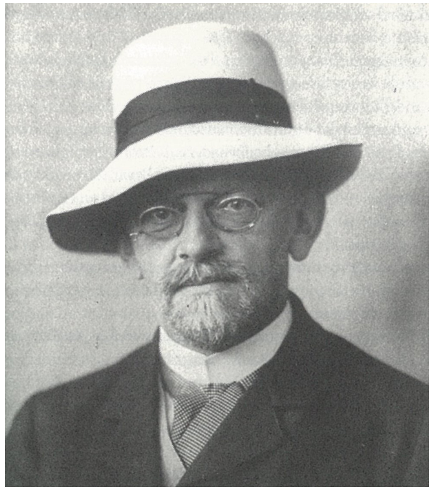
GRUNDLAGEN DER GEOMETRIE

VON

DR. DAVID HILBERT,
O. PROFESSOR AN DER UNIVERSITÄT GÖTTINGEN.

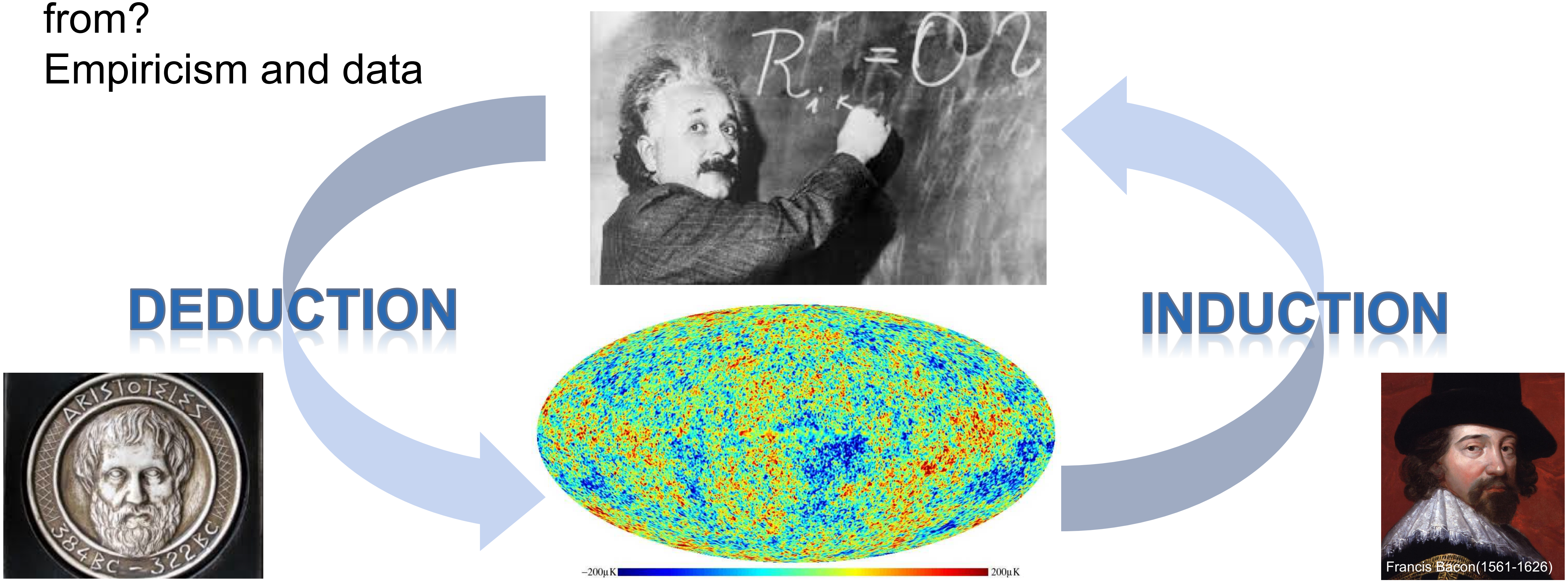
MIT 50 IN DEN TEXT GEDRUCKTEN FIGUREN.

LEIPZIG,
VERLAG VON B. G. TEUBNER.
1899.



Induction – reference to “reality”

Axioms are generally accepted **fundamental principles**! Where do they come from?
Empiricism and data



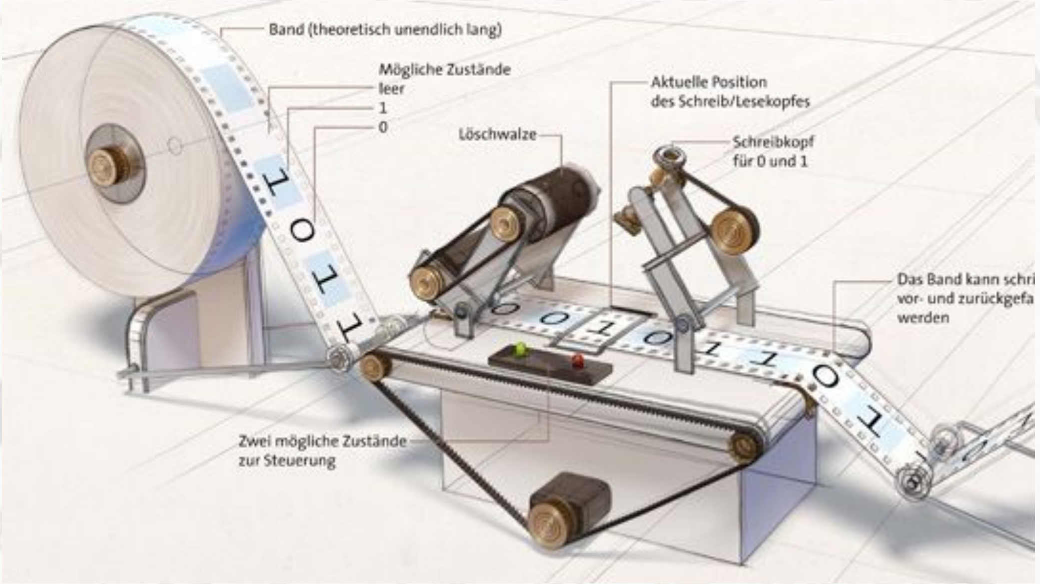
Deduction and induction as computational processes

- **Leibniz**: ***Characteristica Universalis***

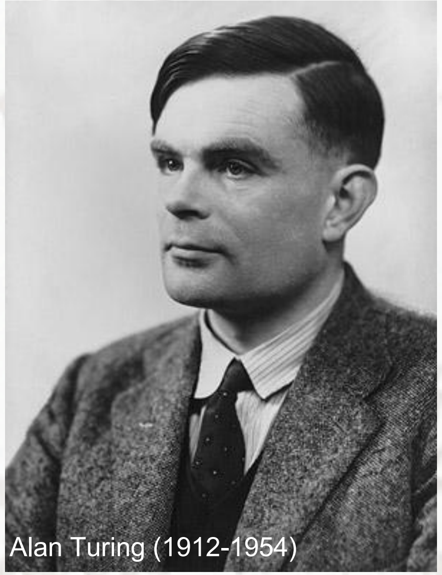
  Thinking is achieved by computing

  - First mechanical calculators
  - Binary system

- T… is … computer

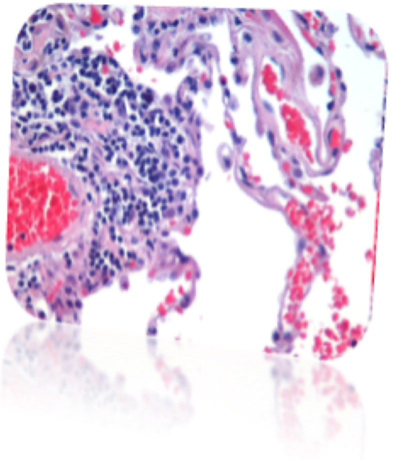


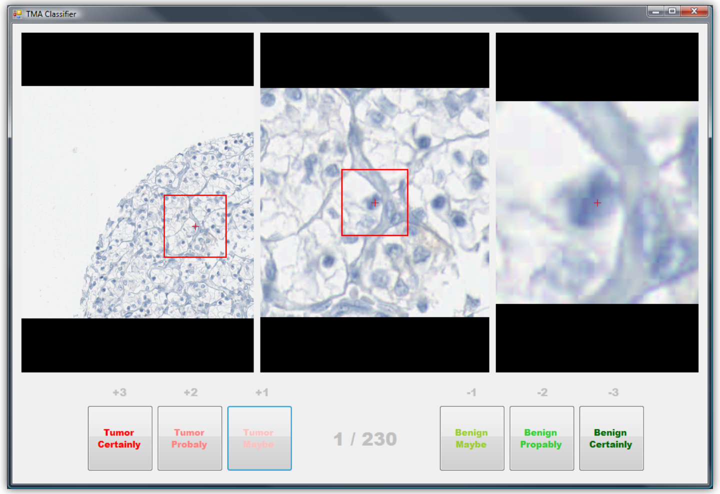

Gottfried Wilhelm Leibniz (1646-1716)

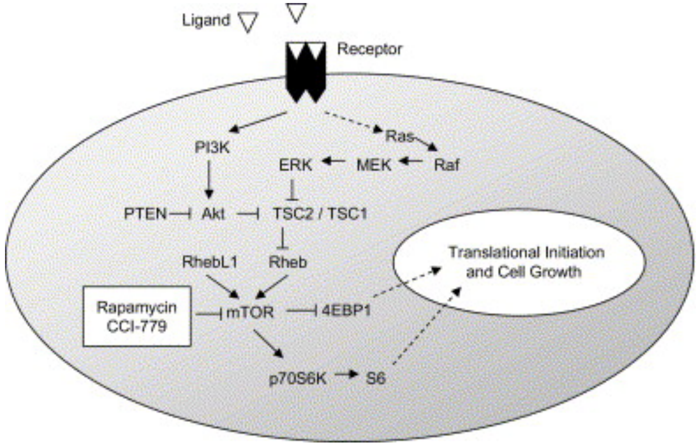

Alan Turing (1912-1954)

IT value generation in personalized medicine

Thomas Fuchs
MSKCC, PAIGE.AI

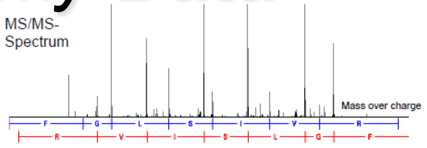


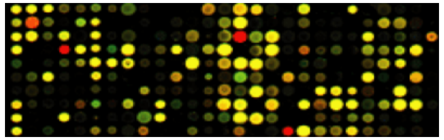


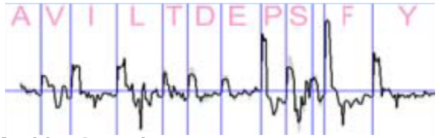


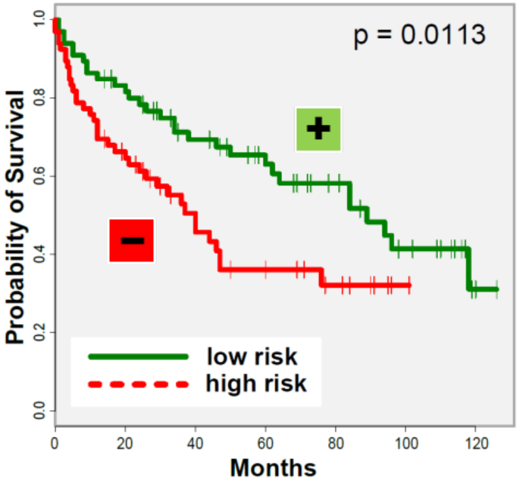
Activation of the mTOR Signaling Pathway in Renal Clear Cell Carcinoma. Robb et al., J Urology 177:346 (2007)

my **Data** → *my* **Information** → *our* **Knowledge** → *my* **Value**



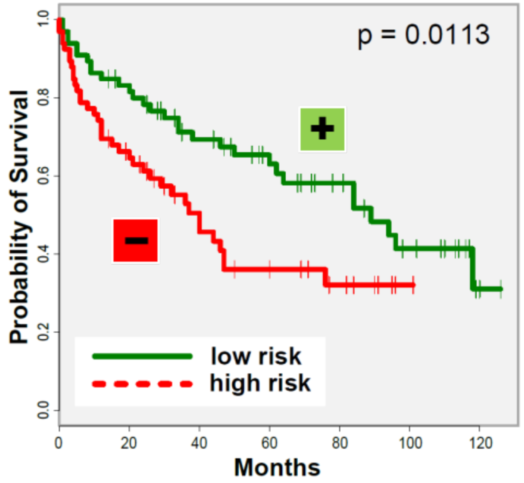




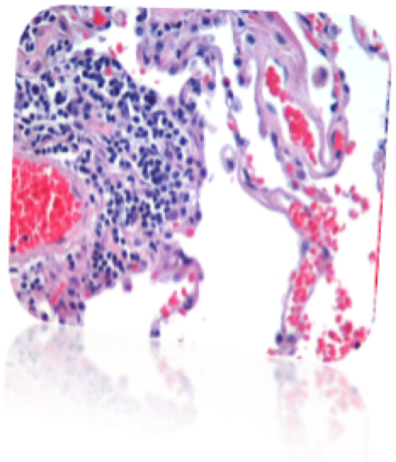


Fundamental data science questions - Which posterior distribution is encoder by algorithm $\mathcal{A}$?

$\mathbb{P}^{\mathcal{A}}\Big($



$\Big|$



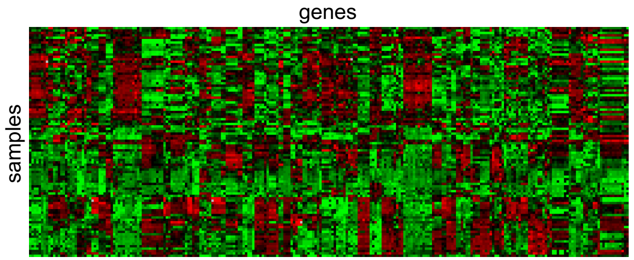
$,$

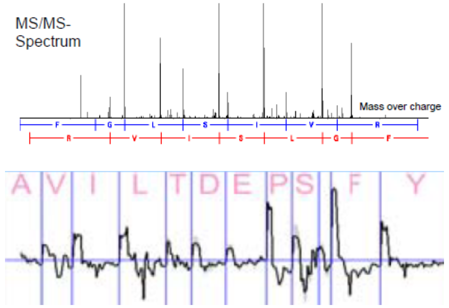
$\Big)$

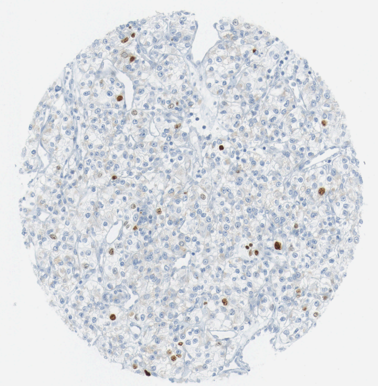
Big Data in der Medizin
Worüber reden wir eigentlich?

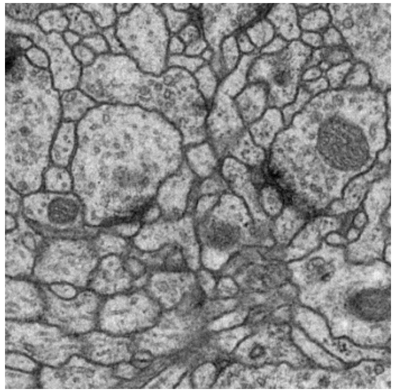
Heterogene Datenquellen

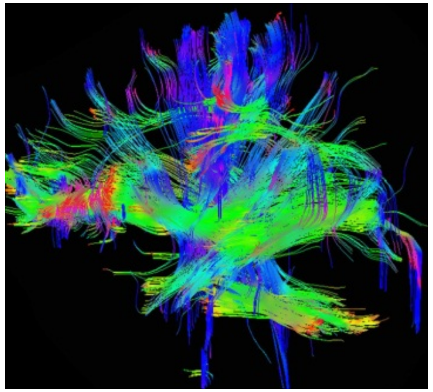
Patientendaten



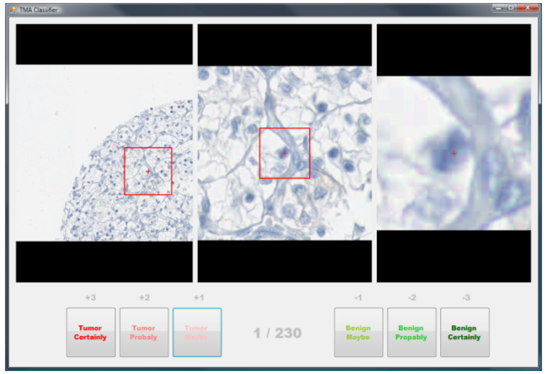


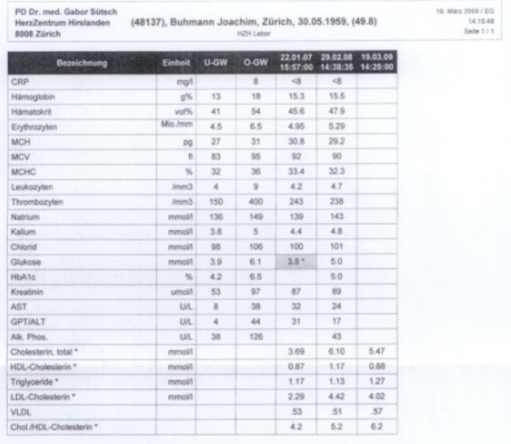






Diagnose-, Prognose- und Therapiedaten der Mediziner





PD Dr. med. Gabor Sütsch
HerzZentrum Hirslanden
8008 Zürich

(48137), Buhmann Joachim, Zürich, 30.05.1959, (49.8)
HZH Labor

19. März 2009 / EG
14:15:48
Seite 1 / 1

Bezeichnung	Einheit	U-GW	O-GW	22.01.07 15:57:00	29.02.08 14:38:35	19.03.09 14:25:00
CRP	mg/l		8	<8	<8	
Hämoglobin	g%	13	18	15.3	15.5	
Hämatokrit	vol%	41	54	45.6	47.9	
Erythrozyten	Mio./mm	4.5	6.5	4.95	5.29	
MCH	pg	27	31	30.8	29.2	
MCV	fl	83	95	92	90	
MCHC	%	32	36	33.4	32.3	
Leukozyten	/mm3	4	9	4.2	4.7	
Thrombozyten	/mm3	150	400	243	238	
Natrium	mmol/l	136	149	139	143	
Kalium	mmol/l	3.8	5	4.4	4.8	
Chlorid	mmol/l	98	106	100	101	
Glukose	mmol/l	3.9	6.1	3.5 *	5.0	
HbA1c	%	4.2	6.5		5.0	
Kreatinin	umol/l	53	97	87	89	
AST	U/L	8	38	32	24	
GPT/ALT	U/L	4	44	31	17	
Alk. Phos.	U/L	38	126		43	
Cholesterin, total *	mmol/l			3.69	6.10	5.47
HDL-Cholesterin *	mmol/l			0.87	1.17	0.88
Triglyceride *	mmol/l			1.17	1.13	1.27
LDL-Cholesterin *	mmol/l			2.29	4.42	4.02
VLDL				.53	.51	.57
Chol./HDL-Cholesterin *				4.2	5.2	6.2

„Selbstvermessung"

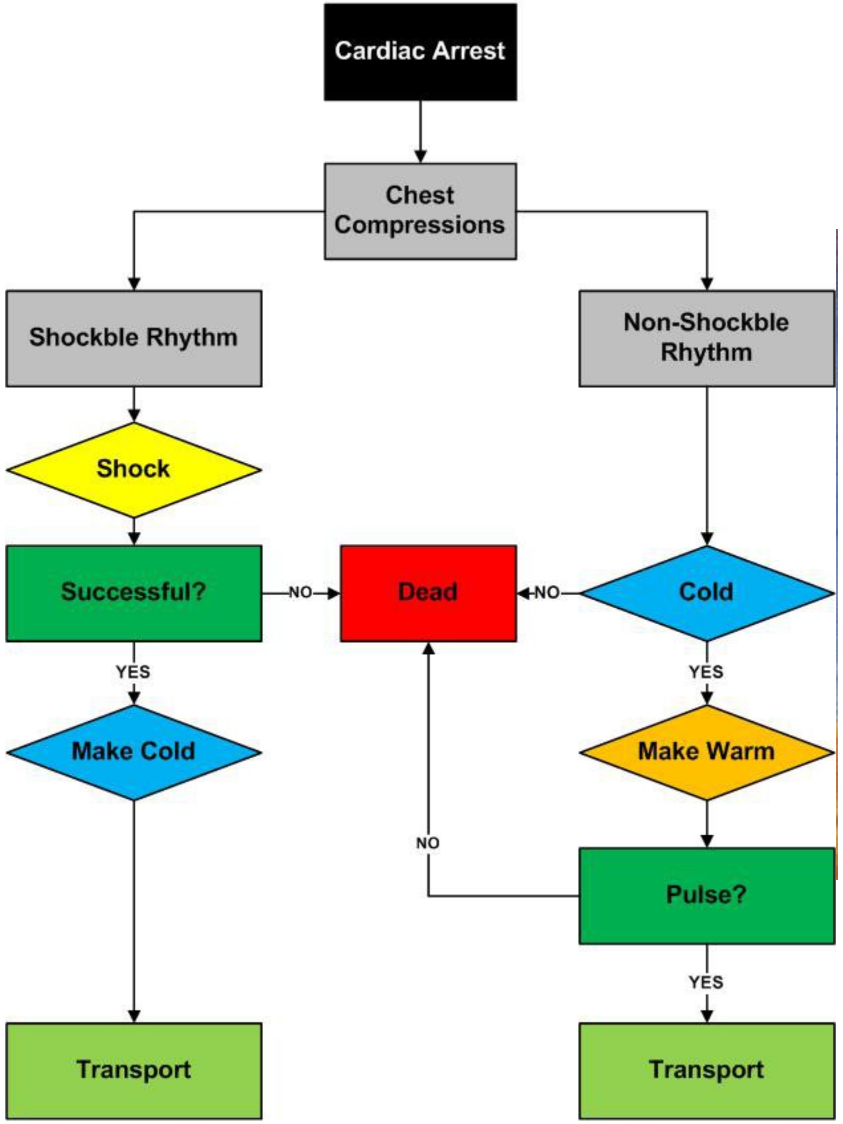
The Algorithm: Idiom of Modern Science

(Bernard Chazelle)

- Informally, an **algorithm** is any well-defined computational procedure, that takes some value as **input** and produces some value as **output**. (CLRS)
- **Analysis of algorithms**
  - ✓ **Runtime**
  - ✓ **Memory consumption**
  - ✗ **Robustness**
  - ✗ **Generalization**

➢ **Learning algorithms „explore" a complex stochastic reality!**

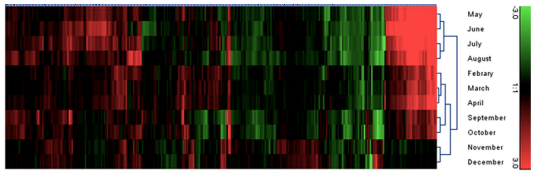
2015 ACLS CARDIAC ARREST ALGORITHM

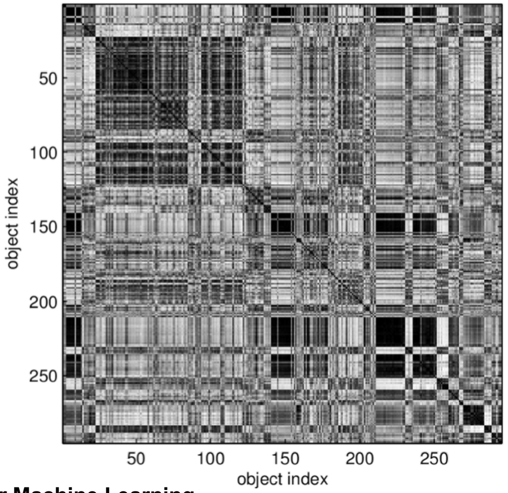


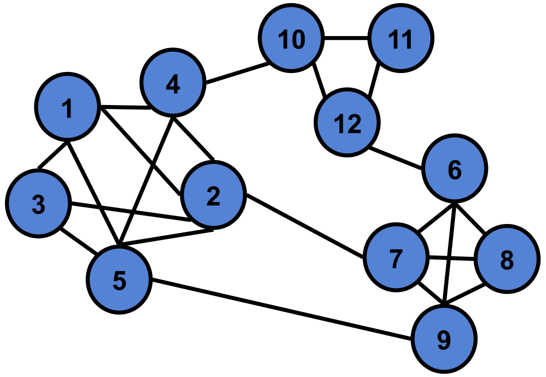
Algorithmics for Data Science – what is the problem?

- **Random inputs imply random outputs**

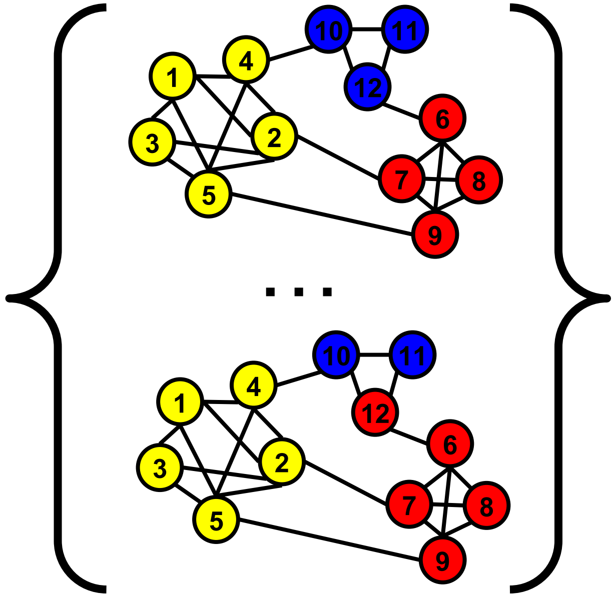
$$\underbrace{\text{input } \mathbf{X}}_{\text{given}} \sim P(\mathbf{X}) \Longrightarrow \underbrace{\mathcal{A}}_{\text{algorithm}} \Longrightarrow \underbrace{\text{output } c \sim P(c|\mathbf{X})}_{\text{design}}$$



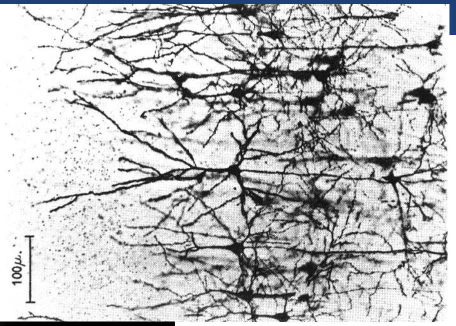




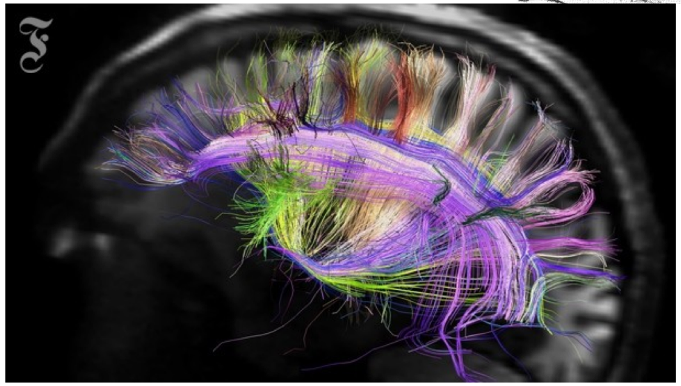
$\mathcal{A}$?



Learning machines master algorithmic induction and «imitate» humans

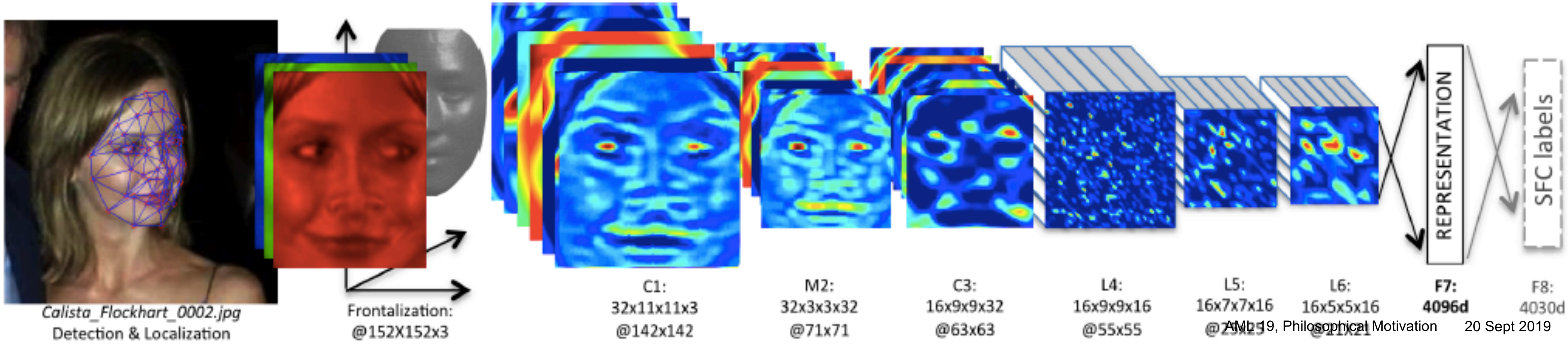


- **Biological neural networks** are adaptive and can learn.
- **Artificial neural networks** mimic these learning capabilities.



Neural networks visualized by brain scans. © VAN WEDEEN

- **DeepFace** network of FaceBook



"Deep Network" Halluzinationen

(Courtesy of **Sebastian Nowozin, 2016**)

Image interpolation with neural networks

What is missing?

The Scientific Method

Step 1: **Ask questions**

$\mathcal{A}$?

Step 2:
Propose hypotheses

$\mathcal{A}$?

Step 3:
Conduct experiment

$\mathcal{A}$?

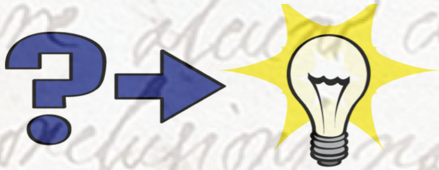
Step 4: **Analyze results**

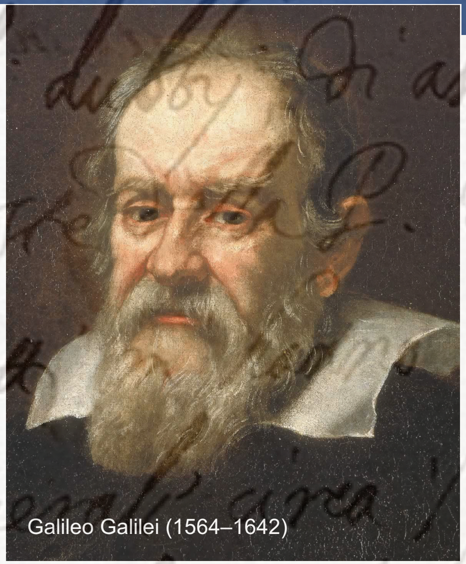
$\mathcal{A}$?

Step 5: Draw conclusions –
postulate a theory

$\mathcal{A}$?

$\mathcal{A}$?





Galileo Galilei (1564–1642)

Core question for computer science:
How can we validate (data science) algorithms?

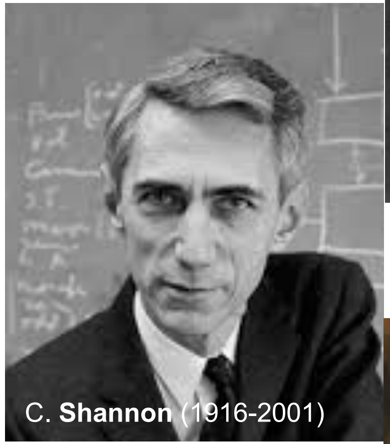
I. **Algorithms with random variables as input compute random variables as output!**
   How can we prove correctness of such algorithms?

II. **Algorithms** have to compute **typical solutions**!
   What does this mean for algorithm design?

III. **When do algorithms generalize** **over noise/model mismatch?**

IV. **How can algorithms autonomously improve performance?**

A. **Kolmogorov** (1903-1987)



C. **Shannon** (1916-2001)

V. **Vapnik** (1936 -)