Through the Philosopher's Glass

Scattered Reflections on the Philosophical and Socio-ethical Aspects of Machine Learning

Marcello Pelillo
University of Venice, Italy

Winter School on Quantitative Systems Biology: Learning and Artificial Intelligence
Trieste, Italy, November 23, 2017

The European Centre for Living Technology

http://www.ecltech.org

Canaletto, Grand Canal from Santa Maria della Carità (1726)

Mission

Established in 2004, ECLT is an *international* and *interdisciplinary* research Centre dedicated to the creation of technologies and methodologies which embody the essential properties of **living systems**, such as:

- Self-organization
- Evolution
- Adaptability
- Learning
- Perception & language

Members of ECLT

biology
computer science
statistics
physics
engineering
chemistry
mathematics
economics
philosophy
sociology
complex systems

ECLT is a consortium of Universities, Laboratories, Centres

Some Past Projects

PACE - Programmable Artificial Cell Evolution (2004-2009)
EU 6th Framework Program Cooperation ICT

ECCell – Electronic Chemical Cell (2008-2011)
EU 7th Framework Program Cooperation ICT

ASSYST - Action for the Science of complex SYstems for Socially intelligent (2009-2012)
EU 7th Framework Program Cooperation ICT

COBRA - Coordination of Biological & Chemical IT Research Activities (2010 -2014)
EU 7th Framework Program Cooperation

GSDP - Global Systems Dynamics and Policy (2010-2014)
EU 7th Framework Program Cooperation ICT

INSITE - The Innovation Society, Sustainability, and ICT (2011 -2014)
EU 7th Framework Program Cooperation ICT

iNSPiRe - Development of Systemic Packages for Deep Energy Renovation of Residential and Tertiary Buildings including Envelope and Systems (2012-2016)
EU 7th Framework Program Cooperation

MATCHIT - Matrix for Chemical IT (2010-2013)
EU 7th Framework Program Cooperation ICT

MD - Emergence by Design (2011-2014)
EU 7th Framework Program Cooperation

MICREAGENTS - Microscale Chemically Reactive Electronic Agents (2012-2015)
EU 7th Framework Program Cooperation ICT

Current Projects

Green Growth and Win-win Strategies For Sustainable Climate Action

Consortium

- Global Climate Forum
- The Institute of Environmental Sciences and Technology
- Autonomous University of Barcelona
- E3-Modelling
- Environmental Change Institute, Oxford University
- Ecole d'Economie de Paris
- University College London
- The Ground_Up Project
- Deltares
- Institute for Advanced Sustainability Studies
- Global Green Growth Institute
- Jill Jaeger
- European Centre for Living Technology
- Institute of Environmental Sciences at Boğaziçi University
- Center for Remote Sensing and Ocean Sciences, Udayana University
- University of Cape Town
- 2° Investing Initiative

Current Projects

New Pathways for Sustainable Urban Development in China’s Medium-sized Cities

Consortium

Centre National de la Recherche Scientifique
Hangzhou Normal University
Institut d’Etudes Politiques d’Aix-en-Provence
European Centre for Living technology
Spatial Foresight GmbH

Current Projects

Welcome to the AI4EU initiative

The AI4EU proposal adressing [ICT-26 2018 H2020 call] has successfully passed the evalution process.

The project should start early this autumn

https://ai4eu.org

Current Projects

SAMSUNG

Hume-Nash Machines: Context-aware Models of Learning and Recognition

gsk GlaxoSmithKline

Statistical Procedures for Lead Optimization in Drug Discovery Processes

Events


SINS 2016
Social Impact through Network Science
IEEE
SMC
Systems, Man, and Cybernetics Society
eclt
European Centre for Living Technology
The Human Use of Machine Learning:
22 December 2016
Christmas Lecture
Luc Steels
"Will Artificial Intelligence Rule the World?"
ECLT Christmas Lecture, 22 December 2016

ICCV 2017

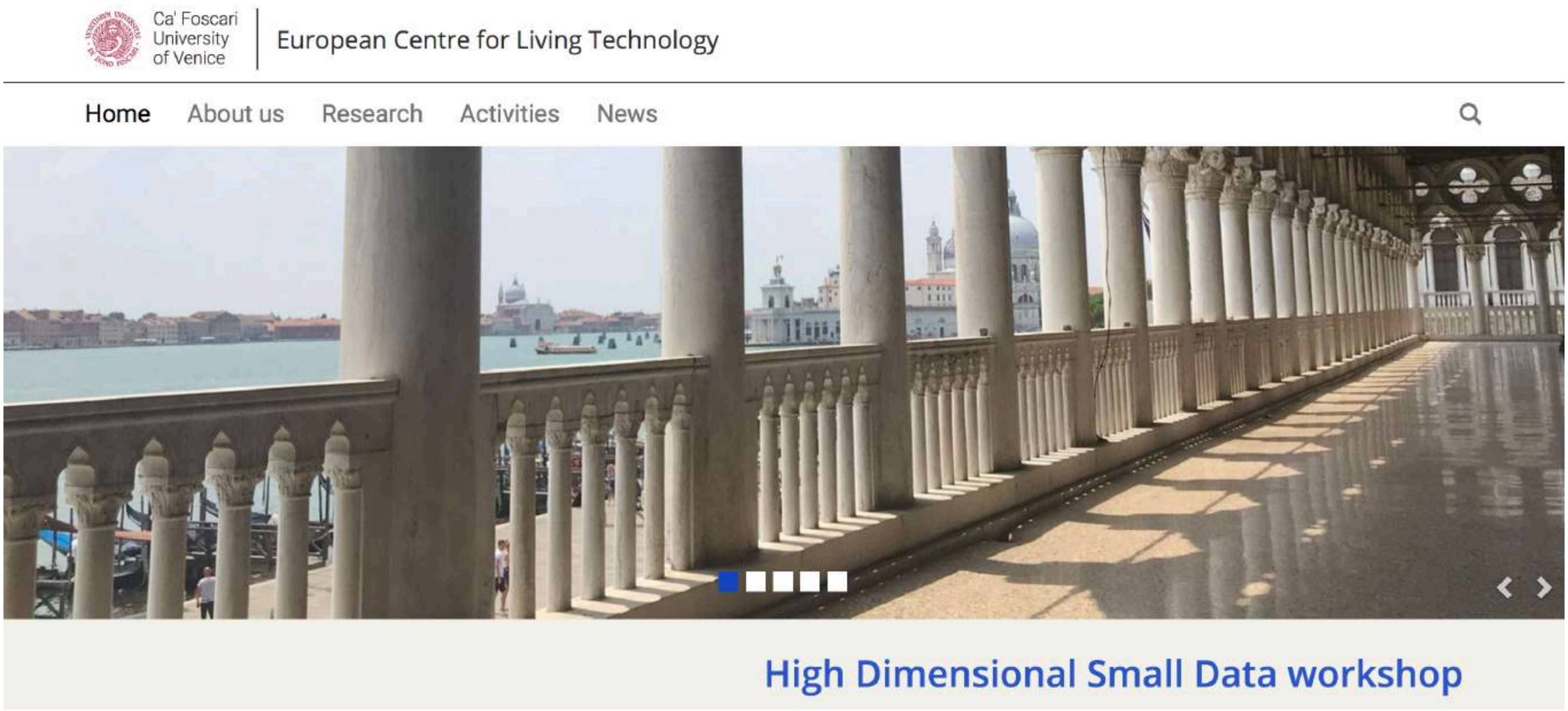
http://iccv2017.thecvf.com/

OVERALL MEETING SPONSORS

CVF IEEE computer society

If you want to know more …

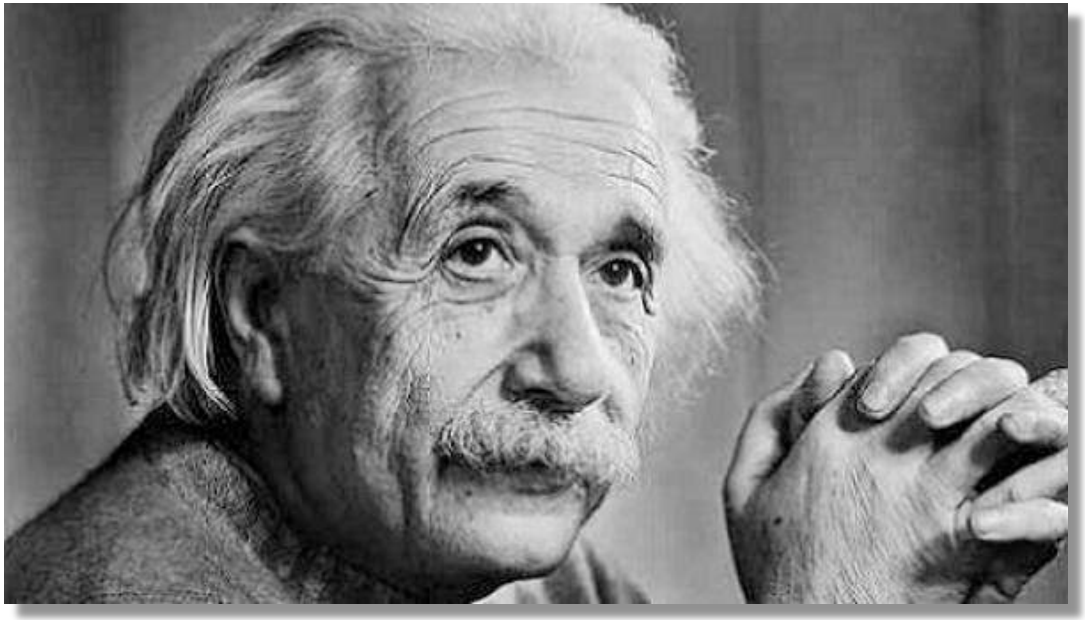
http://www.ecltech.org



Two attitudes towards philosophy

«Science without epistemology is,
insofar as it is thinkable at all, primitive and muddled.»

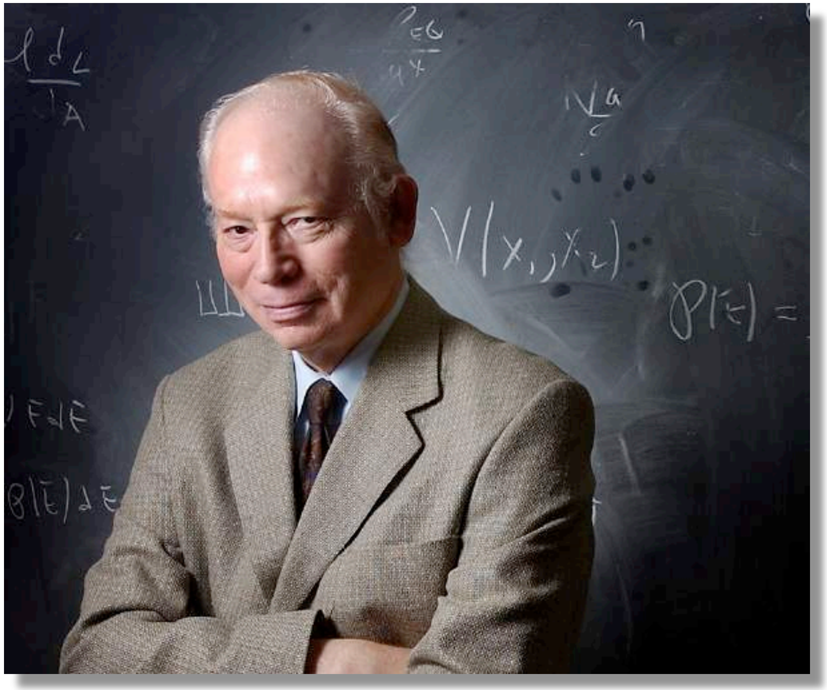
Albert Einstein (1949)



Two attitudes towards philosophy

«We should not expect [philosophy] to provide today's scientists with any useful guidance about how to go about their work or about what they are likely to find.»

Steven Weinberg
Dreams of a Final Theory (1993)



Why philosophy?

«It is not just that the philosophy of science is safe for scientists. **A little of it may even do you good.** Like spending time in another culture, the pursuit of the philosophy of science, and of science studies generally, **helps to reveal contingencies in scientific practices that may look like necessities** from within the practices themselves.»

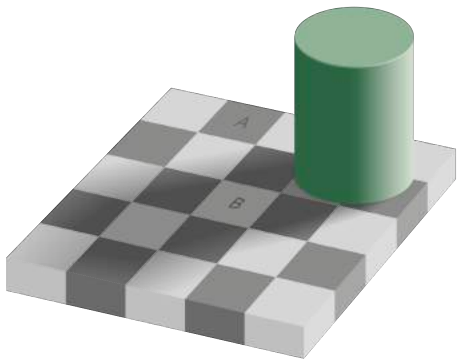
Peter Lipton
The truth about science (2005)

A personal view

«Machine learning is the continuation of epistemology by other means.»

Liberally adapted from Carl von Clausewitz

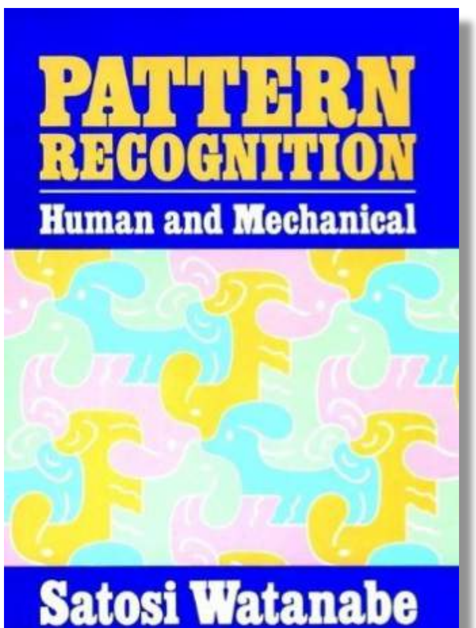
Our essentialist assumption



The heirs of Aristotle?

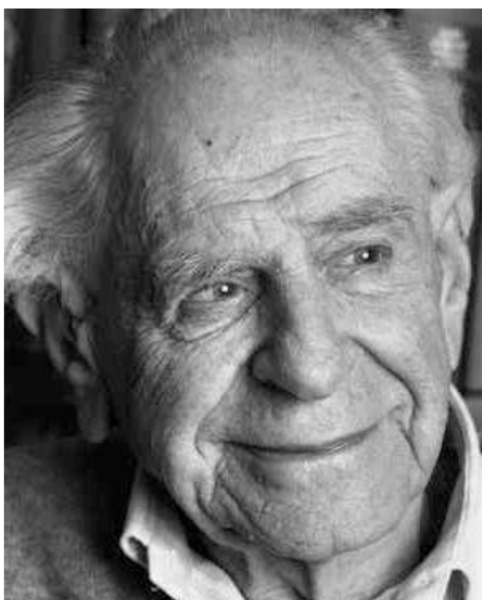
«Whether we like it or not, under all works of pattern recognition lies tacitly the Aristotelian view that the world consists of a discrete number of self-identical objects provided with, other than fleeting accidental properties, a number of fixed or very slowly changing attributes. Some of these attributes, which may be called "features," determine the class to which the object belongs.»

Satosi Watanabe
Pattern Recognition: Human and Mechanical (1985)



Essentialism and its discontents

«The development of thought since Aristotle could be summed up by saying that every discipline, as long as it used the Aristotelian method of definition, has remained arrested in a state of empty verbiage and barren scholasticism, and that **the degree to which the various sciences have been able to make any progress depended on the degree to which they have been able to get rid of this essentialist method.**»



Karl Popper
The Open Society and Its Enemies (1945)

Essentialism under attack

During the XIX and the XX centuries, the *essentialist* position was subject to a massive assault from several quarters and it became increasingly regarded as an impediment to scientific progress.

Strikingly enough, this conclusion was arrived at independently in various different disciplines:

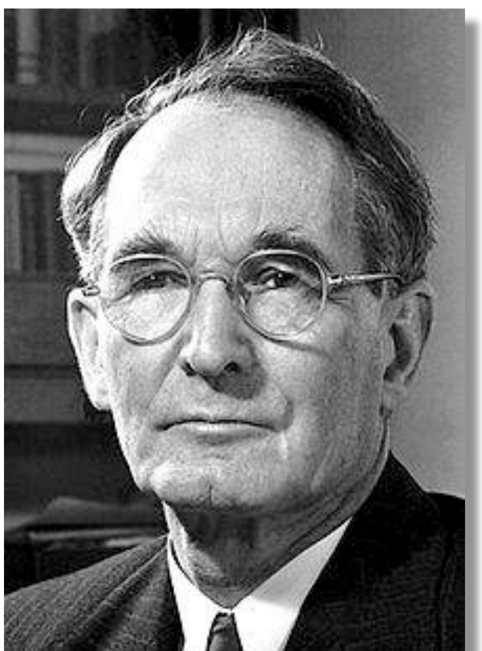
- ✓ Physics
- ✓ Biology
- ✓ Psychology
- ✓ Mathematics

not to mention Philosophy …

Definitions in physics

«**What do we mean by the length of an object?**
[...]
To find the length of an object, we have to perform certain physical operations. The concept of length is therefore fixed when the operations by which length is measured are fixed
[...]
In general, we mean by any concept nothing more than a set of operations; **the concept is synonymous with the corresponding set of operations.**»

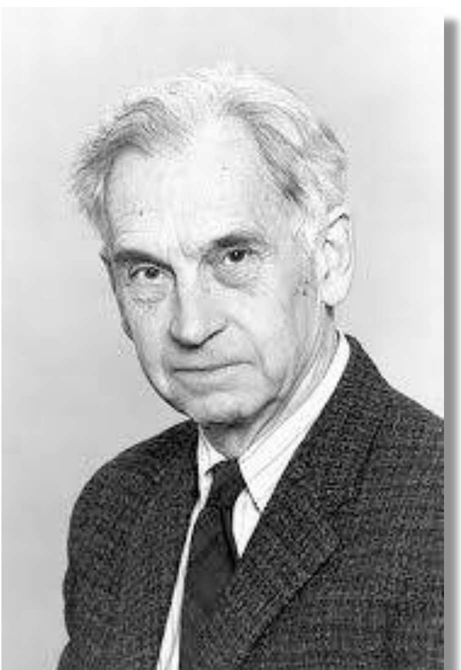
Percy W. Bridgman
The Logic of Modern Physics (1927)



Can we be essentialist after Darwin?

«Essentialism [...] dominated the thinking of the western world to a degree that is still not yet fully appreciated by the historians of ideas.
[...]
It took more than two thousand years for biology, under the influence of Darwin, to escape the paralyzing grip of essentialism.»

Ernst Mayr
The Growth of Biological Thought (1982)



Against "classical" categories

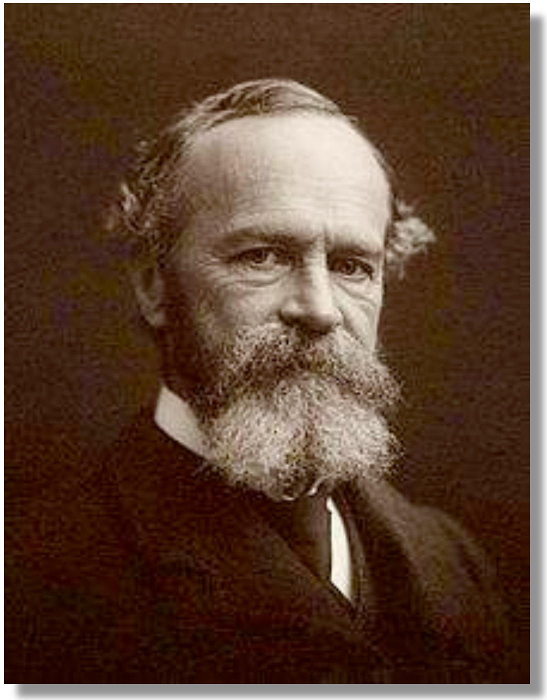
«Categorization is a central issue. The traditional view is tied to the classical theory that categories are defined in terms of common properties of their members. **But a wealth of new data on categorization appears to contradict the traditional view of categories**. In its place there is a new view of categories, what Eleanor Rosch has termed *the theory of prototypes and basic-level categories*.»

George Lakoff
Women, Fire, and Dangerous Things (1987)

"Signal" *vs.* "noise"

«*There is no property* ABSOLUTELY *essential to any one thing.* The same property which figures as the essence of a thing on one occasion becomes a very inessential feature upon another.»

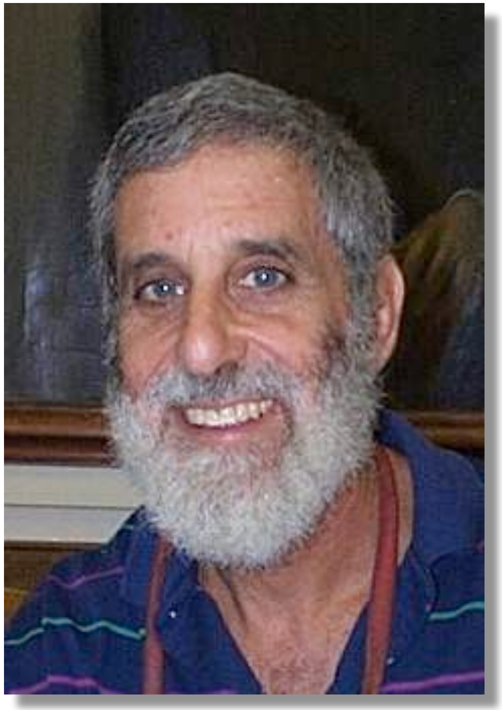
William James
The Principles of Psychology (1890)



What is the subject-matter of math?

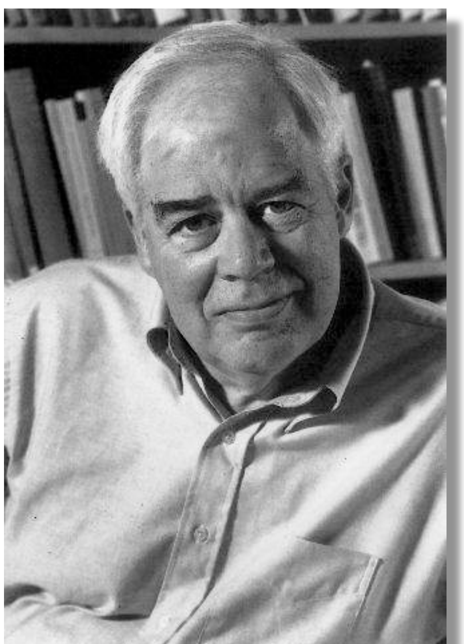
«In mathematics the primary subject-matter is not the individual mathematical objects but rather the structures in which they are arranged.»

Michael D. Resnik
Mathematics as a Science of Patterns (1997)



Radical anti-essentialism

«We antiessentialists would like to convince you that it [...] does not pay to be essentialist about tables, stars, electrons, human beings, academic disciplines, social institutions, or anything else. We suggest that you think of all such objects as resembling numbers in the following respect: **there is nothing to be known about them except an initially large, and forever expandable, web of relations to other objects**.



There are, so to speak, relations all the way down, all the way up, and all the way out in every direction: you never reach something which is not just one more nexus of relations.»

Richard Rorty
A World Without Substances or Essences (1994)

Two consequences of the essentialist assumption in PR/ML

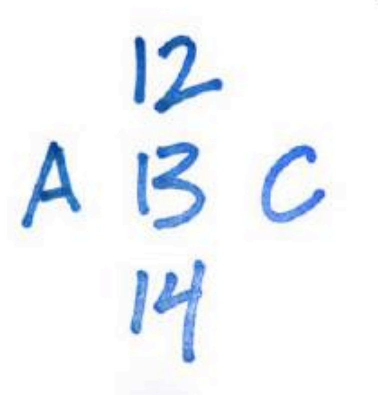
Our essentialist attitude has had two major consequences which greatly contributed to shape the ML/PR fields in the past few decades.

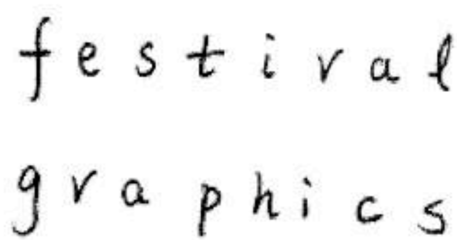
- ✓ it has led the community to focus mainly on **feature-vector representations**, where, each object is described in terms of a vector of numerical attributes and is therefore mapped to a point in a Euclidean (geometric) vector space

- ✓ it has led researchers to maintain a **reductionist position**, whereby objects are seen in isolation and which therefore tends to overlook the role of contextual, or relational, information

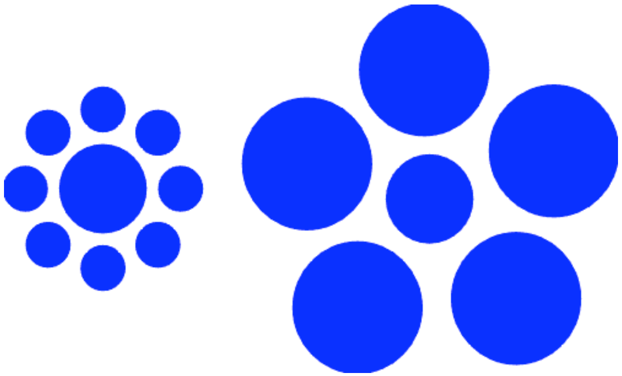
Context helps …

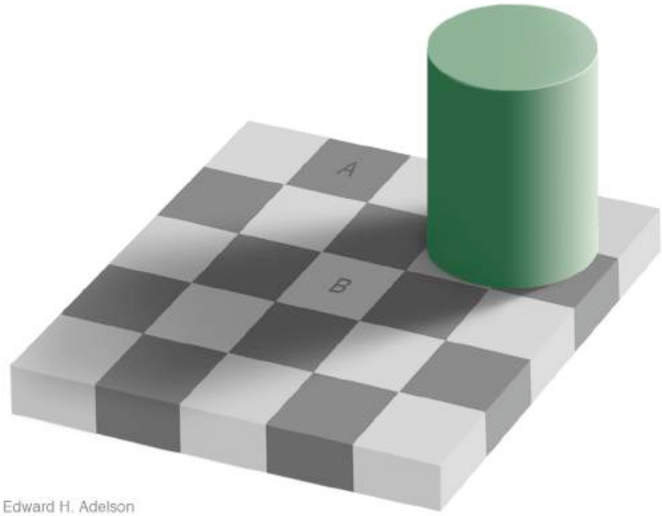
c	→	cat
	→	circus
i	→	sin
	→	fine
e	→	red
	→	read

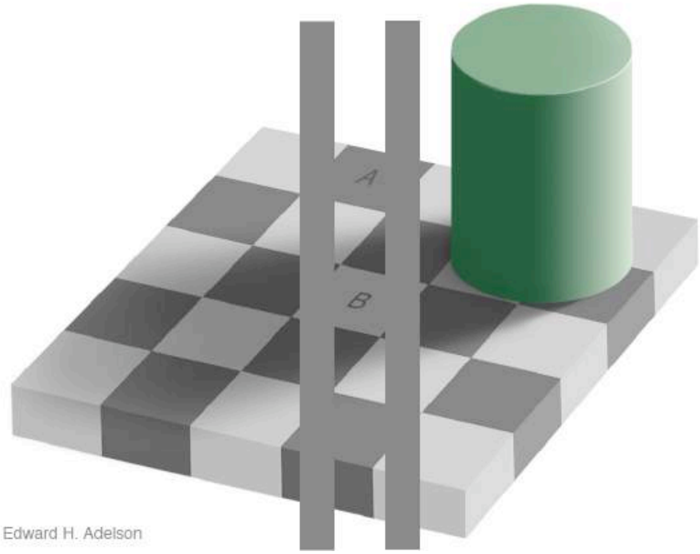




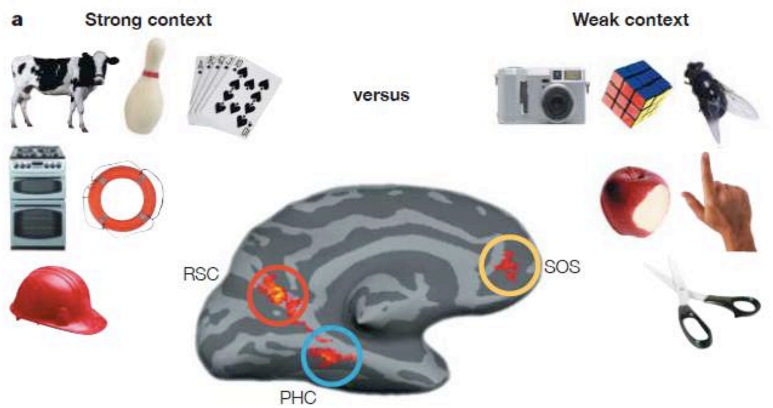
… but can also deceive







Context and the brain



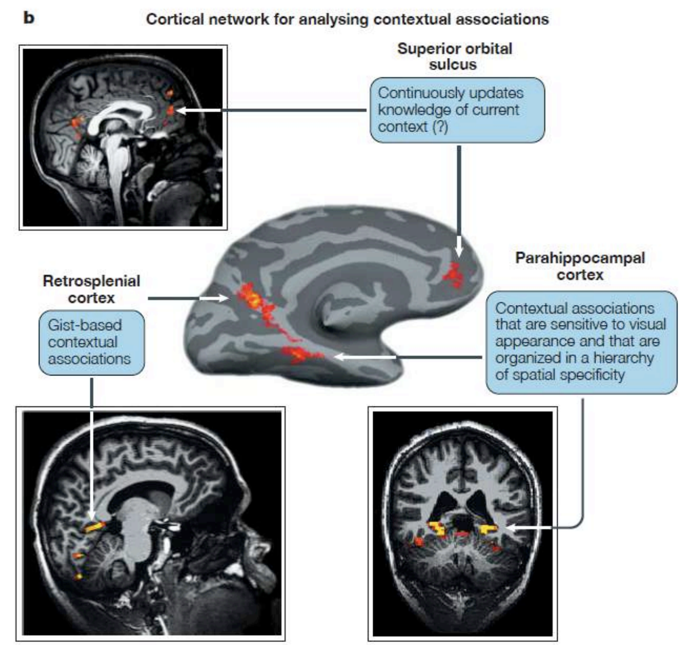


Figure 3 | **Cortical areas involved in processing context. a** | A functional magnetic resonance imaging (fMRI) statistical activation map representing the difference between perceiving objects that are strongly associated with a specific context and perceiving objects that are not associated with a unique context. This is a medial view of the left hemisphere, shown using a precise computer reconstruction where the sulci have been exposed by 'inflation'. The parahippocampal cortex (PHC) is circled in blue; the retrosplenial cortex (RSC) is circled in red; the superior orbital sulcus (SOS) is circled in yellow. Note that in all experimental conditions, subjects viewed similar looking colour photographs of meaningful, everyday common objects that were equally recognizable. Consequently, activation due to low-level processes was presumably subtracted out, and the differential activation map shown here represents only processes that are related to the level of contextual association. **b** | The cortical network for contextual associations among visual objects, suggested on the basis of existing evidence. Other types of context might involve additional regions (for example, hippocampus for navigation[125] and Broca's area for language-related context). Modified, with permission, from REF. 12 © (2003) Elsevier Science.

From: M. Bar, "Visual objects in context", Nature Reviews Neuroscience, August 2004.

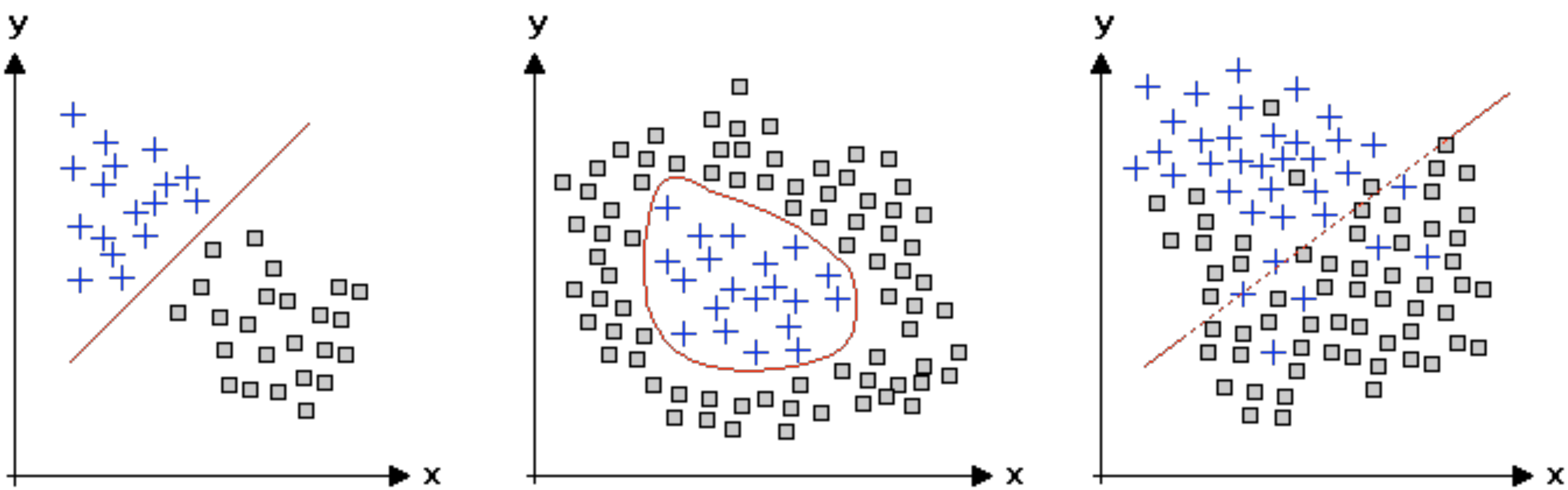
The importance of similarities

«Surely there is nothing more basic to thought and language than our sense of similarity.
[...]
And every reasonable expectation depends on resemblance of circumstances, together with our tendency to expect similar causes to have similar effects.»

Willard V. O. Quine
Natural Kinds (1969)

Today's view: Similarity as a by-product

Traditional machine learning and pattern recognition techniques are centered around the notion of **feature-vector**, and derive object similarities from vector representations.



Limitations of feature-vector representations

There are situations where either it is not possible to find satisfactory feature vectors or they are inefficient for learning purposes.

This is typically the case, e.g.,

- ✓ when features consist of both numerical and categorical variables
- ✓ in the presence of missing or inhomogeneous data
- ✓ when objects are described in terms of structural properties, such as parts and relations between parts, as is the case in shape recognition
- ✓ in the presence of purely relational data (graphs, hypergraphs, etc.)
- ✓ …

Application domains: Computational biology, adversarial contexts, social signal processing, medical image analysis, social network analysis, document analysis, network medicine, etc.

Signs of a transition?

The field is showing an increasing propensity towards anti-essentialist/ relational approaches, e.g.,

- ✓ Kernel methods
- ✓ Pairwise clustering (e.g., spectral methods, game-theoretic methods)
- ✓ Metric learning
- ✓ Graph transduction
- ✓ Dissimilarity representations (Duin et al.)
- ✓ Theory of similarity functions (Blum, Balcan, …)
- ✓ Relational / collective classification
- ✓ Graph mining
- ✓ Contextual object recognition
- ✓ …

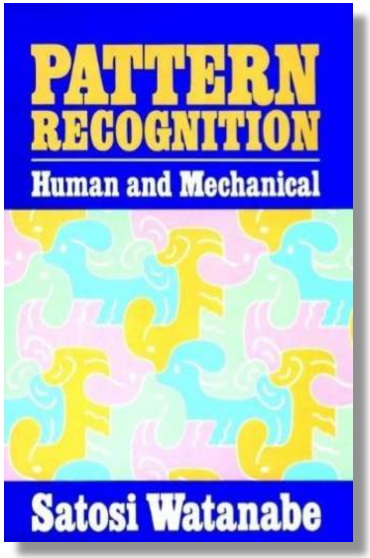
See also “link analysis” and the parallel development of “network science” …

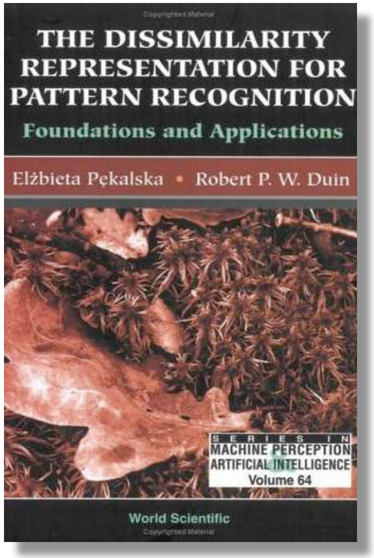
Readings

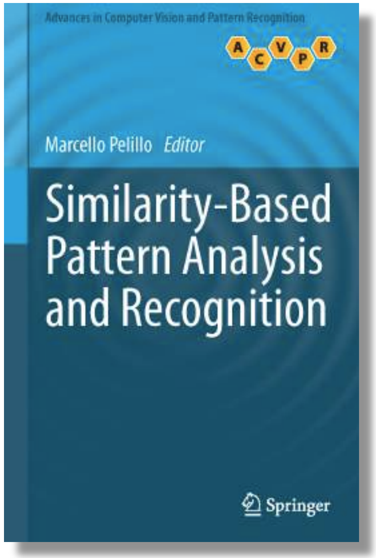
M. Pelillo and T. Scantamburlo. How mature is the field of machine learning? In: *Proc. AI*IA (*2013).

N. Cristianini. On the current paradigm in artificial intelligence. *AI Communication* (2014).

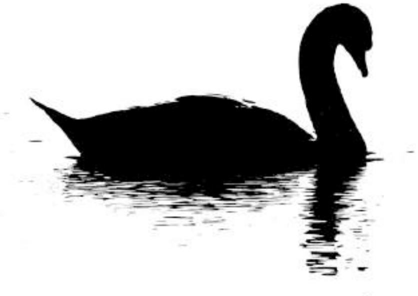
R. P. W. Duin and E. Pekalska: The science of pattern recognition. Achievements and perspectives. *Studies in Computational Intelligence* (2007).







Induction and its discontents



Machine learning as philosophy of science

«Machine learning studies inductive strategies as they might be carried out by algorithms.
The philosophy of science studies inductive strategies as they appear in scientific practice.
[...]
the two disciplines are, in large measure, one, at least in principle.
They are distinct in their histories, research traditions, investigative methodologies; however, the knowledge which they ultimately aim at is in large part indistinguishable.»

Kevin Korb
Machine learning as philosophy of science (2004)

The "problem" of induction

«If we look back at the history of thinking about induction, two figures appear to stand out from the remainder.

Francis Bacon appears, as he would have wished, as the first really systematic thinker about induction;

and **David Hume** appears as perhaps the first and certainly the greatest of all inductive sceptics, as a philosopher who bequeathed to his successors a Problem of Induction.»

John R. Milton
Induction before Hume (1987)

The two ways towards the truth

«There are and can be only two ways of searching into and discovering truth.

The one flies from the senses and particulars to the most general axioms, and from these principles, the truth of which it takes for settled and immovable, proceeds to judgment and to the discovery of middle axioms. And this way is now in fashion.

The other derives axioms from the senses and particulars, rising by a gradual and unbroken ascent, so that it arrives at the most general axioms last of all. **This is the true way, but as yet untried.**»

Francis Bacon

Novum Organum (1620)

No need for geniuses

«Our method of discovering the sciences, does not much depend upon subtlety and strength of genius, but lies level to almost every capacity and understanding.

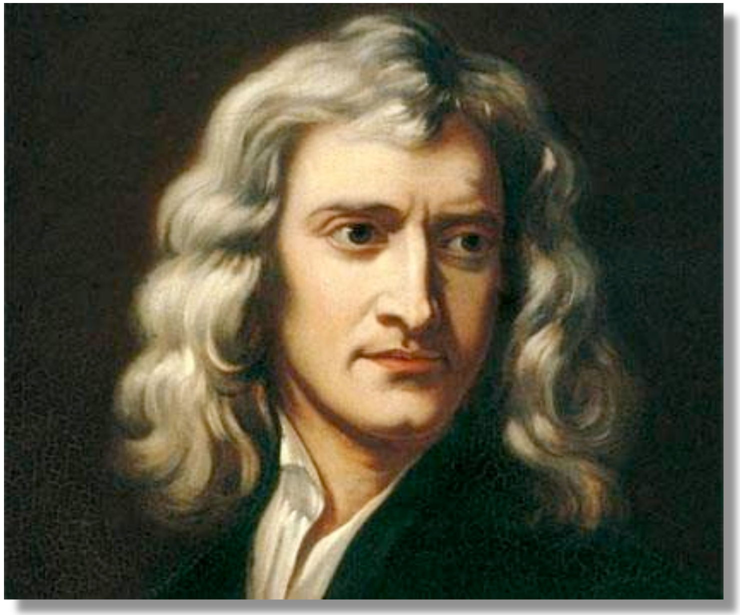
For, as it requires great steadiness and exercise of the hand to draw a true strait line, or a circle, by the hand alone, but little or no practice with the assistance of a ruler or compasses; so it is our method.»

Francis Bacon
Novum Organum (1620)

A great supporter

«In experimental philosophy, **propositions gathered from phenomena by induction should be taken to be either exactly or very nearly true** notwithstanding any contrary hypotheses, until yet other phenomena make such propositions either more exact or liable to exceptions.»

Isaac Newton
Philosophiae Naturalis Principia Mathematica (1726)



Logical necessity?

«The bread, which I formerly eat, nourished me; [...]
but does it follow, that other bread must also nourish me at another time, and that like sensible qualities must always be attended with like secret powers?
The consequence seems nowise necessary.»

David Hume
An Enquiry Concernstinct
g Human Understanding
(1748)

Justifying induction?

«All our experimental conclusions proceed upon the supposition that the future will be conformable to the past.
To endeavour, therefore, the proof of this last supposition by probable arguments, or arguments regarding existence, must be evidently **going in a circle**, and taking that for granted, which is the very point in question.»

David Hume
An Enquiry Concerning Human Understanding
(1748)

Logical paradoxes

«What tends to confirm an induction?
This question has been aggravated on the one hand by
Hempel's puzzle of the non-black non-ravens, and exacerbated
on the other by Goodman's puzzle of the grue emeralds.»

Willard V. O. Quine
Natural kinds (1969)

From black ravens ...

Nicod's principle: Universal generalizations are confirmed by their positive instances and falsified by their negative instances.

Example.

A black raven confirms the hypothesis "*All ravens are black*"

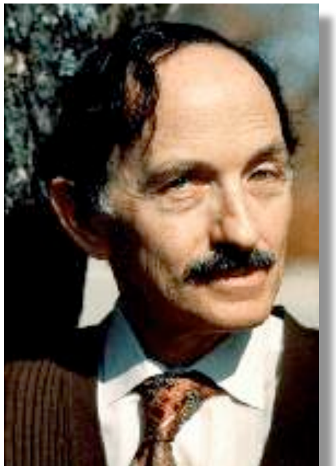
Equivalence principle: Whatever confirms a generalization confirms as well all its logical equivalents.

Example.

$\forall x\, (\mathrm{A}x \rightarrow \mathrm{B}x)$ is logically equivalent to $\forall x\, (\sim\mathrm{B}x \rightarrow \sim\mathrm{A}x)$

Hence, the hypothesis "*All ravens are black*" is logically equivalent to "*All non-black things are non-ravens*"

… to white shoes and indoor ornithology



«Hempel's paradox of confirmation can be worded thus 'A case of a hypothesis supports the hypothesis. Now the hypothesis that all crows are black is logically equivalent to the contrapositive that **all non-black things are non-crows**, and this **is supported by the observation of a white shoe.**'»

Irving J. Good
The white shoe is a red herring (1967)

«The prospect of being able to investigate ornithological theories without going out in the rain is so attractive that we know there must be a catch in it.»

Nelson Goodman
Fact, Fiction, and Forecast (1955)

Lawlike statements?

«That a given piece of copper conducts electricity increases the credibility of statements asserting that other pieces of copper conduct electricity [...]

But the fact that a given man now in this room is a third son does not increase the credibility of statements asserting that other men now in this room are third sons [...]

Yet in both cases our hypothesis is a generalization of the evidence statement. **The difference is that in the former case the hypothesis is a lawlike statement**; while in the latter case, the hypothesis is a merely contingent or accidental generality.»

Nelson Goodman
Fact, Fiction, and Forecast (1955)

Goodman's new riddle

Argument 1:

PREMISE	All the many emeralds observed prior to 2018 AD have been green
CONCLUSION	***All emeralds are green***

Argument 2:

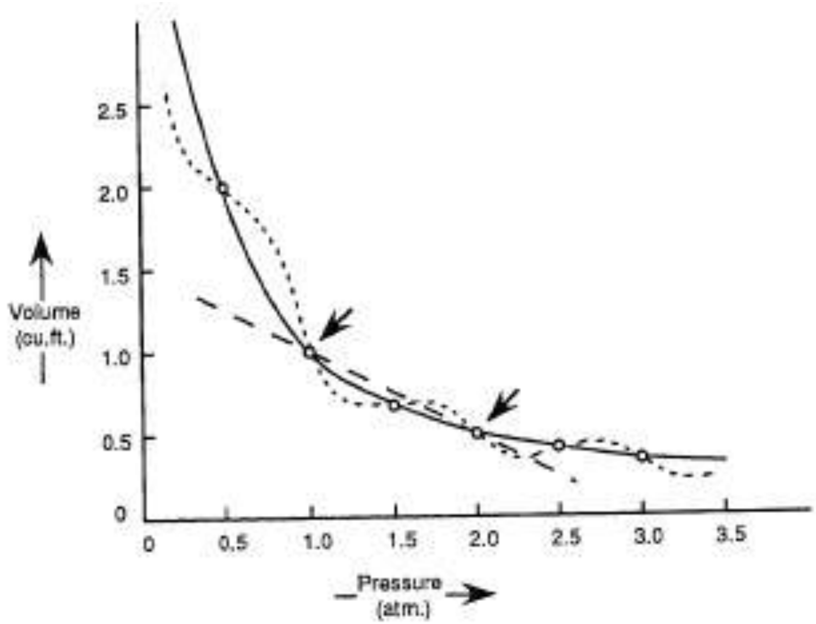
PREMISE	All the many emeralds observed prior to 2018 AD have been "grue"
CONCLUSION	***All emeralds are "grue"***

Definition: Any object is said to be *grue* if:

- ✓ it was first observed before 2018 AD and is green, or
- ✓ it was *not* first observed before 2018 AD and is blue

If all evidence is based on observations made before 2018 AD, then the second argument should be considered as good as the first ...

Goodman's riddle and model selection



Boyle's Law (solid line) and alternative laws.

There's always an infinity of mutually contradictory hypotheses that fit the data, but which is best confirmed?

***Customary answer**:* choose the simplest one (Occam's razor). But… why?

The probabilistic turn

«I am convinced that it is impossible to expound the methods of induction in a sound manner, without resting them upon the theory of probability.

Perfect knowledge alone can give certainty, and in nature perfect knowledge would be infinite knowledge, which is clearly beyond our capacities. We have, therefore, to content ourselves with partial knowledge—knowledge mingled with ignorance, producing doubt.»

William S. Jevons
The Principles of Science (1874)

But … what does "probability" mean?

Classical view (*Laplace, Pascal, J. Bernoulli, Huygens, Leibniz, …*)

Probability = ratio # favorable cases / # possible cases

Frequentist view (*von Mises, Reichenbach, …*)

Probability = limit of relative frequencies

Logical view (*Keynes, Jeffreys, Carnap, …*)

Probability = logical relations between propositions ("partial implication")

Subjectivist view (*Ramsey, de Finetti, Savage, …*)

Probability = a (personal) agent's "degree of belief "

But also: Propensity (Popper), Best-system (Lewis), …

Bayesianism to the rescue?

«Through much of the twentieth century, the unsolved problem of confirmation hung over philosophy of science. What is it for an observation to provide evidence for, or confirm, a scientific theory? [...]
The situation has now changed. Once again a large number of philosophers have real hope in a theory of confirmation and evidence. The new view is called *Bayesianism*.»

Peter Godfrey-Smith
Theory and Reality (2003)

The three tenets of Bayesianism

Bayesian confirmation theory (BCT) makes the following assumptions:

1. It is assumed that agents assigns *degrees of belief*, or credences, to different competing hypotheses, reflecting the agent's level of expectation that a particular hypothesis will turn out to be true

2. The degrees of belief are assumed to behave mathematically like probabilities, thus they can be called *subjective probabilities*

3. Agents are assumed to learn from the evidence by what is called the *Bayesian conditionalization rule*. The conditionalization rule directs one to update his credences in the light of new evidence in a quantitatively exact way

In BCT, evidence e confirms hypothesis h if:

$$P(h \mid e) > P(h)$$

The Bayesian "machine"

- ✓ determine the prior probability of h
- ✓ if e_1 is observed, calculate the posterior probability $P(h \mid e_1)$ via Bayes' theorem
- ✓ consider this posterior probability as your new prior probability of h
- ✓ if e_2 is observed, calculate the posterior probability $P(h \mid e_2)$ via Bayes' theorem
- ✓ consider this posterior probability as your new prior probability of h
- ✓ …

Bayesians' answer to confirmation paradoxes

The ravens: White shoes do in fact confirm the hypothesis that all ravens are black, but only to a negligible degree.

The grue emeralds: Both hypotheses ("green" and grue") are OK, but most people would assign a higher prior to the "green" hypothesis than to the "grue" one. (But… why is it so?)

Challenges to Bayesianism

Priors. Where do they come from? Also, initial set of prior probabilities can be chosen freely ⇒ how could a strange assignment of priors be criticized, so long as it follows the axioms?

Old evidence. Existing evidence can in fact confirm a new theory, but according to Bayesian kinematics it cannot (e.g., the perihelion of Mercury and Einstein's general relativity theory).

If e is known before theory T is introduced, then we have $P(e) = 1 = P(e|T)$, which yields:

$$P_{new}(T \mid e) = \frac{P(T)P(e \mid T)}{P(e)} = P(T)$$

⇒ posterior probability of T is the same as its prior probability!

Solomonoff induction

«Solomonoff completed the Bayesian framework by providing a rigorous, unique, formal, and universal choice for the model class and the prior.»

Marcus Hutter
On universal prediction and Bayesian confirmation (2007)

Basic ingredients:

- ✓ Epicurus (keep all explanations consistent with the data)
- ✓ Occam (choose the simplest model consistent with the data)
- ✓ Bayes (combine evidence and priors)
- ✓ Turing (compute quantities of interest)
- ✓ Kolmogorov (measure simplicity/complexity)

Data expressed as binary sequences
Hypotheses expressed as algorithms (processes that generate data)

Bad news: Solomonoff induction is intractable .... (use approximation)

A never-ending debate

«The dispute between the Bayesians and the anti-Bayesians has been one of the major intellectual controversies of the 20th century.»

Donald Gillies, *Was Bayes a Bayesian?* (2003)

«All that can be said about 'inductive inference' […], essentially, reduces […] to Bayes' theorem.»

Bruno De Finetti, *Teoria della probabilità* (1970)

«The theory of inverse probability is founded upon an error, and must be wholly rejected.»

Ronald A. Fisher
Statistical Methods for Research Workers (1925)

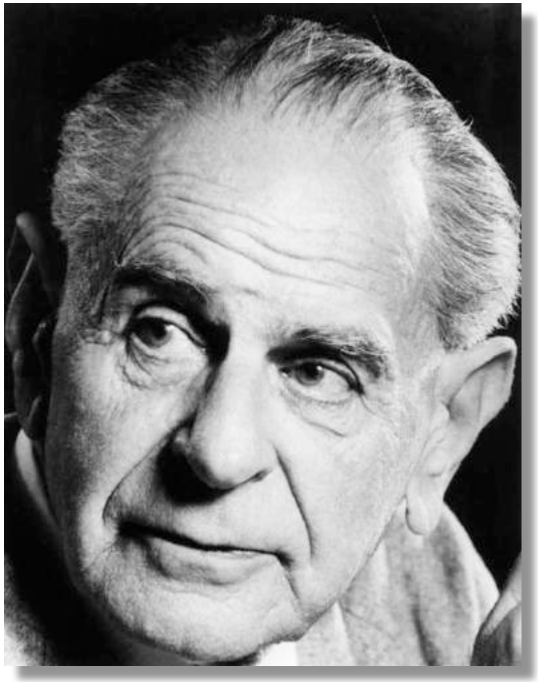
Against induction

«I think that I have solved a major philosophical problem: the problem of induction.»

Karl Popper
Objective Knowledge (1972)

«Induction, i.e. inference based on many observations, is a myth.
It is neither a psychological fact, nor a fact of ordinary life, nor one of scientific procedure.»

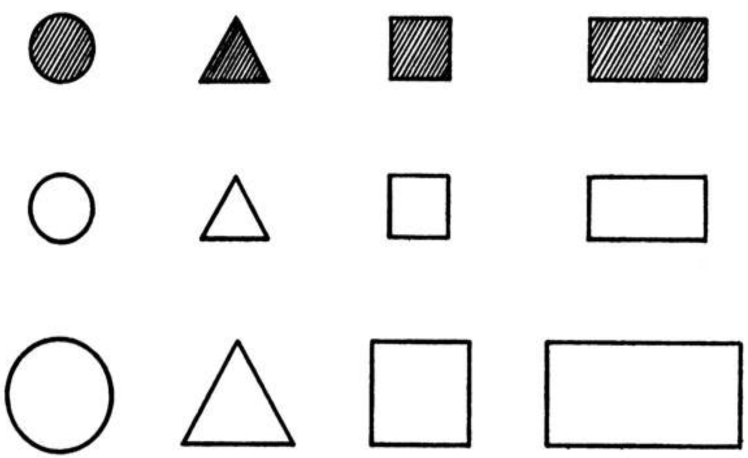
Karl Popper
Conjectures and Refutations (1963)



Observation is selective

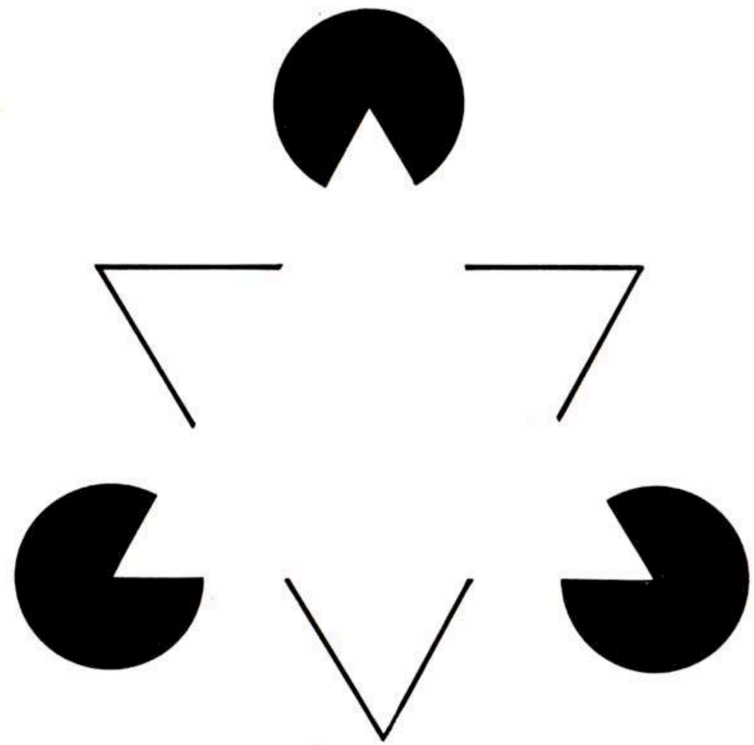
«The fundamental doctrine which underlies all theories of induction
is the doctrine of the **primacy of repetitions**.
[…]
All the repetitions which we experience are **approximate repetitions**;»

«**Repetition presupposes similarity**,
and similarity presupposes a point of view – a theory, or an expectation.»



Karl Popper
The Logic of Scientific Discovery (1959)
Objective Knowledge (1972)

Theory-laden observations



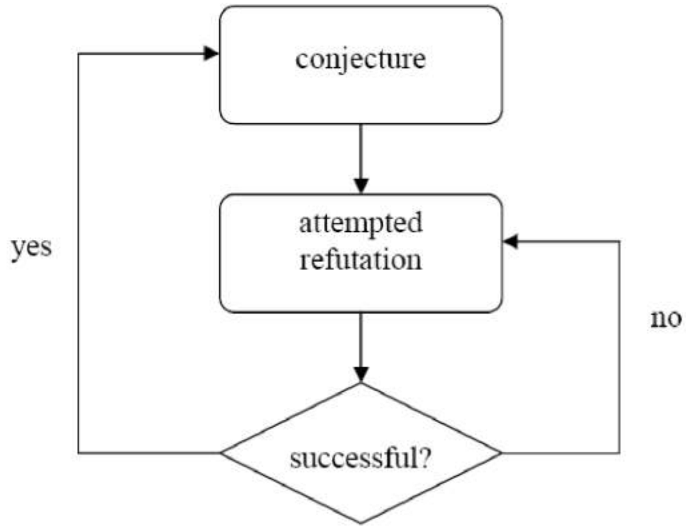
Popper's scientific method

«My whole view of scientific method may be summed up by saying that it consists of these three steps:

1 We stumble over some problem.
2 We try to solve it, for example by proposing some theory.
3 We learn from our mistakes, especially from those brought home to us by the critical discussion of our tentative solutions [...]

Or in three words: *problems – theories – criticism*.»

Karl Popper
The Myth of the Framework (1994)



[Wüthrich, 2010]

Feynman's version

«In general we look for a new law by the following process.
First we guess it.
Then we compute the consequences of the guess to see what would be implied if this law that we guessed is right.
Then we compare the result of the computation to nature, with experiment or experience, compare it directly with observation, to see if it works.

If it disagrees with experiment it is wrong. In that simple statement is the key to science.»

Richard Feynman
The Character of Physical Law (1965)

A "simple" example

By some chance, you come across the relations:

$$3 + 7 = 10, \quad 3 + 17 = 20, \quad 13 + 17 = 30$$

It strikes you that the numbers 3, 7, 13, and 17 are odd primes.

Now, the sum of two odd primes is necessarily an even number, but …
what about the other even numbers?

From: G. Polya, Mathematics and Plausible Reasoning, Vol. 1, (1954)

A "simple" example

The first even number which is a sum of two odd primes is, of course,

$$6 = 3 + 3.$$

Looking beyond 6, we find that:

$$8 = 3 + 5$$

$$10 = 3 + 7 = 5 + 5$$

$$12 = 5 + 7$$

$$14 = 3 + 11 = 7 + 7$$

$$16 = 3 + 13 = 5 + 11$$

Question: *Will it go on like this forever?*

From: G. Polya, Mathematics and Plausible Reasoning, Vol. 1, (1954)

A conjecture

Every even integer greater than 2 can be expressed as the sum of two primes.

Letter from Goldbach to Euler dated 7 June 1742

«Every even integer is a sum of two primes. I regard this as a completely certain theorem, although I cannot prove it.»

Leonhard Euler to Christian Goldbach
30 June 1742

Some (scanty) additional evidence

$$6 = 3 + 3$$
$$8 = 3 + 5$$
$$10 = 3 + 7 = 5 + 5$$
$$12 = 5 + 7$$
$$14 = 3 + 11 = 7 + 7$$
$$16 = 3 + 13 = 5 + 11$$
$$18 = 5 + 13 = 7 + 11$$
$$20 = 3 + 17 = 7 + 13$$
$$22 = 3 + 19 = 5 + 17 = 11 + 11$$
$$24 = 5 + 19 = 7 + 17 = 11 + 13$$
$$26 = 3 + 23 = 7 + 19 = 13 + 13$$
$$28 = 5 + 23 = 11 + 17$$
$$30 = 7 + 23 = 11 + 19 = 13 + 17.$$

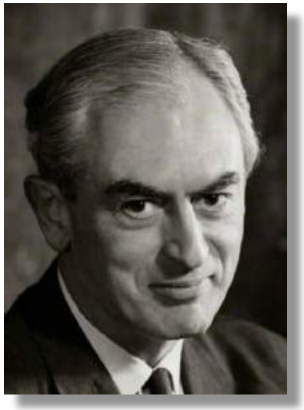
bound	reference
1×10^{4}	Desboves 1885
1×10^{5}	Pipping 1938
1×10^{8}	Stein and Stein 1965ab
2×10^{10}	Granville et al. 1989
4×10^{11}	Sinisalo 1993
1×10^{14}	Deshouillers et al. 1998
4×10^{14}	Richstein 1999, 2001
2×10^{16}	Oliveira e Silva (Mar. 24, 2003)
6×10^{16}	Oliveira e Silva (Oct. 3, 2003)
2×10^{17}	Oliveira e Silva (Feb. 5, 2005)
3×10^{17}	Oliveira e Silva (Dec. 30, 2005)
12×10^{17}	Oliveira e Silva (Jul. 14, 2008)
4×10^{18}	Oliveira e Silva (Apr. 2012)

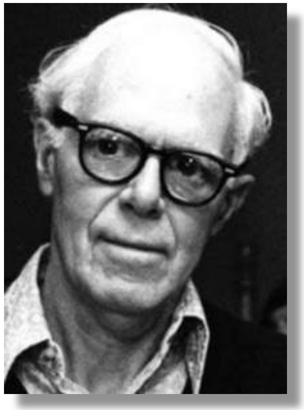
From: http://mathworld.wolfram.com

Reactions to Popper

«I think Popper is incomparably the greatest philosopher of science that has ever been.»

Peter Medawar





«Popper's great and tireless efforts to expunge the word *induction* from scientific and philosophical discourse has utterly failed.»

Martin Gardner

Popper as a precursor of Vapnik

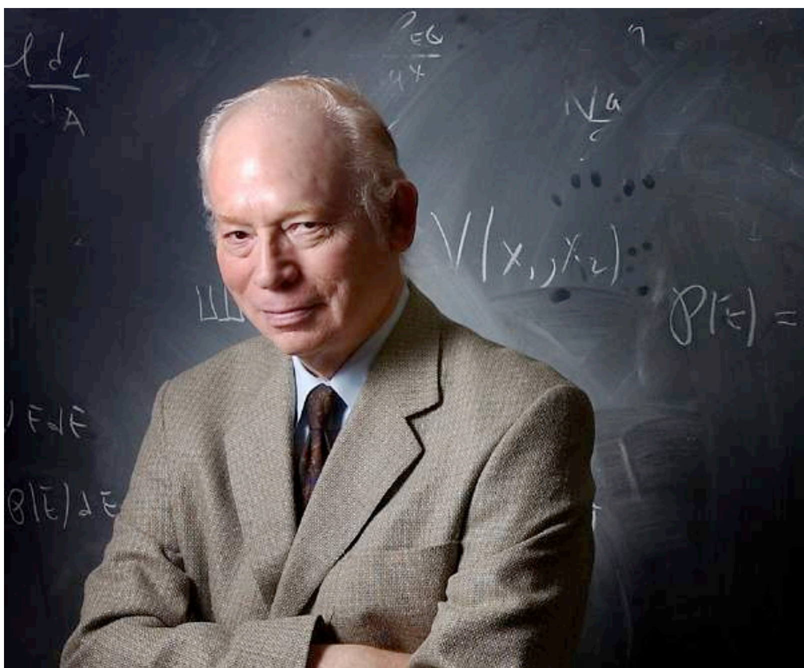
«Let me remark how amazing Popper's idea was. In the 1930's Popper suggested a general concept determining the generalization ability (in a very wide philosophical sense) that in the 1990's turned out to be one of the most crucial concepts for the analysis of consistency of the ERM inductive principles.»

Vladimir Vapnik
The Nature of Statistical Learning Theory (2000)

Let the scientists speak / 1

«Scientists and historians of science have long ago given up the old view of Francis Bacon, that scientific hypotheses should be developed by patient and unprejudiced observation of nature.

It is glaringly obvious that Einstein did not develop general relativity by poring over astronomical data.»



Steven Weinberg

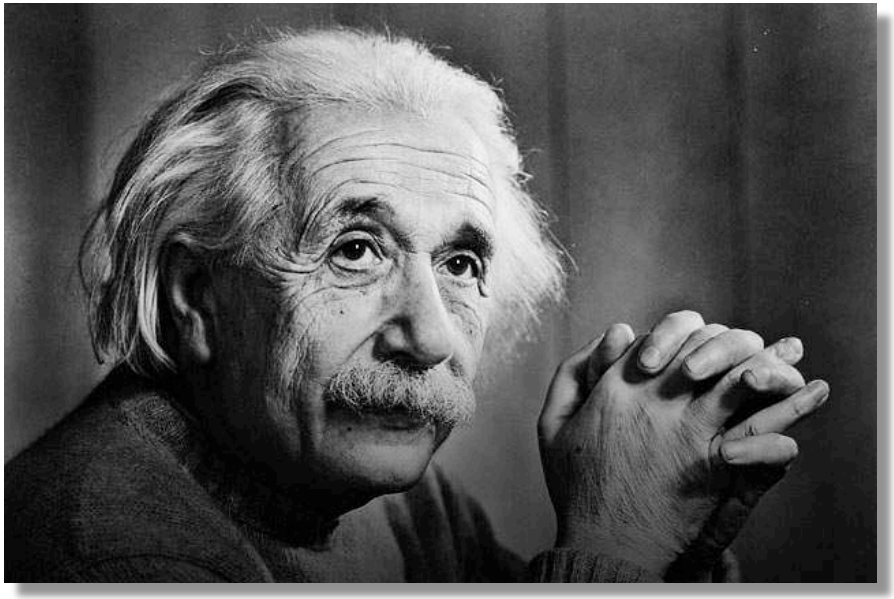
Dreams of a Final Theory (1993)

Let the scientists speak / 2

«The truly great advances in our understanding of nature originated in a manner almost diametrically opposed to induction.»

Albert Einstein

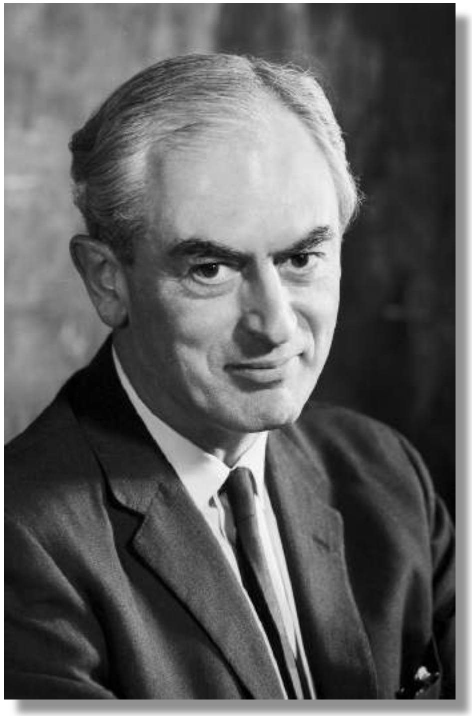
Induction and deduction in physics (1919)



Let the scientists speak / 3

«Deductivism in mathematical literature and inductivism in scientific papers are simply the postures we choose to be seen in when the curtain goes up and the public sees us.

The theatrical illusion is shattered if we ask what goes on behind the scenes. **In real life discovery and justification are almost always different processes.**»



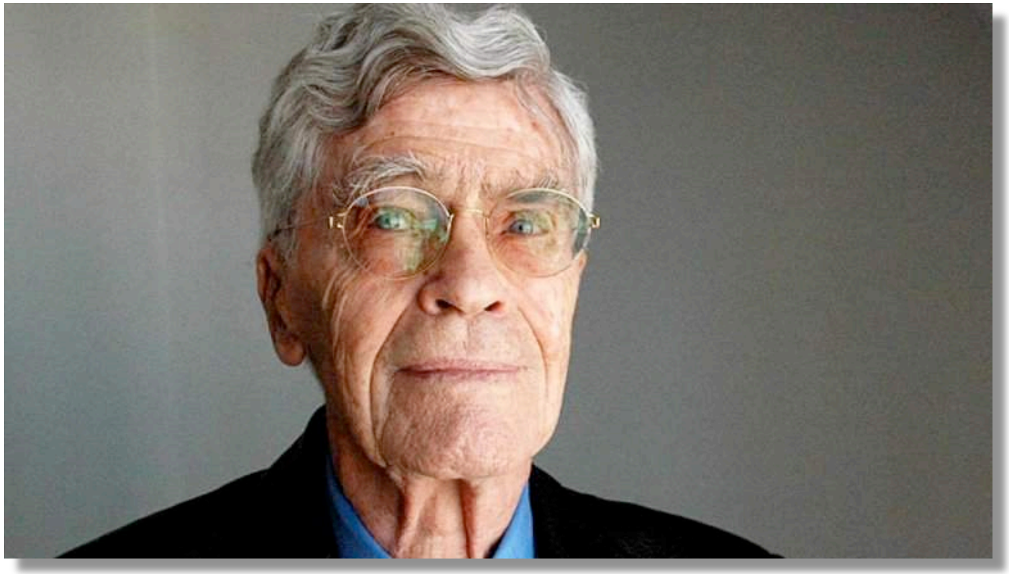
Peter B. Medawar

Induction and Intuition in Scientific Thought (1969)

A role for induction?

«Induction, which is but one of the kinds of plausible reasoning, contributes modestly to the framing of scientific hypotheses, **but is indispensable for their test**, or rather for the empirical stage of their test.»

Mario Bunge
The place of induction in science (1960)



A bag of tricks?

- Enumerative induction
- Deduction
- Eliminative induction
- Abduction (a.k.a. retroduction, or "inference to the best explanation")
- Analogy
- ....

Recall Ramachandran's claim about perception:

> «One could take the pessimistic view that the visual system often cheats, i.e uses rules of thumb, short-cuts, and clever sleight-of-hand tricks that were acquired by trial and error through millions of years of natural selection.»

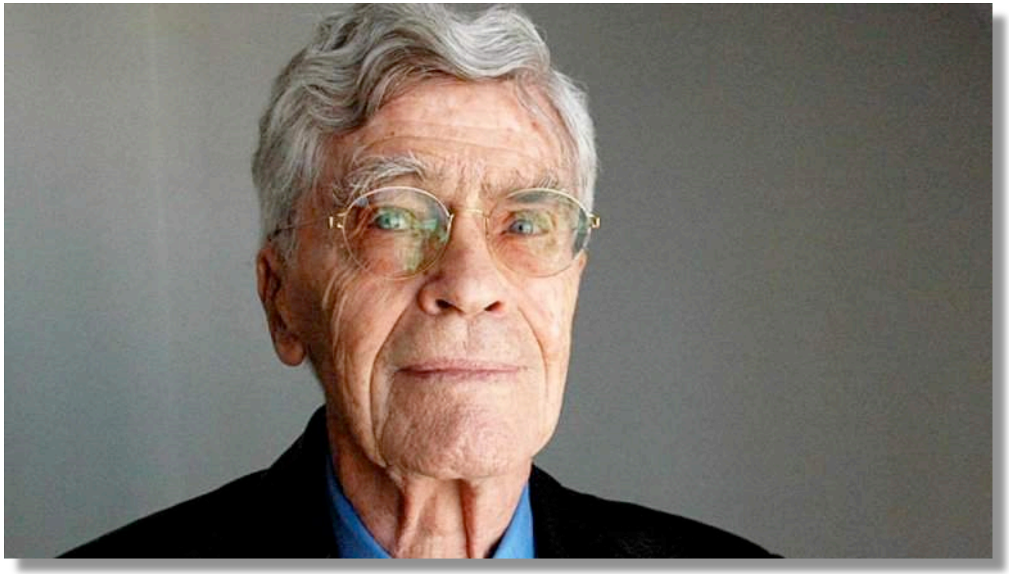
Vilayanur S. Ramachandran

The neurobiology of perception (1985)

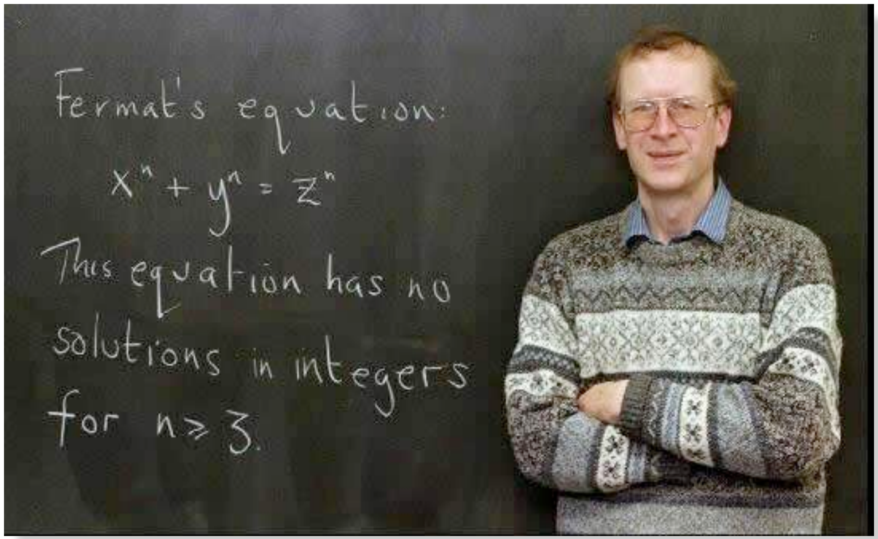
Intuition?

«Intuition is the collection of odds and ends where we place all the intellectual mechanisms which we do not know how to analyze or even name with precision, or which we are not interested in analyzing or naming.»

Mario Bunge
Intuition and Science (1962)



The Aha! Experience



Andrew Wiles
Princeton University

«I have discovered a truly marvelous proof of this, which this margin is too narrow to contain.»

Pierre de Fermat (1601–1665)

From the movie 'The Proof', produced by Nova and aired on PBS on October 28, 1997

The "Aha!" experience

«At this moment I left Caen, where I was then living, to take part in a geological conference arranged by the School of Mines. The incidents of the journey made me forget my mathematical work.

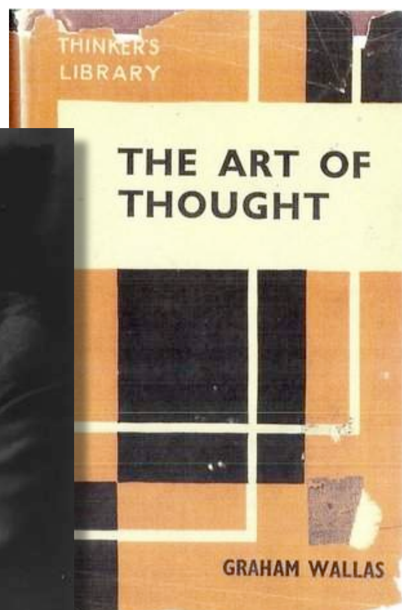
When we arrived at Coutances, we got into a break to go for a drive, and, **just as I put my foot on the step, the idea came to me, though nothing in my former thoughts seemed to have prepared me for it**, that the transformations I had used to define Fuchsian functions were identical with those of non-Euclidian geometry.»

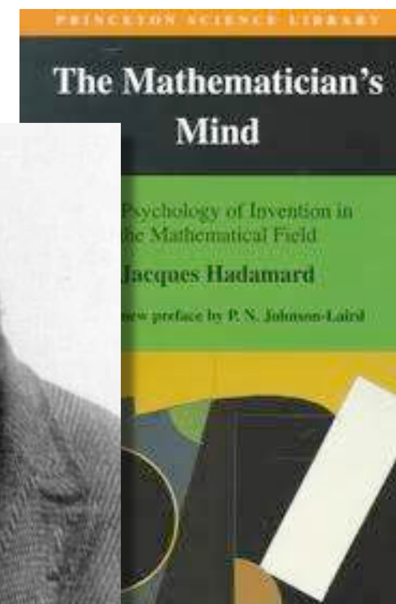
Henri Poincaré
Science and Method (1908)

Poincaré's legacy: Wallas and Hadamard

«Poincaré's observations throw a resplendent light on relations between the conscious and the unconscious, between the logical and the fortuitous, which lie at the base of the problem [of mathematical discovery].»

Jacques Hadamard
The Mathematician's Mind (1945)

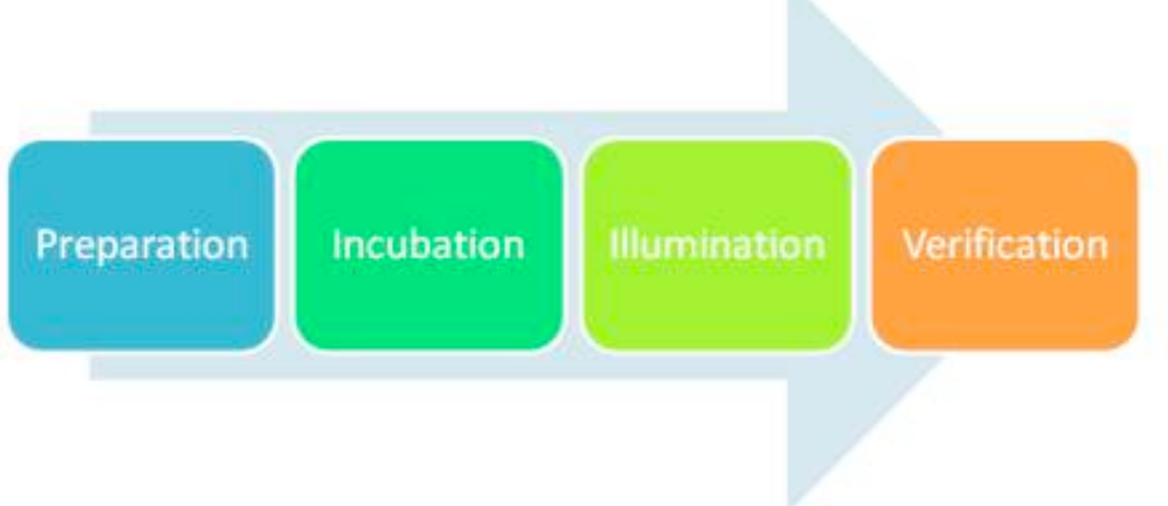




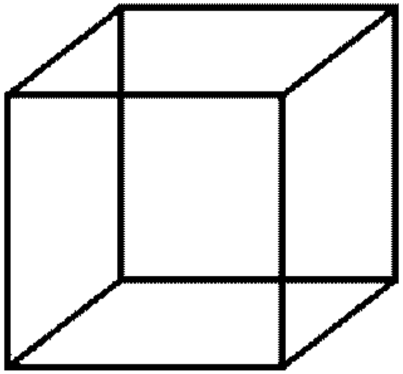
The four stages of invention

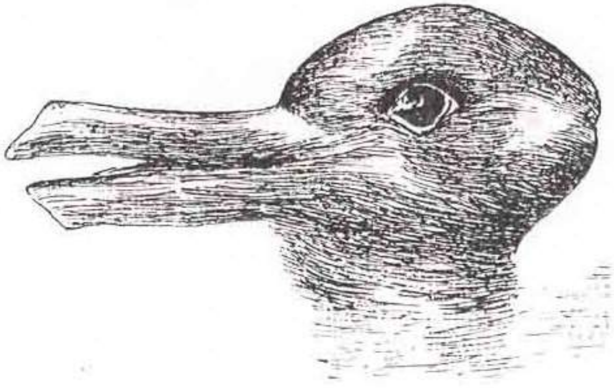
«The same character of suddenness and spontaneousness had been pointed out, some years earlier, by another great scholar of contemporary science. Helmholtz reported it in an important speech delivered in 1896.
[…]
Graham Wallas, in his *Art of Thought*, suggested calling it **illumination**, this illumination being generally preceded by an **incubation** stage wherein the study seems to be completely interrupted and the subject dropped.»

Jacques Hadamard
The Mathematician's Mind (1945)



"Aha!" as Gestalt switches

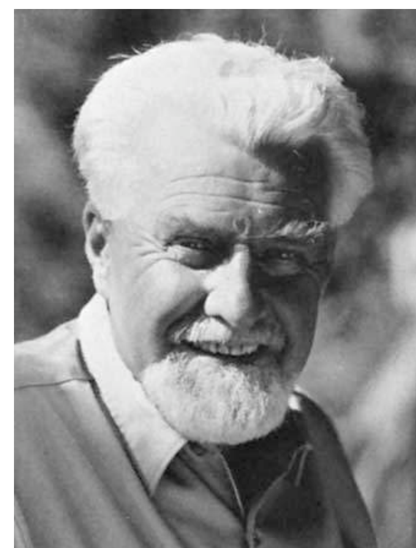




Discovery and *Gestalts*

«The process of discovery is akin to the recognition of shapes as analysed by Gestalt psychology.»

Michael Polanyi
Science, Faith, and Society (1946)



«In my opinion every discovery of a complex regularity comes into being through the function of gestalt perception.»

Konrad Lorenz
Gestalt Perception as Fundamental to Scientific Knowledge (1959)

Is intuition mechanizable?

«The act of discovery escapes logical analysis; there are no logical rules in terms of which a "discovery machine" could be constructed that would take over the creative function of the genius.»

Hans Reichenbach, *The Rise of Scientific Philosophy* (1951)

«The situation has provided a cue; this cue has given the expert access to information stored in memory, and the information provides the answer. **Intuition is nothing more and nothing less than recognition**.»

Herbert A. Simon, *What is an explanation of behavior?* (1992)

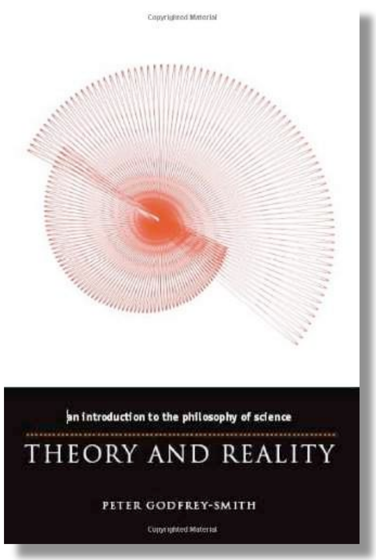
Readings

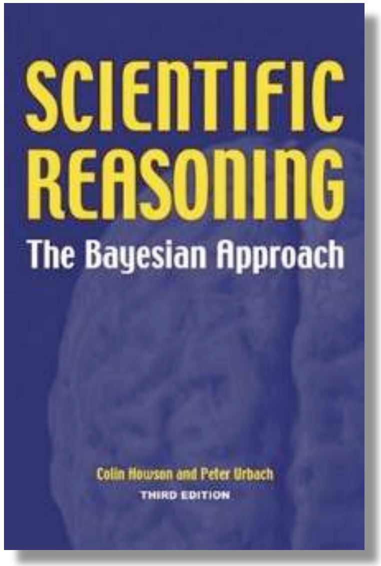
G. Harman and S. Kulkarni. *Statistical learning theory as a framework for the philosophy of induction* (2008).

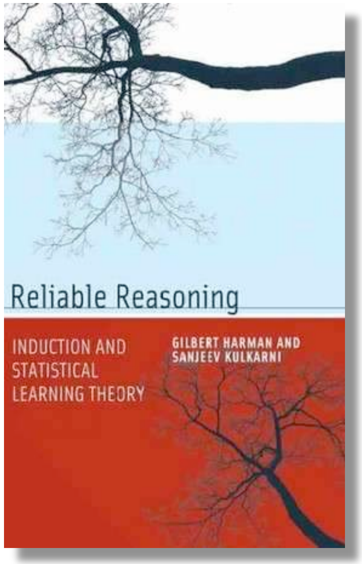
D. Corfield, B. Schölkopf, and V. Vapnik. *Falsificationism and statistical learning theory: Comparing the Popper and the Vapnik-Chervonenkis dimensions* (2009).

M. Hutter. *On universal prediction and Bayesian confirmation* (2007).

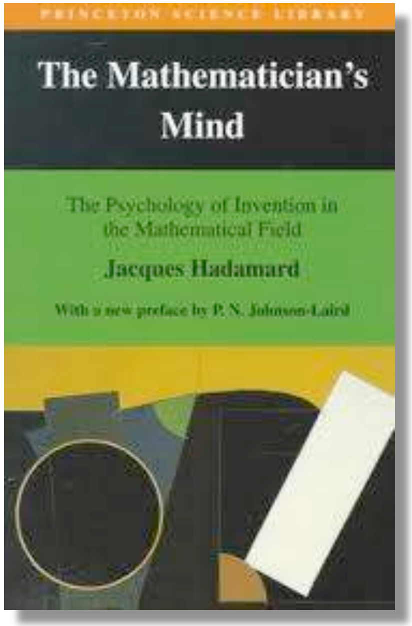
S. Rathmanner and M. Hutter. *A philosophical treatise of universal induction* (2011).

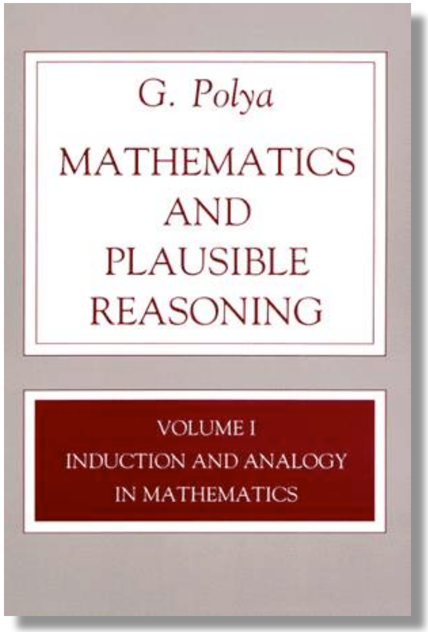


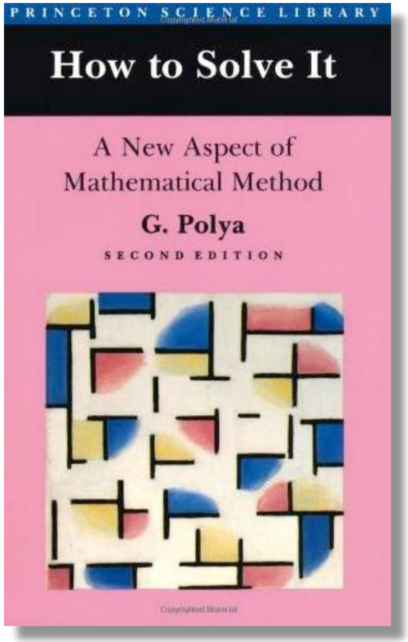


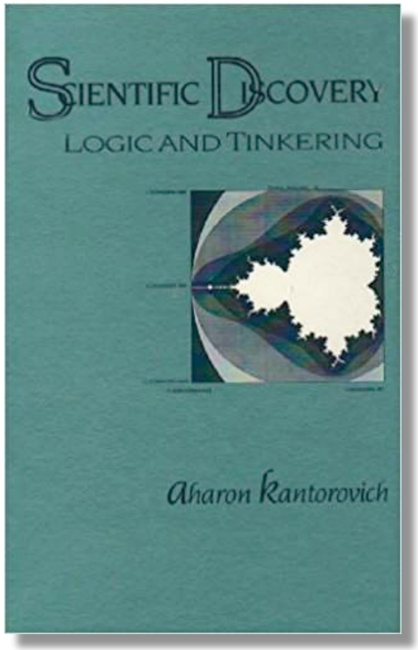


Readings


PRINCETON SCIENCE LIBRARY
The Mathematician's Mind
The Psychology of Invention in the Mathematical Field
Jacques Hadamard
With a new preface by P. N. Johnson-Laird


G. Polya
MATHEMATICS AND PLAUSIBLE REASONING
VOLUME I
INDUCTION AND ANALOGY IN MATHEMATICS


PRINCETON SCIENCE LIBRARY
How to Solve It
A New Aspect of Mathematical Method
G. Polya
SECOND EDITION


SCIENTIFIC DISCOVERY
LOGIC AND TINKERING
Aharon Kantorovich

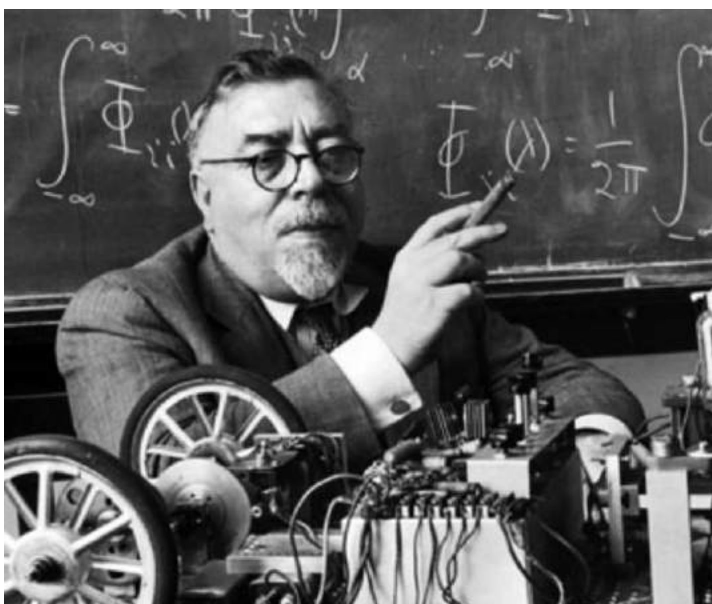
Machine learning and society

Wiener's warning

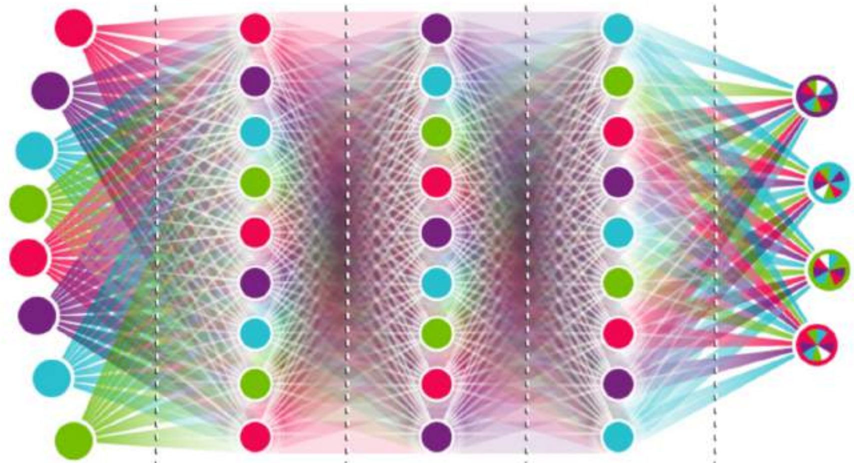
«Any machine constructed for the purpose of making decisions, if it does not possess the power of learning, will be completely literal-minded.

Woe to us if we let it decide our conduct, unless we have previously examined its laws of action, and know fully that its conduct will be carried out on principles acceptable to us!»

Norbert Wiener
The Human Use of Human Beings (1950)



Opacity



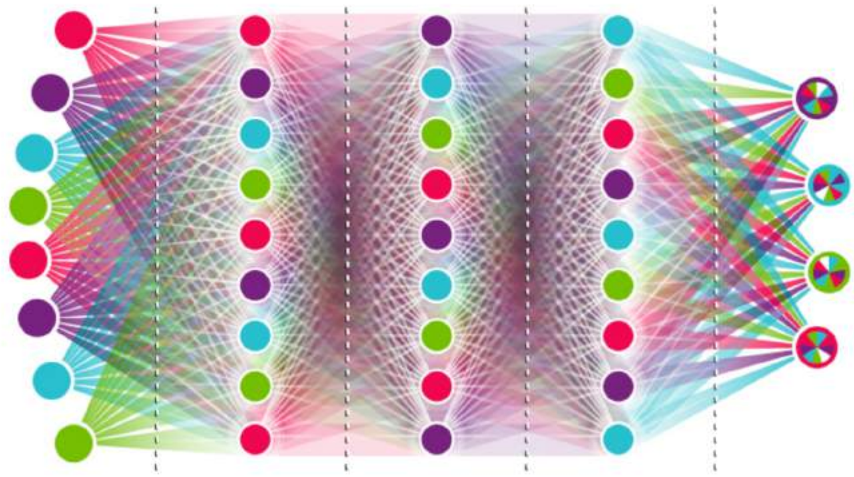
Gorilla!

Google apologises for Photos app's racist blunder

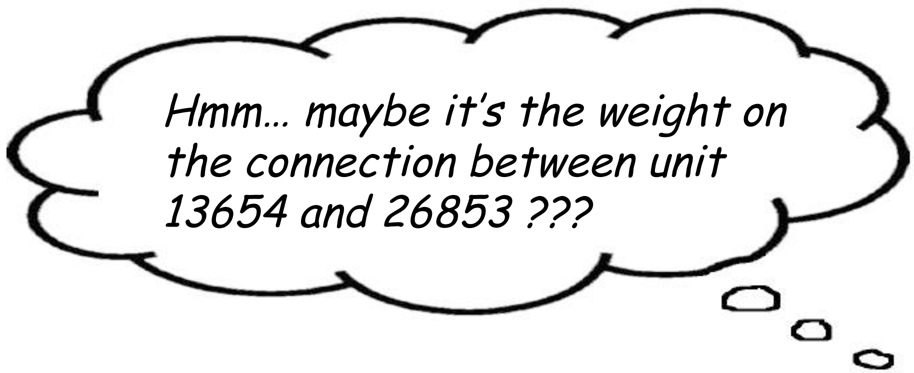
1 July 2015

Share

Debugging?



Gorilla!



After three years ...

WIRED

TOM SIMONITE BUSINESS 01.11.18 07:00 AM

WHEN IT COMES TO GORILLAS, GOOGLE PHOTOS REMAINS BLIND

Towards more frightening scenarios

The New York Times

POLITICS

Sent to Prison by a Software Program's Secret Algorithms

Sidebar
By ADAM LIPTAK MAY 1, 2017

Eric L. Loomis

You're identified, through the COMPAS assessment, as an individual who is at high risk to the community.

Accuracy *vs* transparency

«Deploying unintelligible black-box machine learned models is risky – high accuracy on a test set is NOT sufficient. Unfortunately, the most accurate models usually are not very intelligible (e.g., random forests, boosted trees, and neural nets), and the most intelligible models usually are less accurate (e.g., linear or logistic regression).»

Rich Caruana

Friends don't let friends deploy models they don't understand (2016)

2016 Workshop on Human Interpretability in Machine Learning

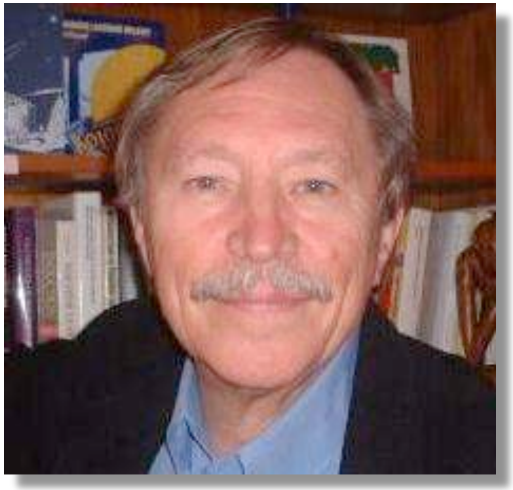
WHI 2016 @ ICML, New York, June 23, 2016

Back to the 1980's

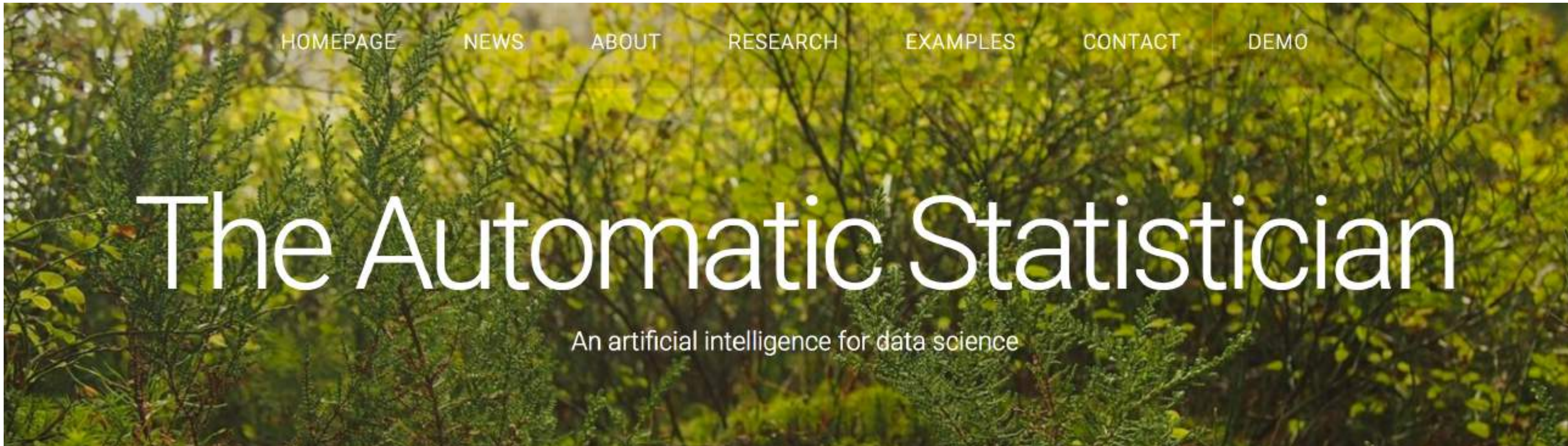
«The results of computer induction should be symbolic descriptions of given entities, semantically and structurally similar to those a human expert might produce observing the same entities.

Components of these descriptions should be comprehensible as single 'chunks' of information, directly **interpretable in natural language**, and should relate quantitative and qualitative concepts in an integrated fashion.»

Ryszard S. Michalski
A theory and methodology of inductive learning (1983)



The “automatic statistician”



«The aim is to find models which have both good predictive performance,
and are somewhat interpretable.
The Automatic Statistician generates a natural language summary of the analysis, producing a 10-15 page report with plots and tables describing the analysis.»

Zoubin Ghahramani (2016)

But why should we care?

«There are things we cannot verbalize.
When you ask a medical doctor why he diagnosed
this or this, he's going to give you some reasons.
But how come it takes 20 years to make a good doctor?
Because the information is just not in books.»

Stéphane Mallat (2016)

«You use your brain all the time; you trust your brain all the time; and you have no idea how your brain works.»

Pierre Baldi (2016)

From: D. Castelvecchi, Can we open the black box of AI? *Nature* (October 5, 2016)

Indeed, sometimes we should …

Explanation is a core aspect of due process (Strandburg, HUML 2016):

- ✓ Judges generally provide either written or oral explanations of their decisions
- ✓ Administrative rule-making requires that agencies respond to comments on proposed rules
- ✓ Agency adjudicators must provide reasons for their decision to facilitate judicial review

Example #1. In many countries, banks that deny a loan have a legal obligation to say why — something a deep-learning algorithm might not be able to do.

Example #2. If something were to go wrong as a result of setting the UK interest rates, the Bank of England can't say: "the black box made me do it".

From: D. Castelvecchi, Can we open the black box of AI? *Nature* (October 5, 2016)

A right to explanation?

General
Data
Protection
Regulation

Art. 13

A data subject has the right to obtain
"meaningful information about the logic involved"

Pedro Domingos
@pmddomingos

Starting May 25, the European Union will require algorithms to explain their output, making deep learning illegal.

5:59 AM - Jan 29, 2018

344 247 people are talking about this

ARTICLE — *International Data Privacy Law*, 2017, Vol. 7, No. 2

Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation

Sandra Wachter*, Brent Mittelstadt** and Luciano Floridi***

Neutrality?

Kranzberg's First Law of Technology
Technology is neither good nor bad; nor is it neutral.

ProPublica

Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica
May 23, 2016

	White	African American
Labeled Higher Risk, But Didn't Re-Offend	23,5%	**44,9%**
Labeled Lower Risk, Yet Did Re-Offend	**47,7%**	28,0%

March 23, 2016

INDEPENDENT

A few hours later …

INDEPENDENT

24 March 2016

The (well-known) question of bias

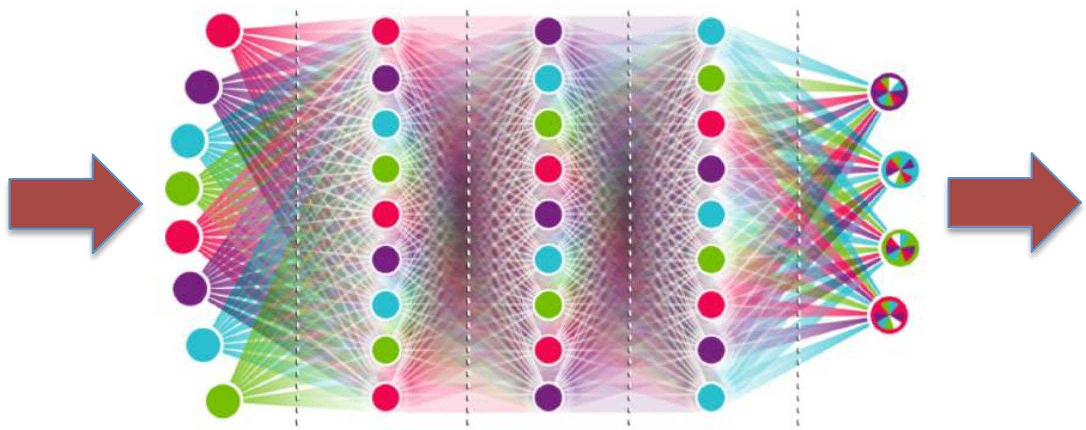
«So, what is the value of current datasets when used to train algorithms for object recognition that will be deployed in the real world?

The answer that emerges can be summarized as: "better than nothing, but not by much".»

Antonio Torralba and Alexei Efros

Unbiased look at dataset bias (2011)

A tale of tanks

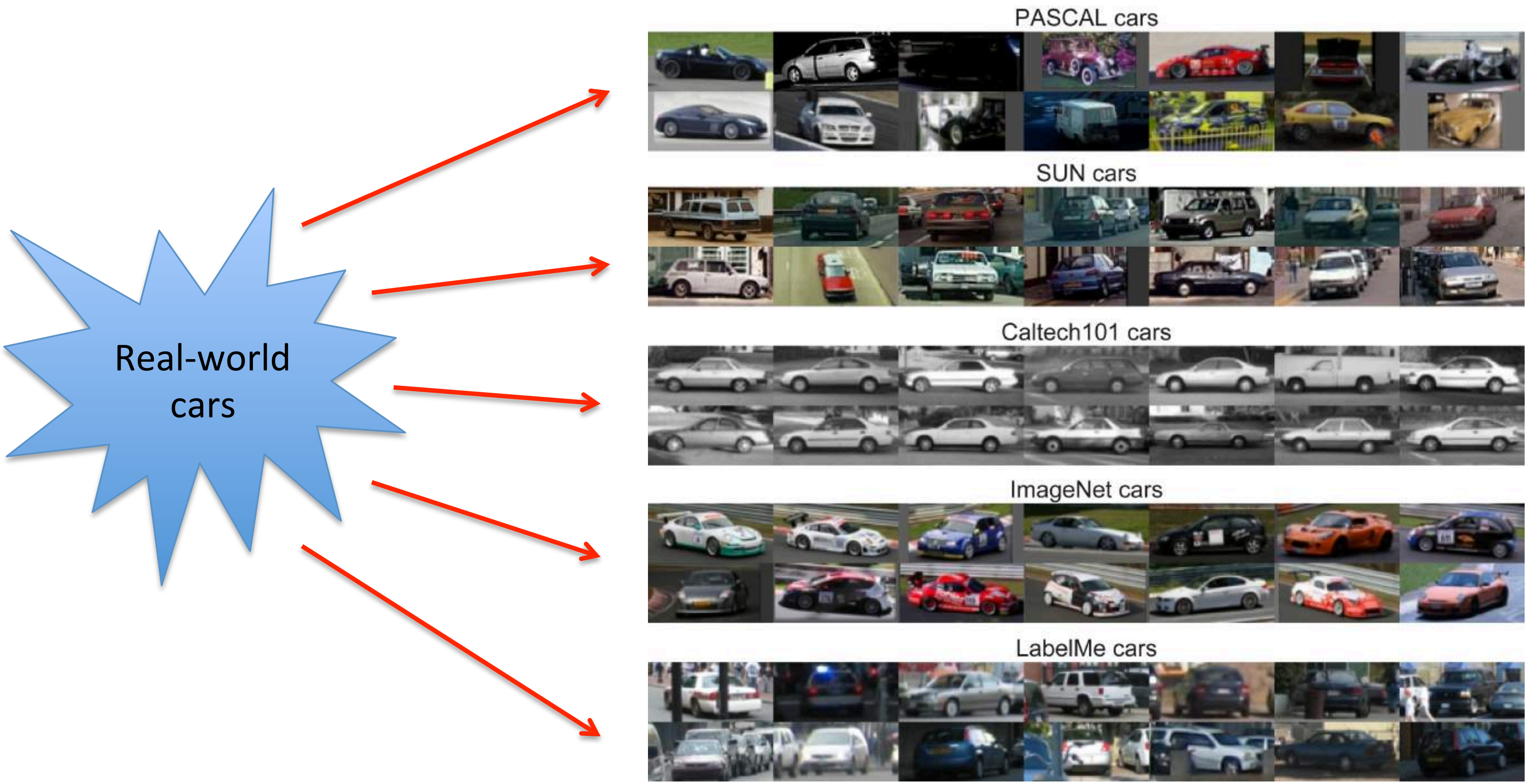


American

Russian

See: https://www.gwern.net/Tanks

The map is not the territory



The curse of biased datasets

«We would like to ask the following question: how well does a typical object detector trained on one dataset generalize when tested on a representative set of other datasets, compared with its performances on the “native” test set?»

A. Torralba and A. Efros (2011)

task	Test on: / Train on:	SUN09	LabelMe	PASCAL	ImageNet	Caltech101	MSRC	Self	Mean others	Percent drop
“car” classification	SUN09	**28.2**	29.5	16.3	14.6	16.9	21.9	28.2	19.8	**30%**
	LabelMe	14.7	**34.0**	16.7	22.9	43.6	24.5	34.0	24.5	**28%**
	PASCAL	10.1	25.5	**35.2**	43.9	44.2	39.4	35.2	32.6	**7%**
	ImageNet	11.4	29.6	36.0	**57.4**	52.3	42.7	57.4	34.4	**40%**
	Caltech101	7.5	31.1	19.5	33.1	**96.9**	42.1	96.9	26.7	**73%**
	MSRC	9.3	27.0	24.9	32.6	40.3	**68.4**	68.4	26.8	**61%**
	Mean others	10.6	28.5	22.7	29.4	39.4	34.1	53.4	27.5	48%
“person” classification	SUN09	**16.1**	11.8	14.0	7.9	6.8	23.5	16.1	12.8	**20%**
	LabelMe	11.0	**26.6**	7.5	6.3	8.4	24.3	26.6	11.5	**57%**
	PASCAL	11.9	11.1	**20.7**	13.6	48.3	50.5	20.7	27.1	**-31%**
	ImageNet	8.9	11.1	11.8	**20.7**	76.7	61.0	20.7	33.9	**-63%**
	Caltech101	7.6	11.8	17.3	22.5	**99.6**	65.8	99.6	25.0	**75%**
	MSRC	9.4	15.5	15.3	15.3	93.4	**78.4**	78.4	29.8	**62%**
	Mean others	9.8	12.3	13.2	13.1	46.7	45.0	43.7	23.4	**47%**

Too big to fail?

Estimate No. 1: The number of meaningful/valid images on a 1200 by 1200 display is at least as high as 10^{400}.

Estimate No. 2: 10^{25} (greater than a trillion squared) is a very conservative lower bound to the number of all possible discernible images.

«These numbers suggest that it is impractical to construct training or testing sets of images that are dense in the set of all images unless the class of images is restricted.»

Theo Pavlidis

The Number of All Possible Meaningful or Discernible Pictures (2009)

The illusion of progress

«An apparent superiority in classification accuracy, obtained in "laboratory conditions," may not translate to a superiority in real-world conditions and, in particular, the apparent superiority of highly sophisticated methods may be illusory, with simple methods often being equally effective or even superior.»

David J. Hand

Classifier Technology and the Illusion of Progress (2006)

Belief in the "law of small numbers"

«People's intuitions about random sampling appear to satisfy the law of small numbers, which asserts that the law of large numbers applies to small numbers as well.»

Amos Tversky and Daniel Kahneman

Belief in the Law of Small Numbers (1971)

Belief in the "law of small numbers"

The believer in the law of small numbers practices science as follows:

1 He gambles his hypotheses on small samples without realizing that the odds against him are unreasonably high. **He overestimates power.**

2 He has undue confidence in early trends and in the stability of observed patterns. **He overestimates significance.**

3 In evaluating replications, he has unreasonably high expectations about the replicability of significant results. **He underestimates the breadth of confidence intervals.**

4 He rarely attributes a deviation of results from expectations to sampling variability, because he finds a causal "explanation" for any discrepancy. Thus, **he has little opportunity to recognize sampling variation in action**.

His belief in the law of small numbers, therefore, will forever remain intact.

From: A. Tversky and D. Kahneman, *Belief in the Law of Small Numbers* (1971)

Bias and social justice

But ML is increasingly being used in several "social" domains:

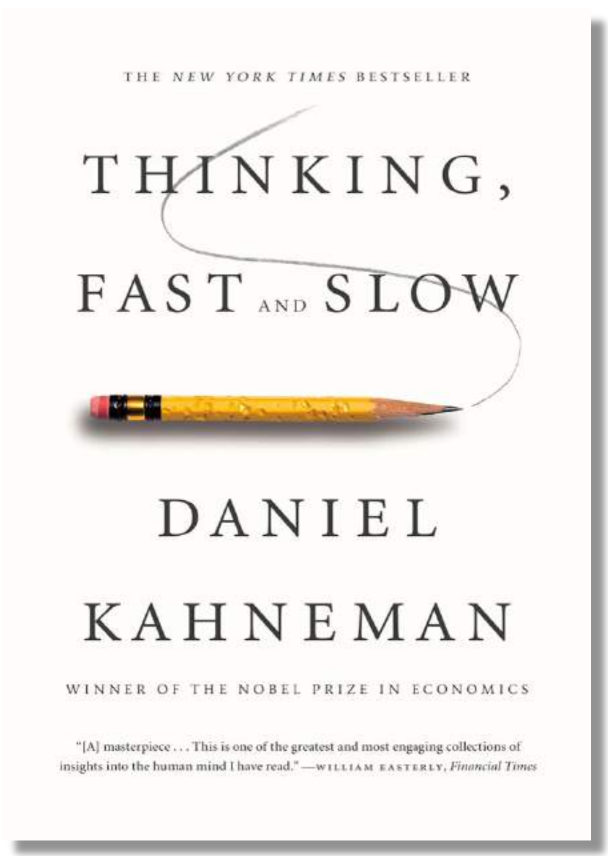
- Recruiting: Screening job applications
- Banking: Credit ratings / loan approvals
- Judiciary: Recidivism risk assessments
- Journalism: News recommender systems
- …

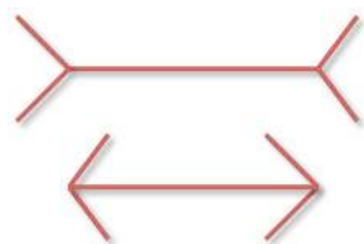
Sources of potential social discrimination:

- Social biases of people collecting the training sets
- Sample size disparity
- Feature selection
- Optimization criteria
- …

M. Hardt, *How big data is unfair.*
Understanding unintended sources of unfairness in data driven decision making (2014)

Bias in humans and machines





Algorithms are biased, but humans also are ...

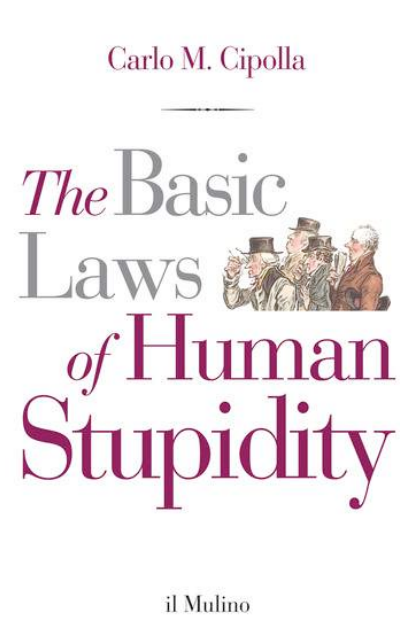
When should we trust humans and when algorithms?

Stupidity (according to C. M. Cipolla)

Third (and golden) basic law of stupidity

A stupid person is a person who causes losses to another person or to a group of persons while himself deriving no gain and even possibly incurring losses.

Carlo M. Cipolla
The Basic Laws of Human Stupidity (2011)

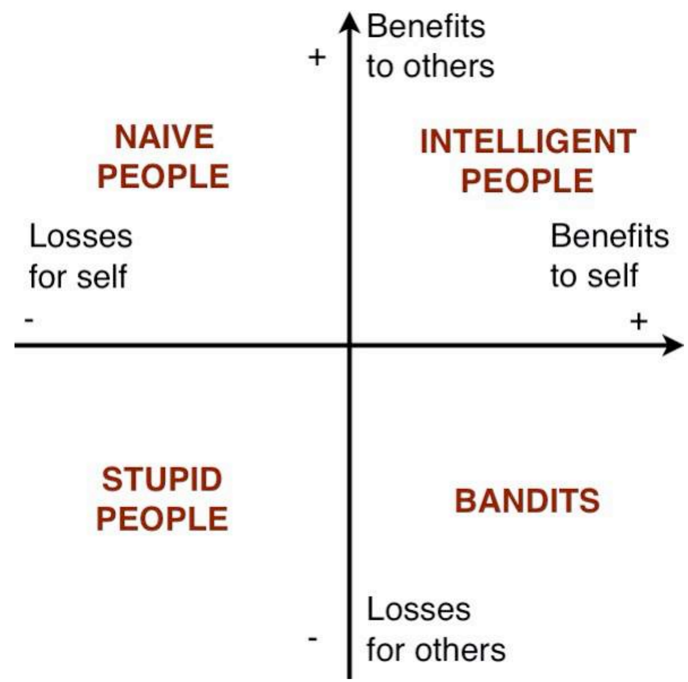


Stupidity
(according to C. M. Cipolla)

Third (and golden) basic law of stupidity

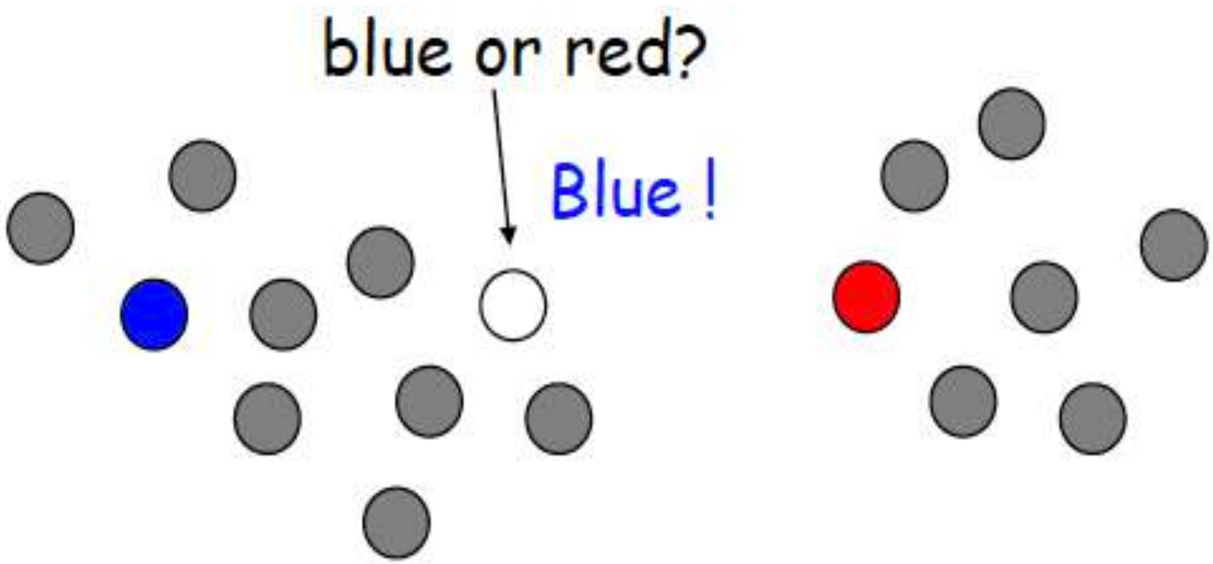
A stupid person is a person who causes losses to another person or to a group of persons while himself deriving no gain and even possibly incurring losses.

Carlo M. Cipolla
The Basic Laws of Human Stupidity (2011)



The smoothness assumption

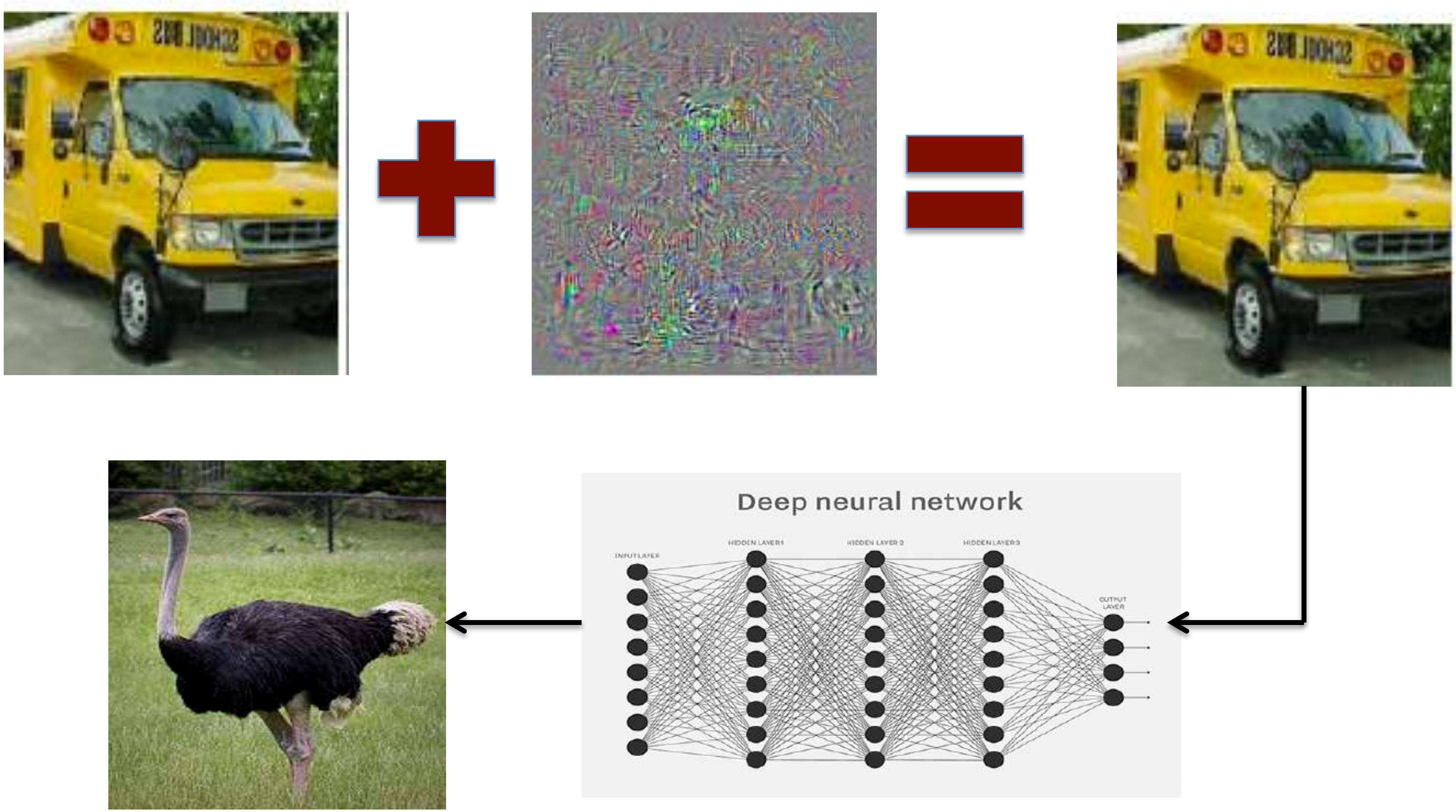
Points close to each other are more likely to share the same label



What about the performance of deep networks on image data that have been modified only slightly?

Courtesy Fabio Roli

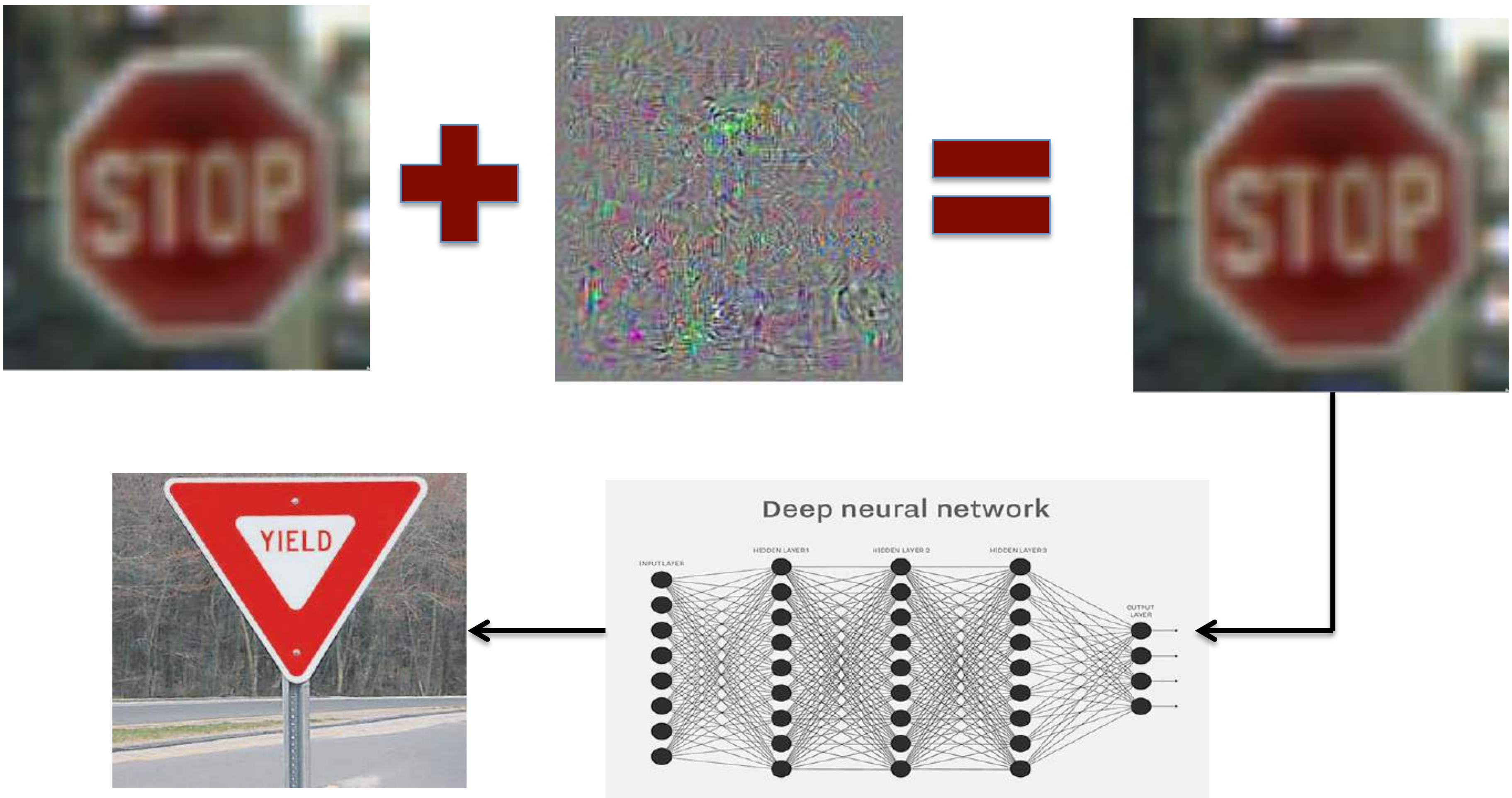
High accuracy = high robustness?



Szegedy et al., *Intriguing properties of neural networks* (2014)

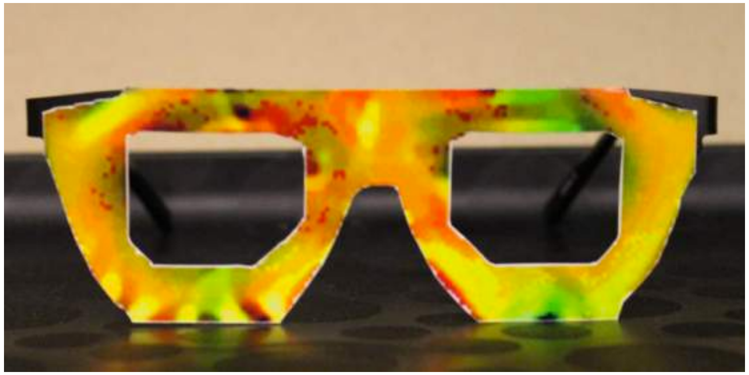
Courtesy Fabio Roli

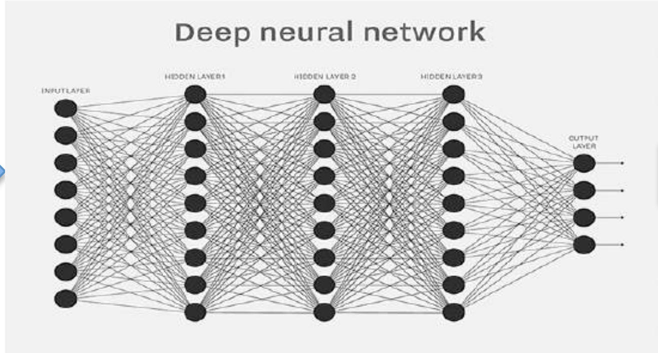
What if …



Courtesy Fabio Roli

Fashionable glasses

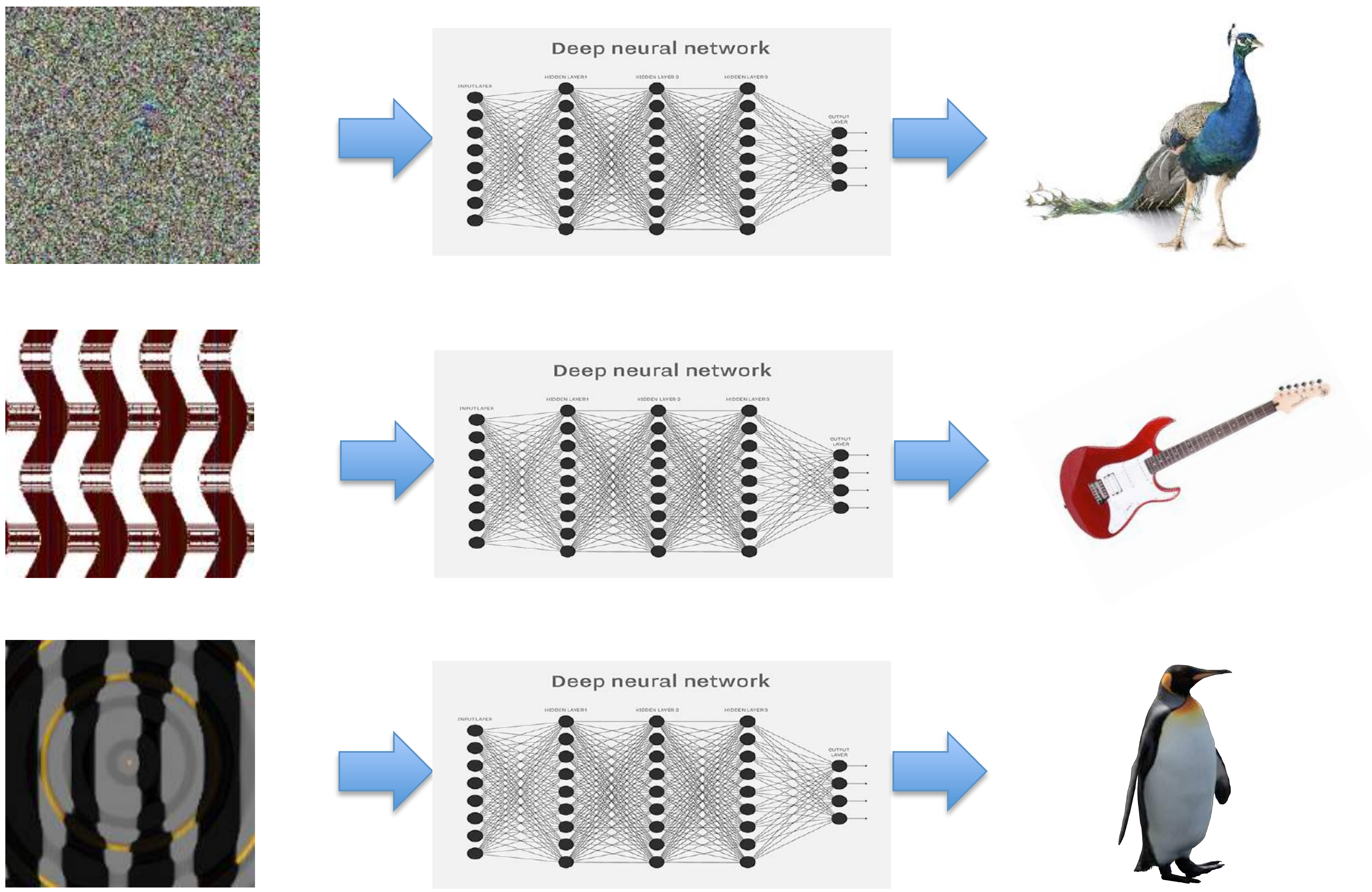




M. Sharif et al., *Accessorize to a crime: Real and stealthy attacks on state-of-the-art face recognition* (2016)

Courtesy Fabio Roli

What does a machine see here?



A. Nguyen et al., *Deep neural networks are easily fooled: High confidence predictions for unrecognizable images* (2015)

Different similarity spaces?

«Different creatures will have different similarity-spaces, hence different ways of grouping things […]

Such perceived similarities (or, for what matter, failure to perceive similarities) will manifest themselves in behavior and are a crucial part of explaining what is distinctive in each individual creature's way of apprehending the world.»

José Luis Bermùdez
Thinking Without Words (2003)

Different similarity spaces?

Great Dialogue, Karel Nepras, Museum of Modern and Contemporary Art, Prague

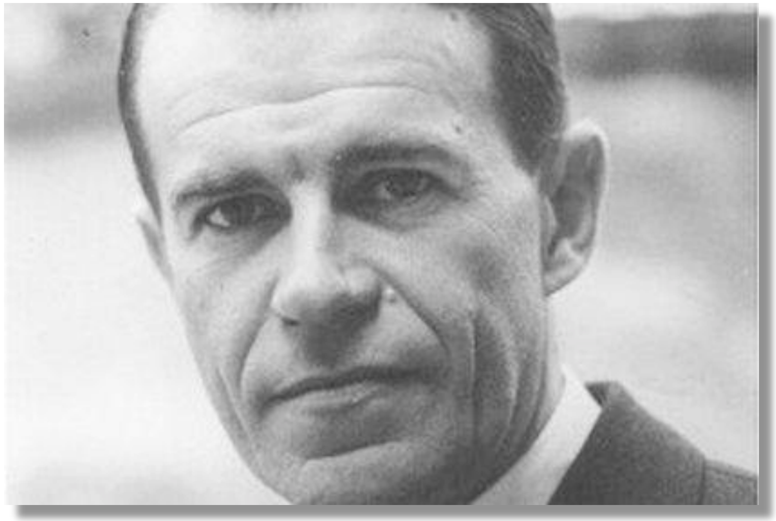
Courtesy Sven Dickinson

Cipolla, again

Fifth basic law of stupidity
A stupid person is the most dangerous type of person.

Corollary
A stupid person is more dangerous than a bandit.

Carlo M. Cipolla
The Fundamental Laws of Human Stupidity (2011)



By way of conclusion

On being impertinent

«That is the essence of science: ask an impertinent question,
and you are on the way to the pertinent answer.»

Jacob Bronowski
The Ascent of Man (1973)

Philosophy and machine learning

Philosophical topics of interest to the machine learning community (not treated, or just touched upon, today):

- ✓ Causality (Pearl, Spirtes, Glymour, Schölkopf, …)
- ✓ Complexity and information (Kolmogorov, Solomonoff, Hutter, …)
- ✓ Model selection
- ✓ Emergentism
- ✓ Scientific method
- ✓ Abstraction and categorization
- ✓ Decision theory
- ✓ Philosophy of technology
- ✓ Ethics

and many more …

If you want to know more ...

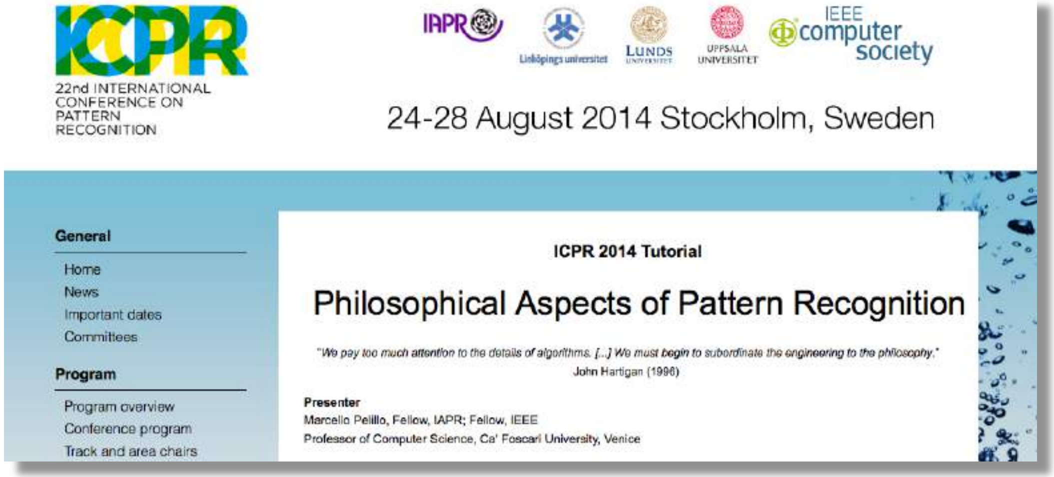
http://www.dsi.unive.it/PhiMaLe2011/

Special issue on
"Philosophical aspects of pattern recognition"

Vol. 64, October 2015

Guest editor: M. Pelillo

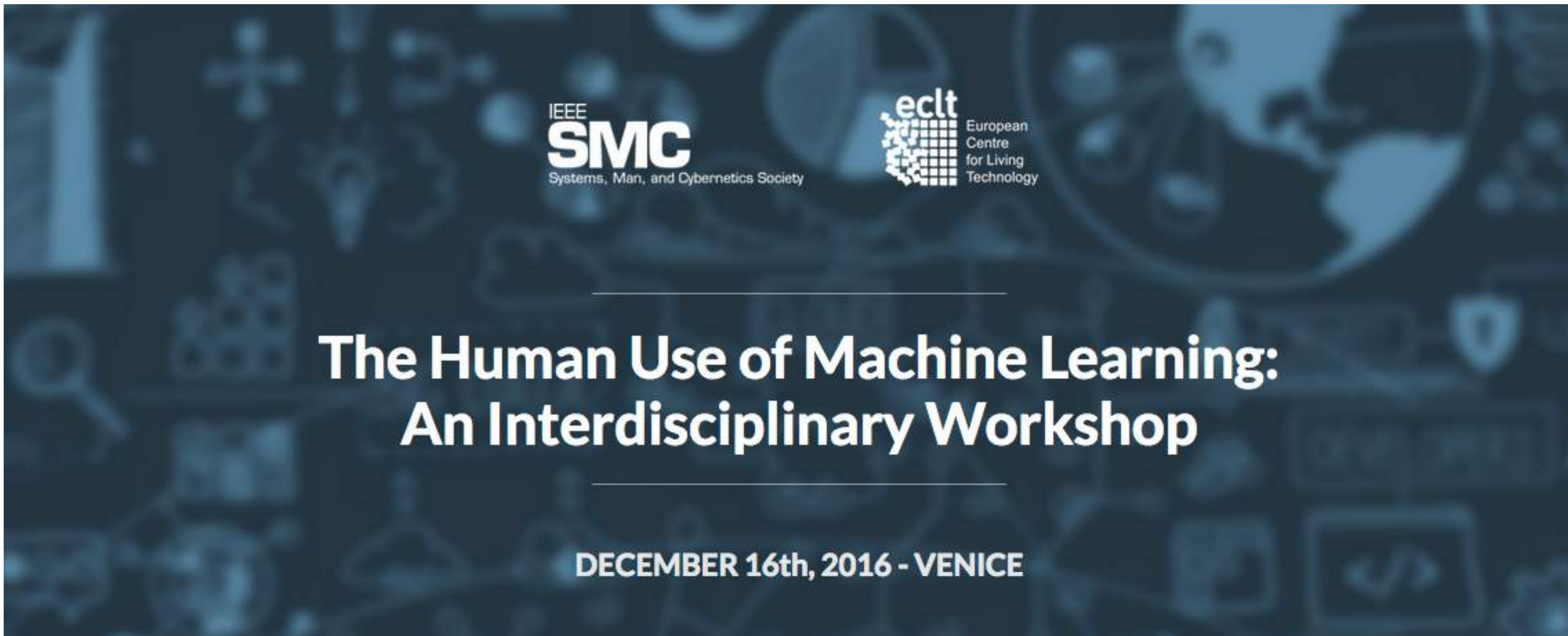
If you want to know more ...



ICPR 2016
23rd INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION.
Cancun, Mexico 4-8 December

ECML–PKDD
2016 RIVA DEL GARDA

If you want to known more …



http://www.dsi.unive.it/HUML2016

Welcome to the AI4EU initiative

The AI4EU proposal adressing [ICT-26 2018 H2020 call] has successfully passed the evalution process.

The project should start early this autumn

https://ai4eu.org


Thanks!