An Overview of Artificial Intelligence

Professor Barry O'Sullivan, PhD(NUI), FEurAI, MRIA
European Artificial Intelligence Association &
Department of Computer Science, University College Cork

Who am I?

@BarryOSullivan

- Professor (Constraint Programming) at UCC.
- Founding Director, Insight Centre for Data Analytics.
- Deputy President, European Artificial Intelligence Association.
- Editor, Springer Book Series
  "Artificial Intelligence: Foundations, Theory, and Algorithms"
  (with Prof. M.Wooldridge, Oxford University).
- Science Foundation Ireland Researcher of the Year 2016.
- Member, Royal Irish Academy.
- CEO and Founder, AI Machina.
- AI Expert-in-Residene, SOSV - http://www.sosv.com

Insight

http://www.insight-centre.org

Big Data

VS

Artificial Intelligence

Science Foundation Ireland

sfi

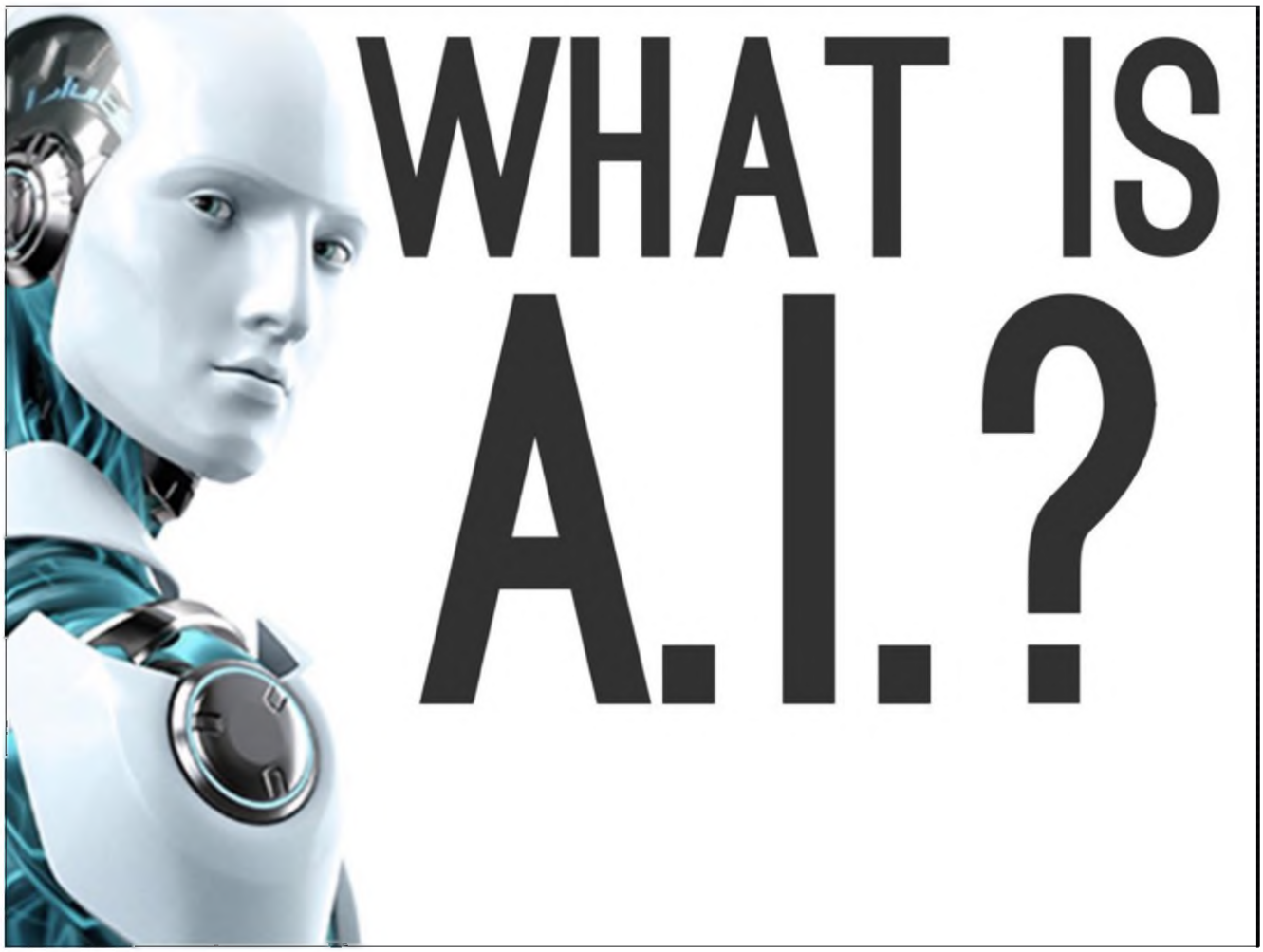
For what's next

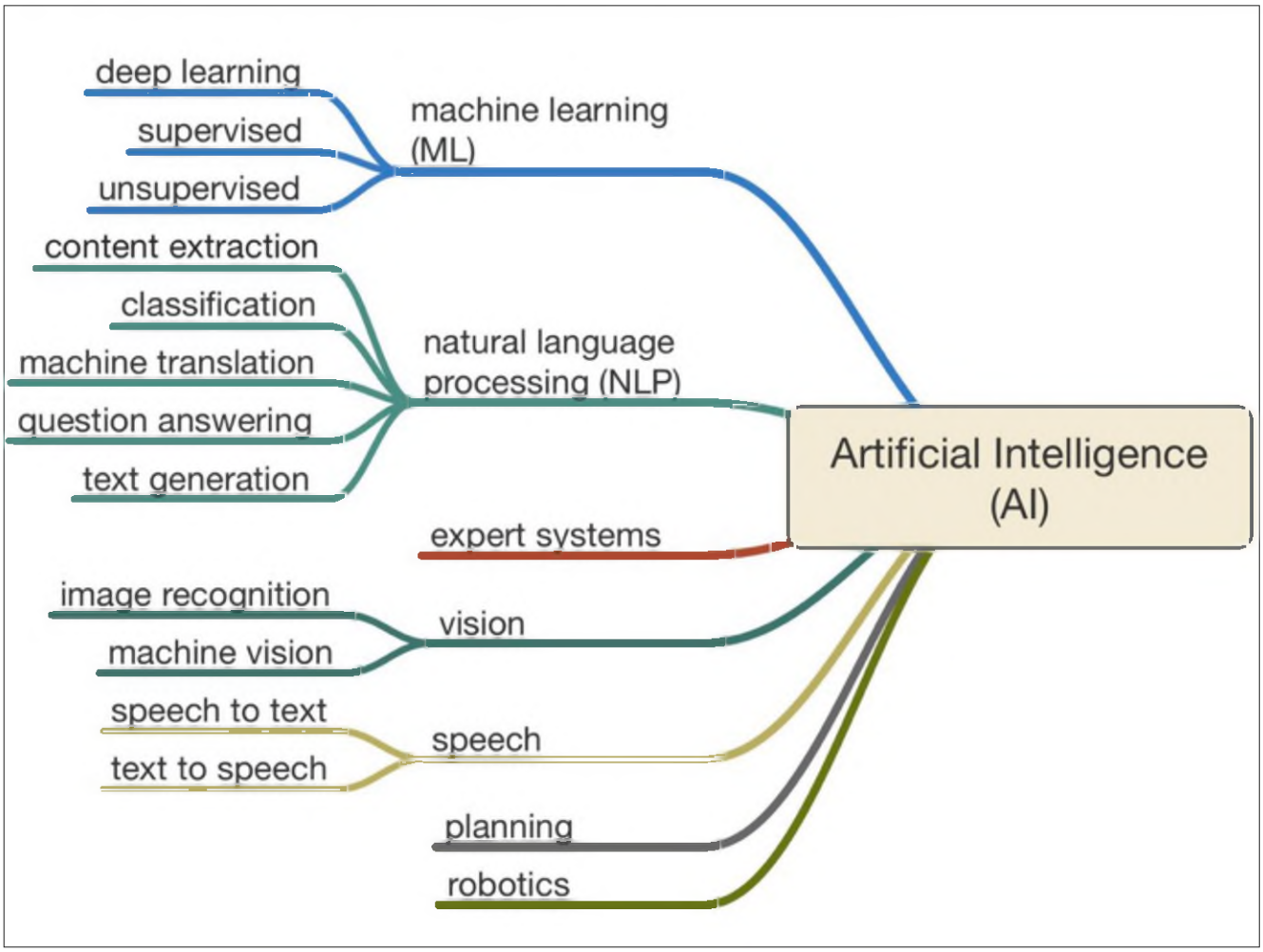
http://www.sfi.ie

EurAi

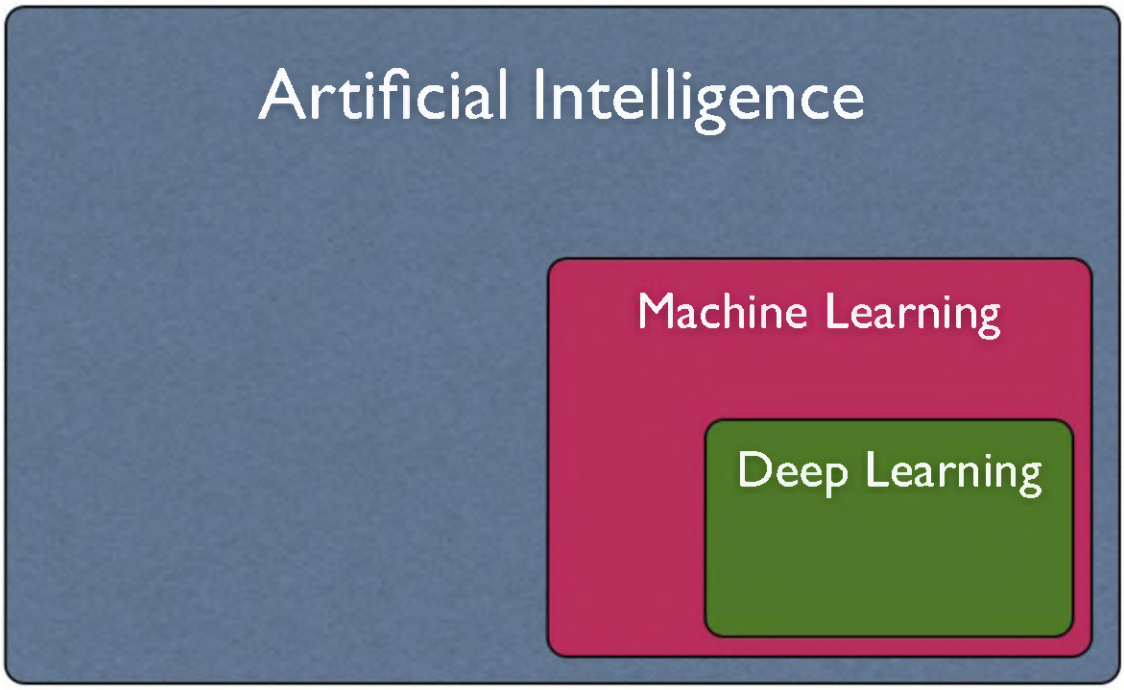
Largest AI association
in the world:
~4500 individual members
~30 countries

Deputy President (since August 2016)
European Association for Artificial Intelligence


WHAT IS
A.I.?


deep learning
supervised
unsupervised
machine learning (ML)
content extraction
classification
machine translation
question answering
text generation
natural language processing (NLP)
Artificial Intelligence (AI)
expert systems
image recognition
machine vision
vision
speech to text
text to speech
speech
planning
robotics

AI is not Machine Learning is not Deep Learning



Textbook Machine Learning

Data

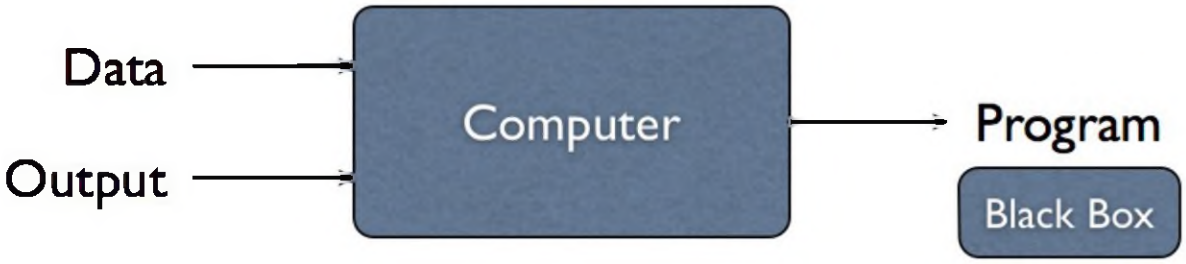
Algorithm

Model

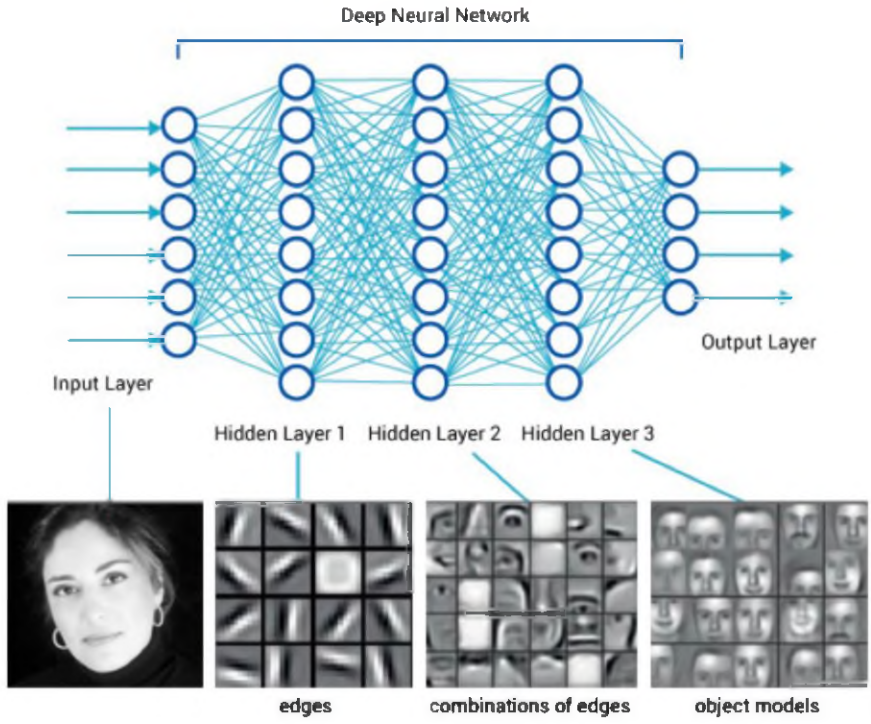
$$f(\mathbf{x})$$

Traditional Programming

Machine Learning



Deep Learning



A. M. Turing (1950) Computing Machinery and Intelligence. *Mind 49:* 433-460.

COMPUTING MACHINERY AND INTELLIGENCE

By A. M. Turing

1. The Imitation Game

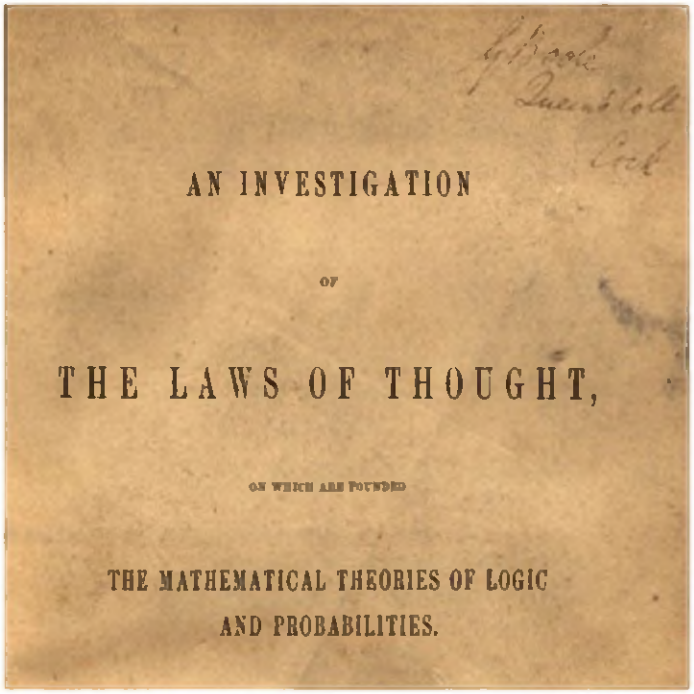
I propose to consider the question, "Can machines think?" This should begin with definitions of the meaning of the terms "machine" and "think." The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous, If the meaning of the words "machine" and "think" are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, "Can machines think?" is to be sought in a statistical survey such as a Gallup poll. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.

The new form of the problem can be described in terms of a game which we call the 'imitation game." It is played with three people, a man (A), a woman (B), and an interrogator (C) who may be of either sex. The interrogator stays in a room apart front the other two. The object of the game for the interrogator is to determine which of the other two is the man and which is the woman. He knows them by labels X and Y, and at the end of the game he says either "X is A and Y is B" or "X is B and Y is A." The interrogator is allowed to put questions to A and B thus:

George Boole

Father of Boolean Algebra

First Professor of Mathematics, UCC


AN INVESTIGATION

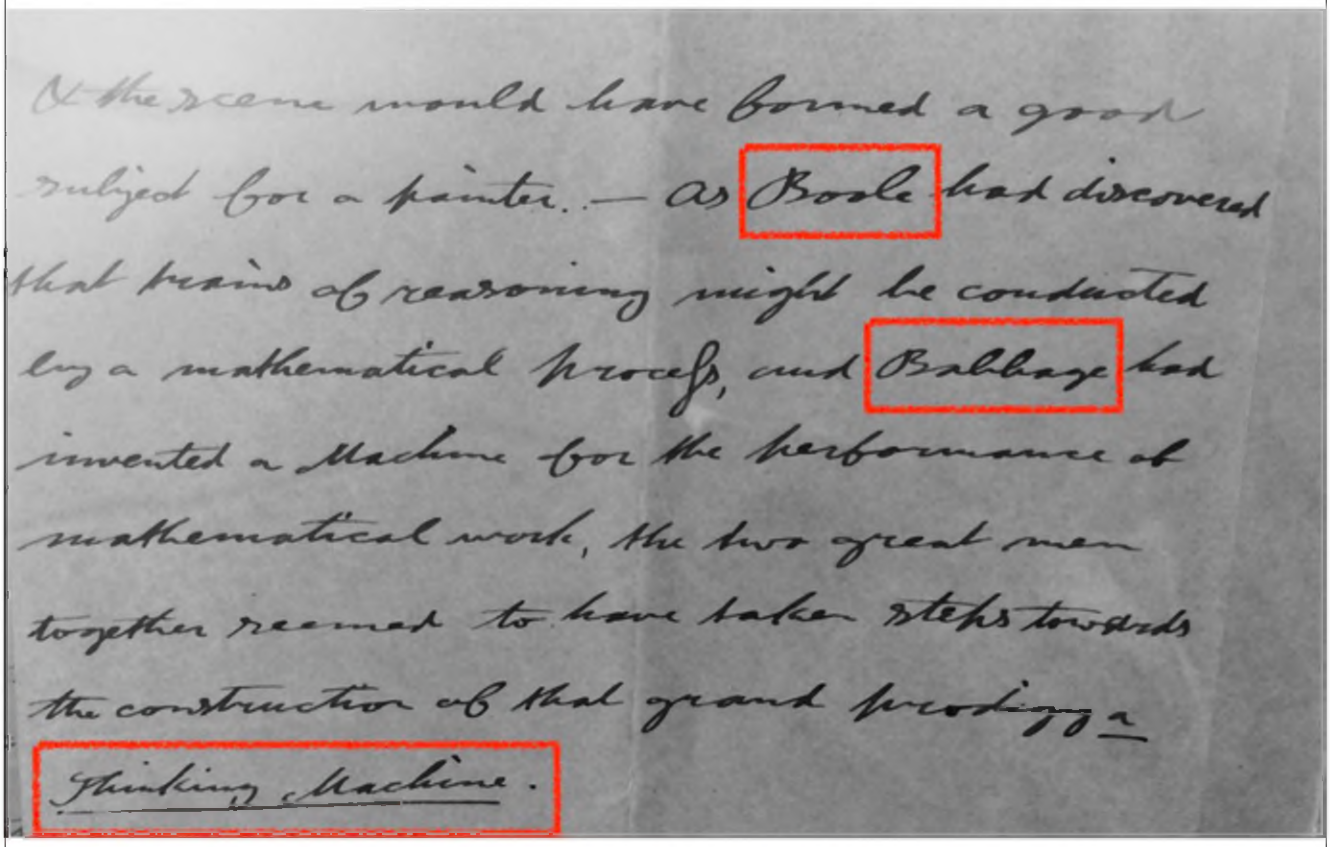
OF

THE LAWS OF THOUGHT,

ON WHICH ARE FOUNDED

THE MATHEMATICAL THEORIES OF LOGIC

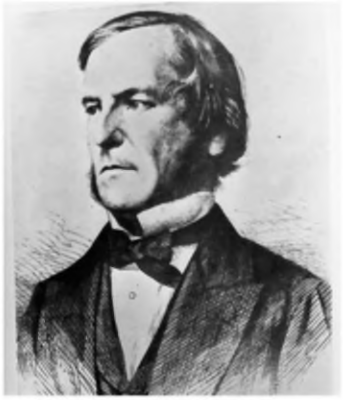
AND PROBABILITIES.



of the scene would have formed a good subject for a painter. — As Boole had discovered that trains of reasoning might be conducted by a mathematical process, and Babbage had invented a Machine for the performance of mathematical work, the two great men together seemed to have taken steps towards the construction of that grand prodigy a Thinking Machine.

Letter from J.Hill to Mary Boole, 1865 (just after the death of George Boole)

From Boole to Deep Learning



George Boole

Geoff Hinton
University of Toronto & Google
(great-great-grandson)

Boole, the Hintons & China

Voynich Manuscript

Ethel Boole Voynich

The Gadfly

Joan Hinton

寒春

Great-grandaughter of Boole, Aunt to Geoff Hinton

Sun Yat-Sen
孫中山 / 孫逸仙
other names

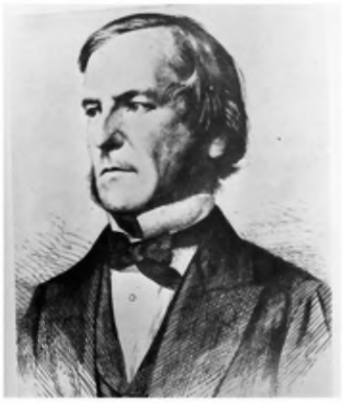
Provisional President of the Republic of China

Soong Ching-ling
宋庆龄
宋慶齡

Honorary President of the People's Republic of China

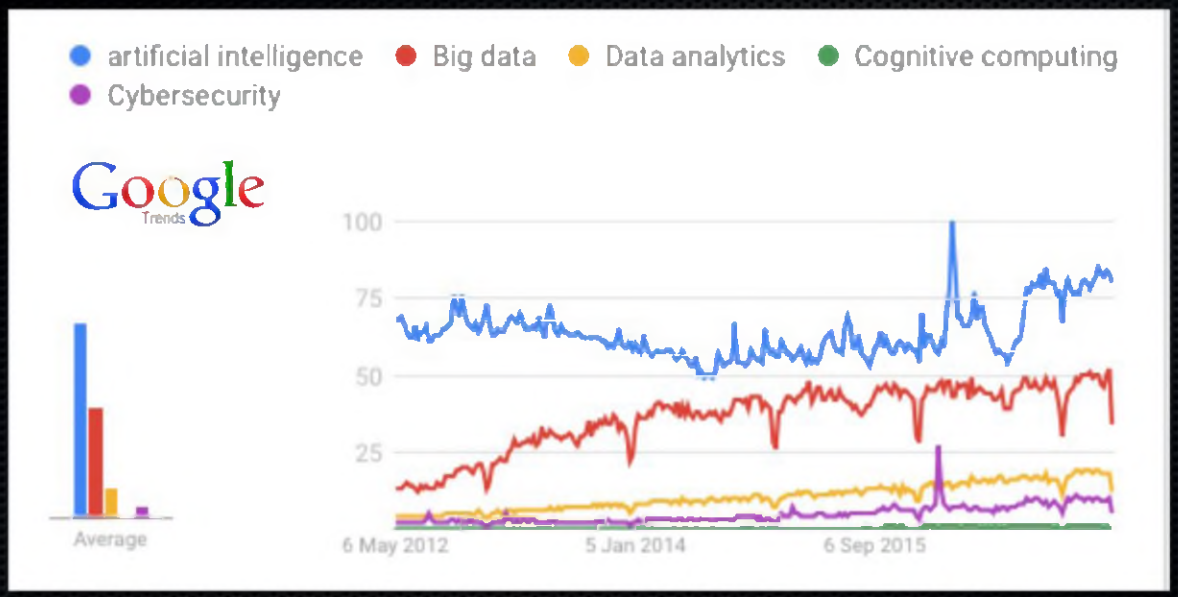
Joan Hinton (second left), Sidney Shapiro (second right) and Eunice Moe Brock (first right) receive "You Bring Charm to China" awards presented on Saturday in Beijing. Wang Jing

From Boole to Deep Learning



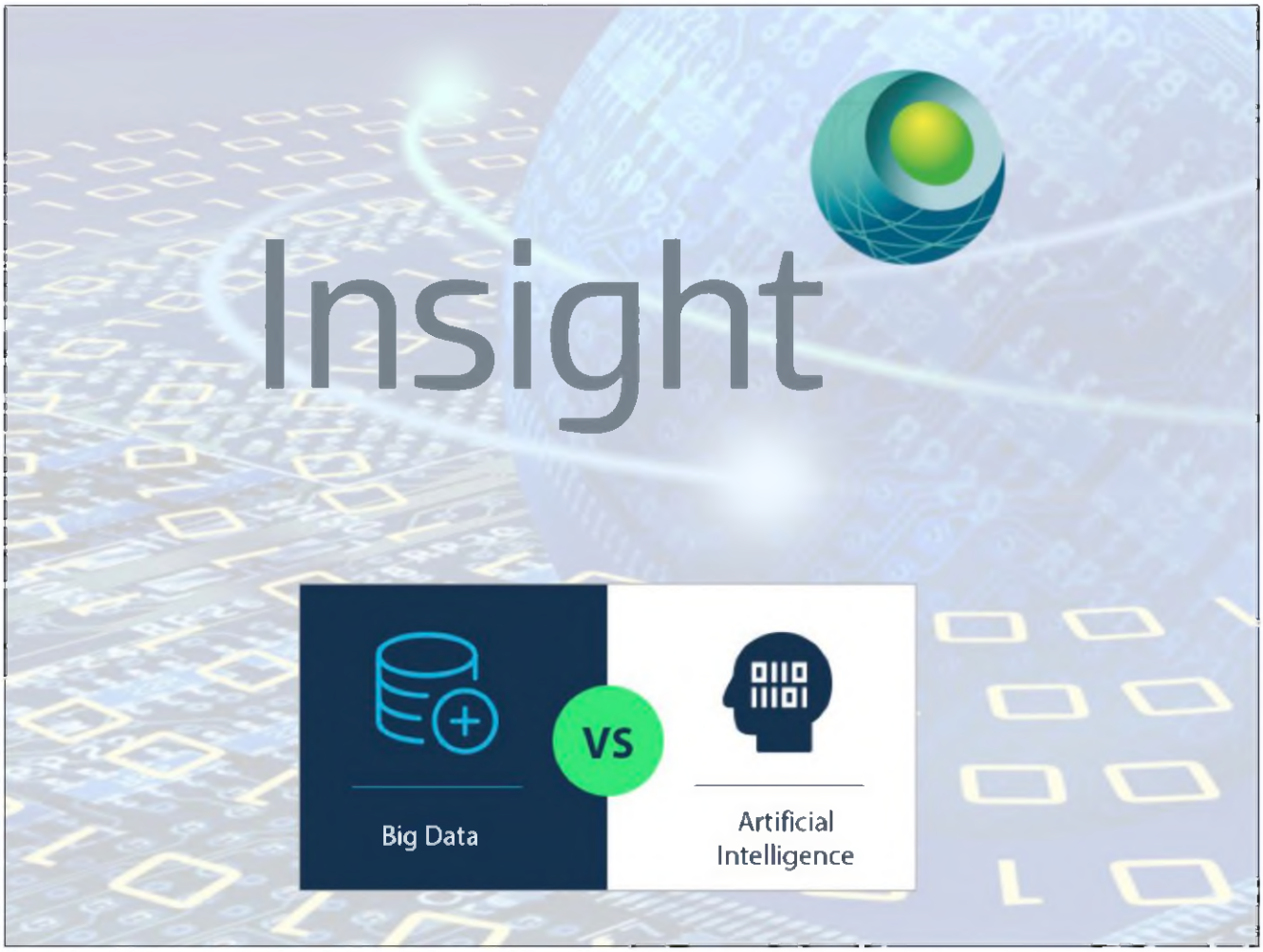
George Boole

Geoff Hinton
Univerity of Toronto & Google
(great-great-grandson)



Interest in AI worldwide. Past 5 years.

AI has dwarfed interested in big data, data analytics, cognitive computing, and cybersecurity combined.


Insight
Big Data
VS
Artificial Intelligence


GARRY
KASPAROV
1996
DEEP
BLUE


Ali Eslami & Phil Laak played the
first AI system for Texas Hold'em.
gettyimages
AFP
2007

engadget 3 related articles

Latest in Culture

How to be a human being in the comments: A refresher

05.01.17

Portland wants to get driverless cars on its roads this year

04.19.17

Libratus, the poker-playing AI, destroyed its four human rivals

Sorry, fellow humans: It wasn't even close.

Chris Velazco, @chrisvelazco
01.31.17 in Robots

14 Comments

1149 Shares


2016


Challenge Match
8 - 15 March 2016
LEE SEDOL
00:01:00
2016

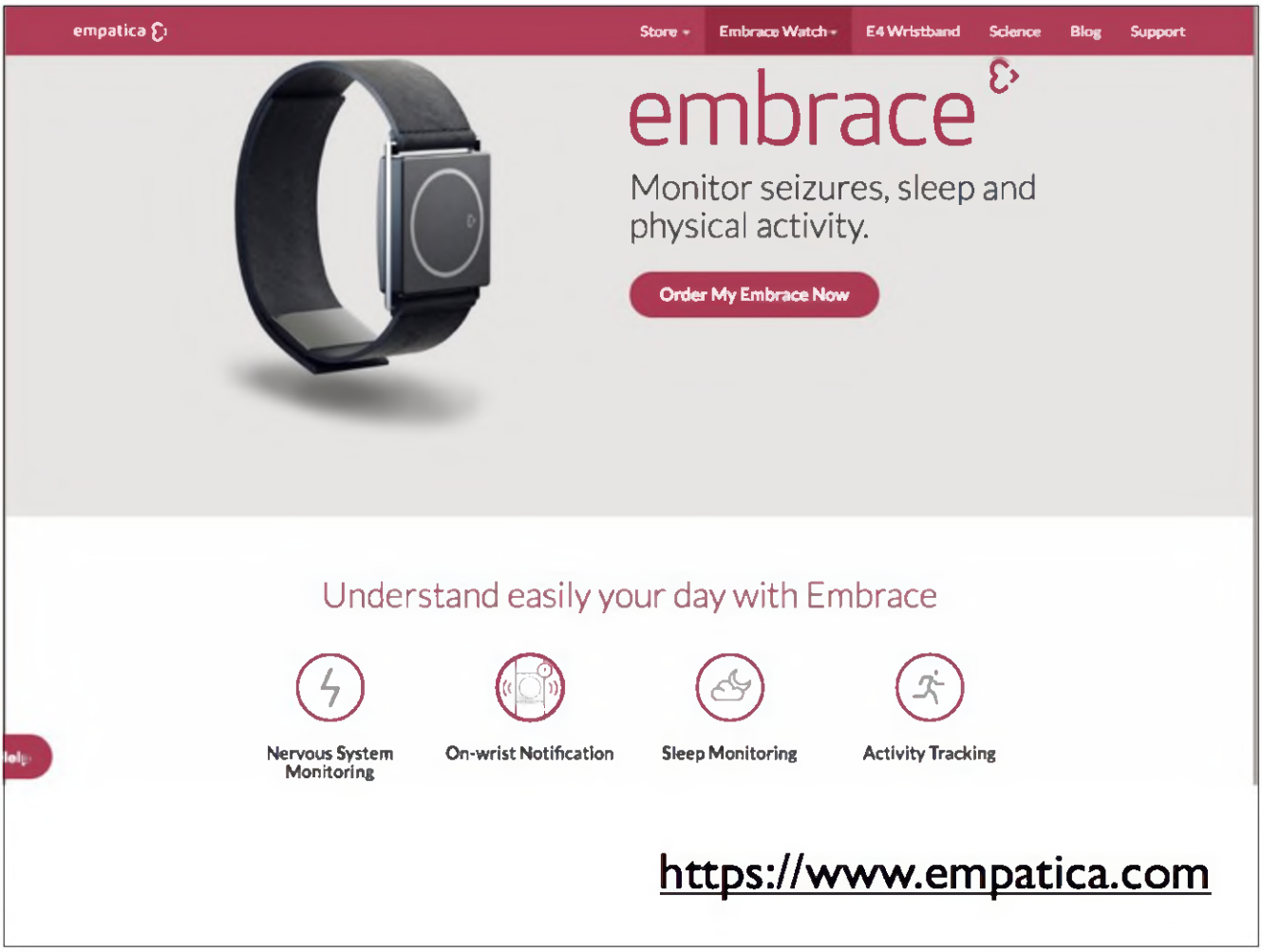
2005, >130 miles, desert

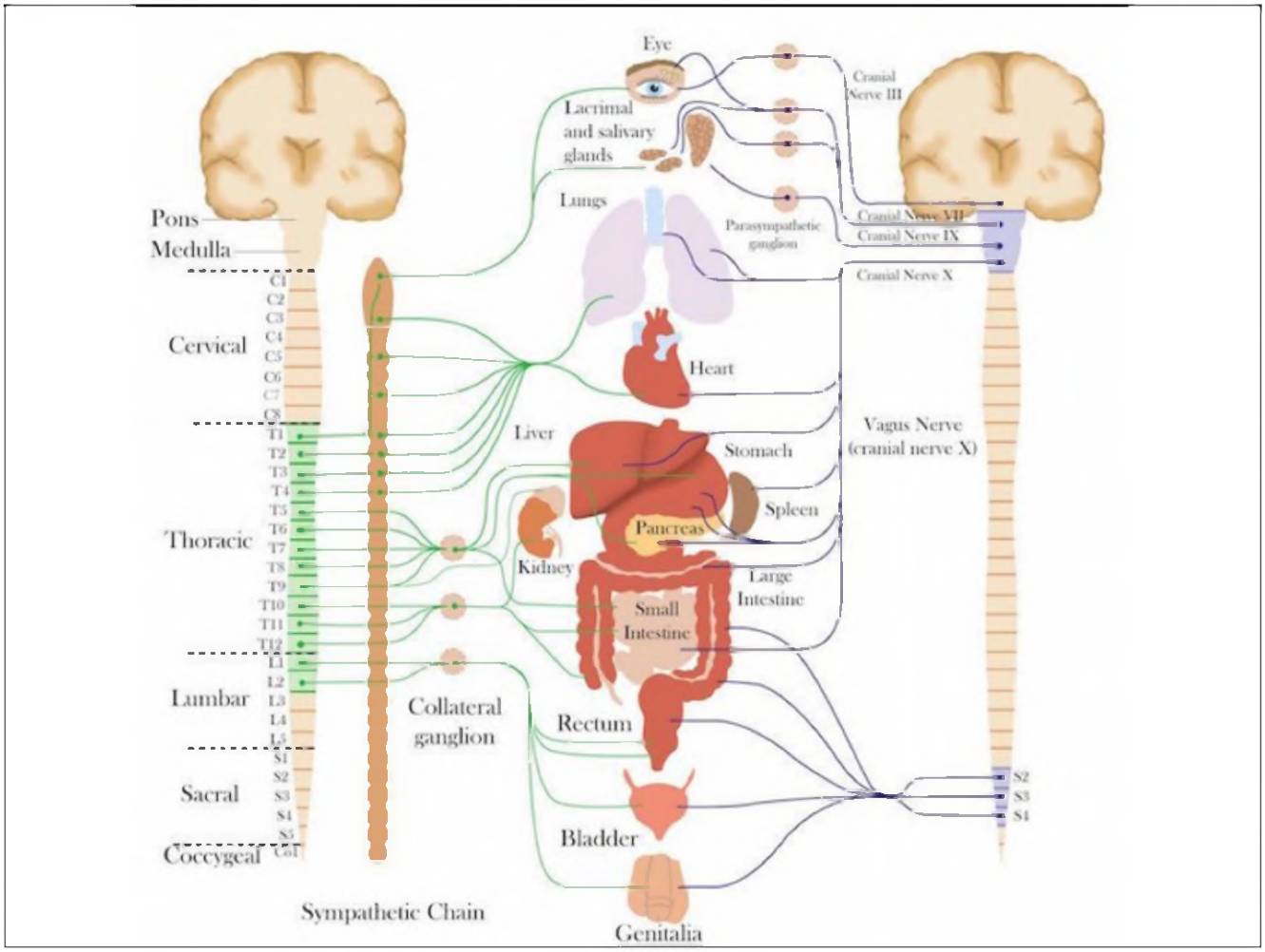
2007, >50 miles, urban

20??, on-demand, autonomous cars

Explanation in AI/ML is Important :)

"Does your car have any idea why my car pulled it over?"


empatica
Store
Embrace Watch
E4 Wristband
Science
Blog
Support
embrace
Monitor seizures, sleep and physical activity.
Order My Embrace Now
Understand easily your day with Embrace
Nervous System Monitoring
On-wrist Notification
Sleep Monitoring
Activity Tracking
https://www.empatica.com

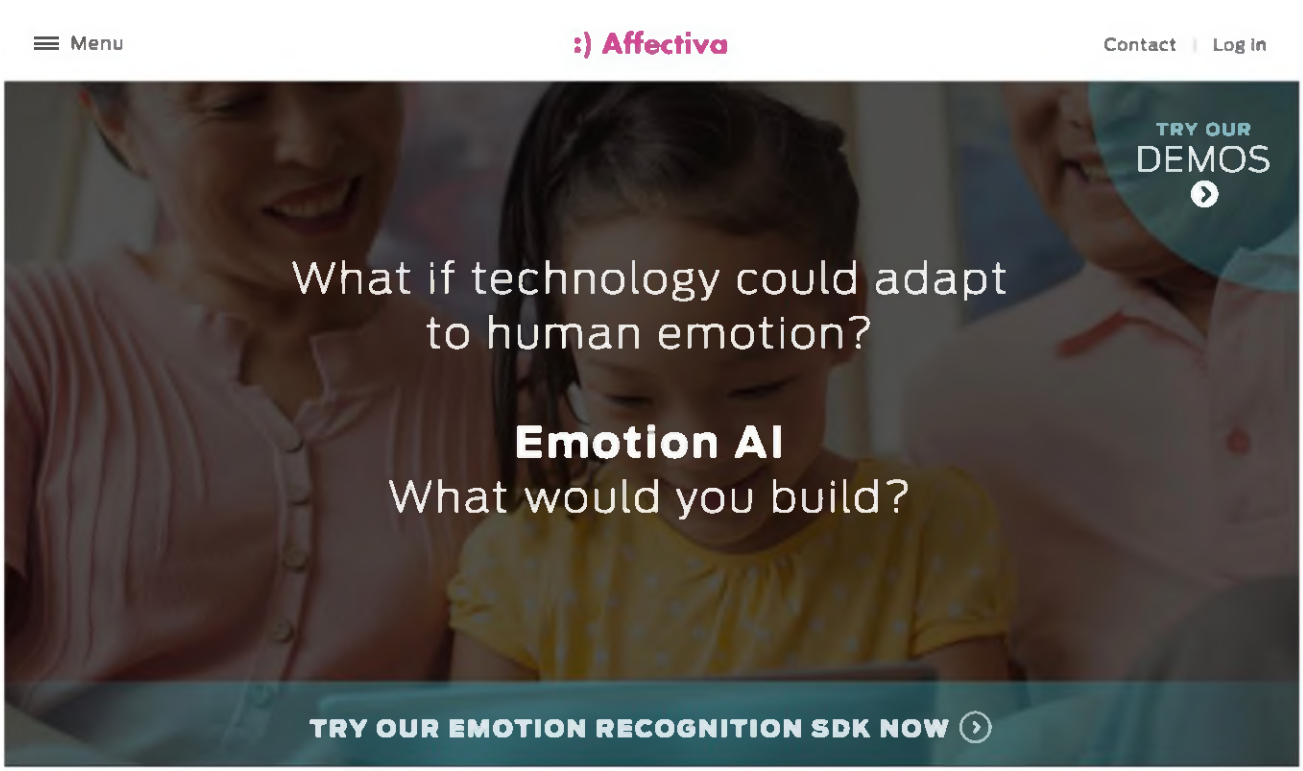

Eye
Cranial Nerve III
Lacrimal and salivary glands
Lungs
Parasympathetic ganglion
Cranial Nerve VII
Cranial Nerve IX
Cranial Nerve X
Pons
Medulla
Cervical
Heart
Liver
Stomach
Vagus Nerve (cranial nerve X)
Spleen
Pancreas
Thoracic
Kidney
Large Intestine
Small Intestine
Lumbar
Collateral ganglion
Rectum
Sacral
Bladder
Coccygeal
Sympathetic Chain
Genitalia

Empathica - Electrodemeral Activity

Projects using Electrodermal Activity

Some of the research applications that Electrodermal Activity monitoring can enable

Sleep disorders and insomnia	Parkinson's disease	Depression	Epilepsy
Addiction	Anxiety disorder	Anesthesia	Schizophrenia
Hypoglycemia	Phobias (and desensitization)	Hot flashes	Dementia
Affect dysregulation	Drug trials	Psychiatric counseling	Alcoholism
Sensory Processing Modulation	Thyroid dysfunction	Bipolar disorder	Sexual Dysfunction
Oppositional defiant disorder	Endocrine disorders	Stroke	Dysthymia
Alzheimer's	Quadriplegia (or paraplegia)	Dialysis	Dermatology
Pain management	Attention deficit hyperactivity disorder	Cystic fibrosis	Sociopathy
Autism spectrum disorders	Post-traumatic stress disorder	Temporary Paralysis	Diabetes



http://www.affectiva.com/

100 Year Study on AI

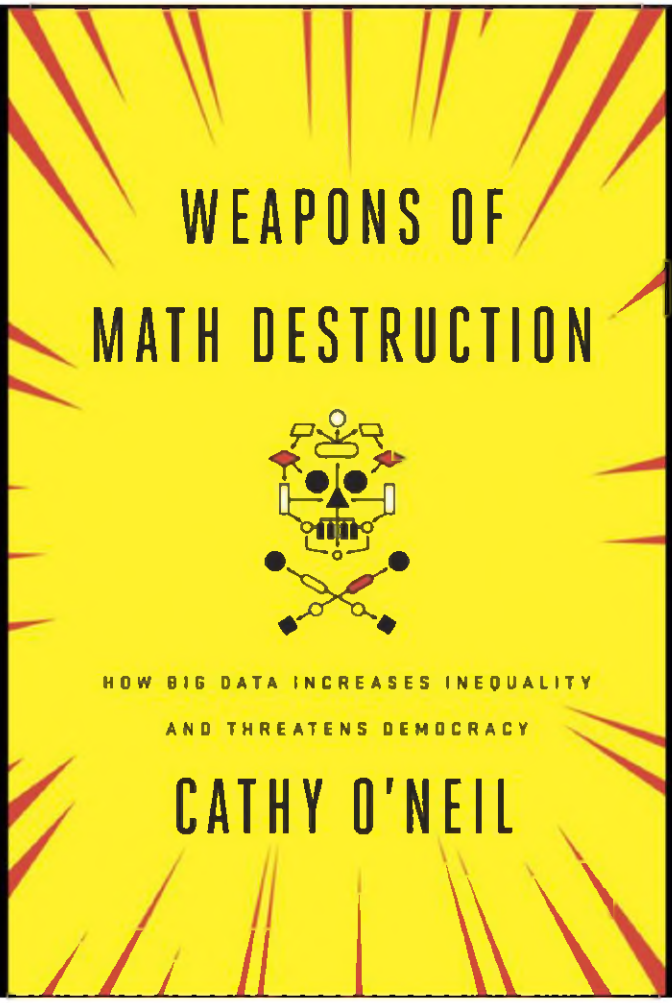
AI revolution has been led by access to compute power, massive data collection, and deep learning

Big Trend	Example Focus
Large-scale Machine Learning	Large data sets, Complexity
Deep Learning	Object recognition, Perception, NLP
Reinforcement Learning	Learning to act in the real world
Robotics	Interacting with the real world
Computer Vision	Visual Classification, Captioning
Natural Language Processing	Voice query, Realtime translation
Collaborative Systems	Human-Machine, Augmentation
Human Computation	Harness human-level intelligence
Algorithmic Game Theory	Econ/social impact & incentives
Internet of Things (IOT)	Connected sensing & reasoning
Neuromorphic Computing	Biological neural net-inspired

Domain	Focus
Transportation	Smart car, self-driving, on-demand
Home/Service Robots	Smart home systems, home robotics
Healthcare	Analytics, robotics, mobile, elder care
Education	Teacher robots, ITS, online, analytics
Sustainable Communities	Achieving societal good
Public Safety/Security	Training, policing, cybersecurity
Employment & Work	Impact on jobs and work
Entertainment	Interaction, personalised, engaging

AI: Bias-in-bias-out

But isn't AI technology independent, cool, mathematical, and calculating? No, it isn't.

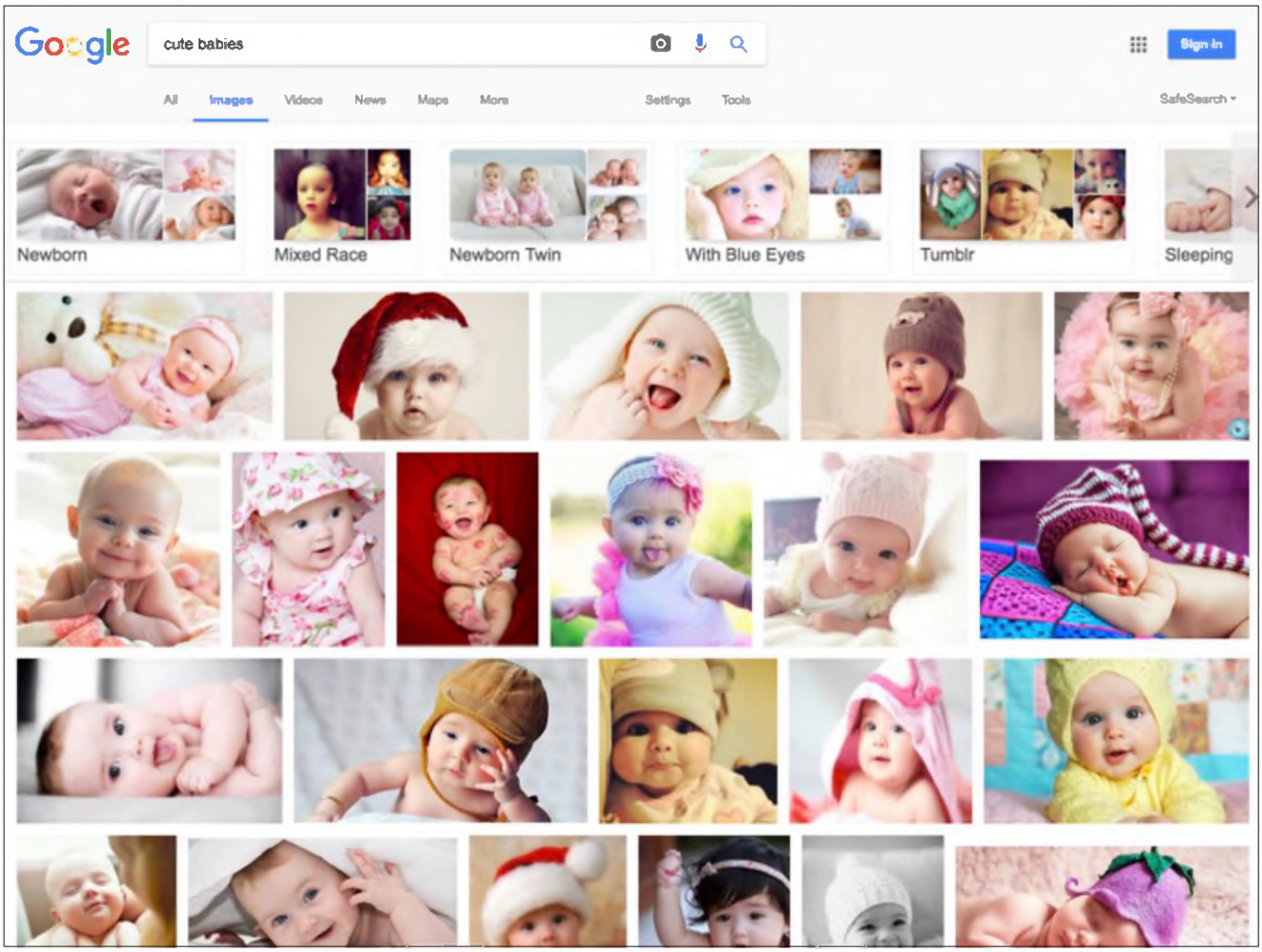


Cathy's blog:

https://mathbabe.org

Bias and Artificial Intelligence

- Data-driven bias
- Interaction bias
- Personalisation
- Similarity bias
- We're biased
- Language is biased



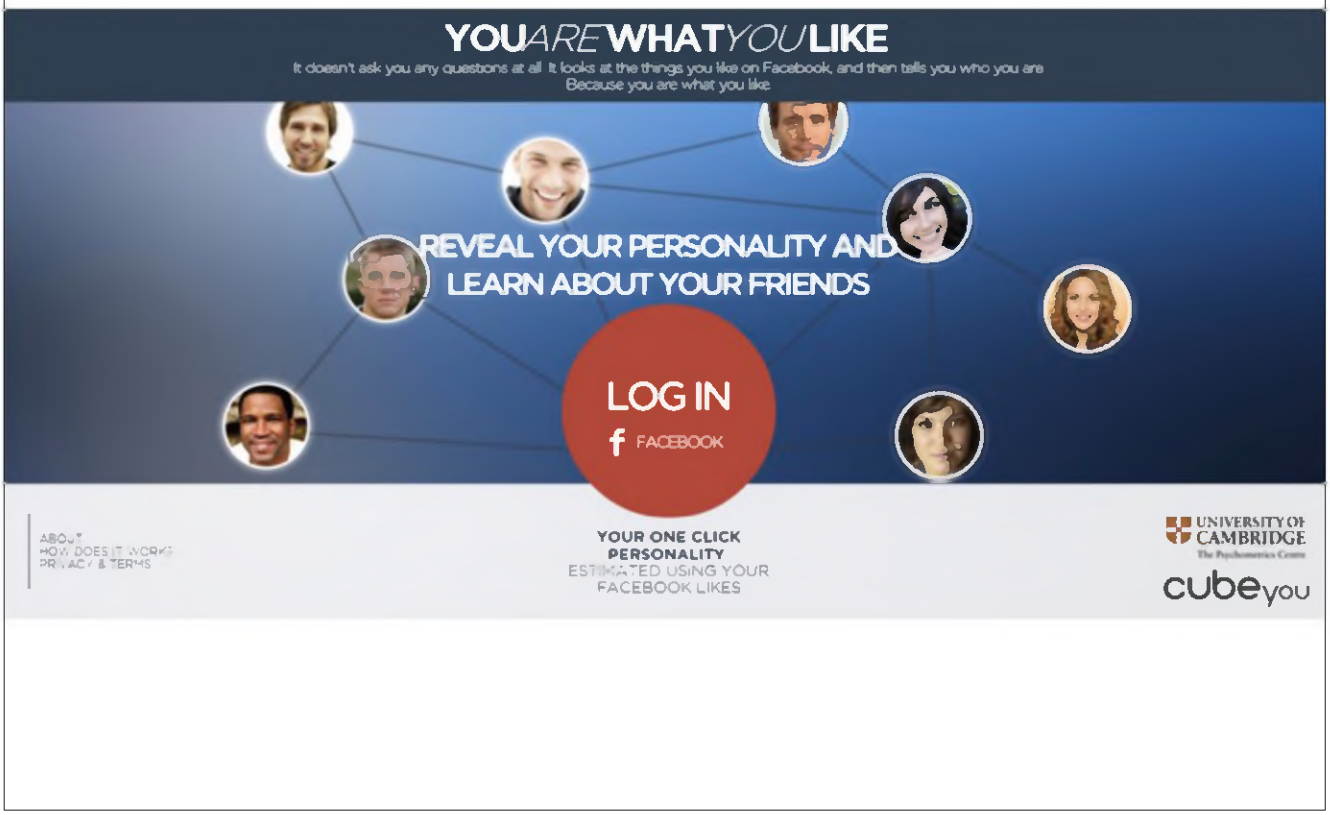
TL;DR MICROSOFT WEB

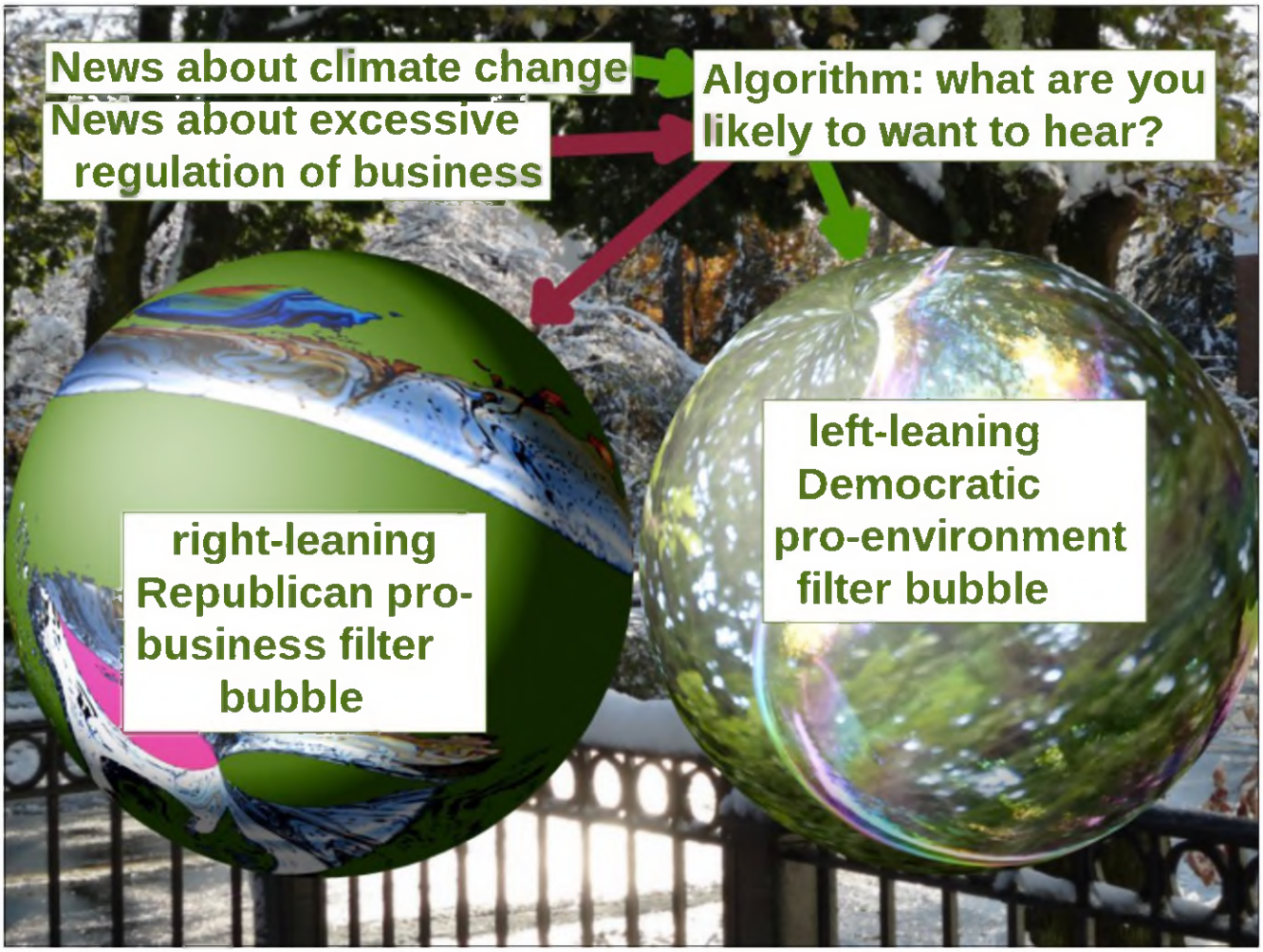
Twitter taught Microsoft's AI chatbot to be a racist asshole in less than a day

by James Vincent | @jvincent | Mar 24, 2016, 6:43am EDT

SHARE TWEET LINKEDIN

CASA


YOUARE WHAT YOU LIKE
It doesn't ask you any questions at all. It looks at the things you like on Facebook, and then tells you who you are
Because you are what you like
REVEAL YOUR PERSONALITY AND
LEARN ABOUT YOUR FRIENDS
LOG IN
FACEBOOK
ABOUT
HOW DOES IT WORK?
PRIVACY & TERMS
YOUR ONE CLICK
PERSONALITY
ESTIMATED USING YOUR
FACEBOOK LIKES
UNIVERSITY OF
CAMBRIDGE
cubeyou


News about climate change
News about excessive regulation of business
Algorithm: what are you likely to want to hear?
left-leaning Democratic pro-environment filter bubble
right-leaning Republican pro-business filter bubble

Science AAAS

Authors | Members

Home News Journals Topics Careers Search

Science Science Advances Science Immunology Science Robotics Science Signaling Science Translational Medicine

SHARE

REPORT

Semantics derived automatically from language corpora contain human-like biases

Aylin Caliskan[1,*], Joanna J. Bryson[1,2,*], Arvind Narayanan[1,*]

+ See all authors and affiliations

Science 14 Apr 2017:
Vol. 356, Issue 6334, pp. 183-186
DOI: 10.1126/science.aal4230

Article Figures & Data

You are currently viewing the abstract.

Science
Vol 356, Issue 6
14 April 2017

Machines learn what people

AlphaGo has demonstrated that a m
many years of concentrated study le
than any human can. Caliskan et al. r
from written texts and that these ass
measured by the Implicit Association

Human-like bias is baked into language!

REPORT

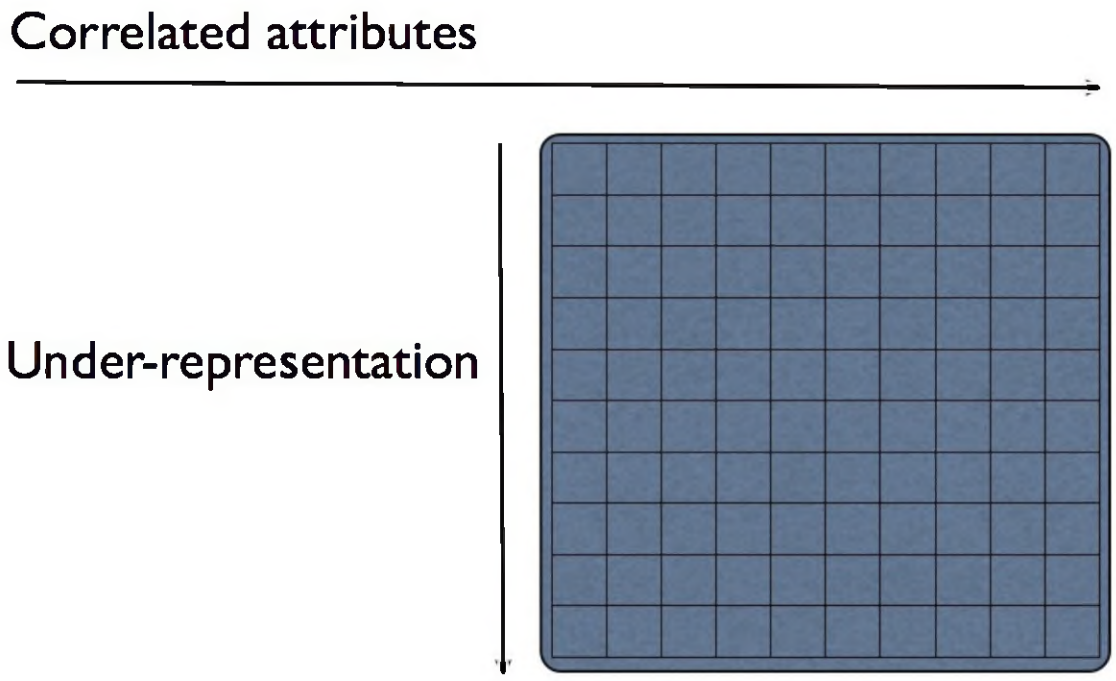
COGNITIVE SCIENCE

Semantics derived automatically from language corpora contain human-like biases

Aylin Caliskan,[1*] Joanna J. Bryson,[1,2*] Arvind Narayanan[1*]

Machine learning is a means to derive artificial intelligence by discovering patterns in existing data. Here, we show that applying machine learning to ordinary human language results in human-like semantic biases. We replicated a spectrum of known biases, as measured by the Implicit Association Test, using a widely used, purely statistical machine-learning model trained on a standard corpus of text from the World Wide Web. Our results indicate that text corpora contain recoverable and accurate imprints of our historic biases, whether morally neutral as toward insects or flowers, problematic as toward race or gender, or even simply veridical, reflecting the status quo distribution of gender with respect to careers or first names. Our methods hold promise for identifying and addressing sources of bias in culture, including technology.

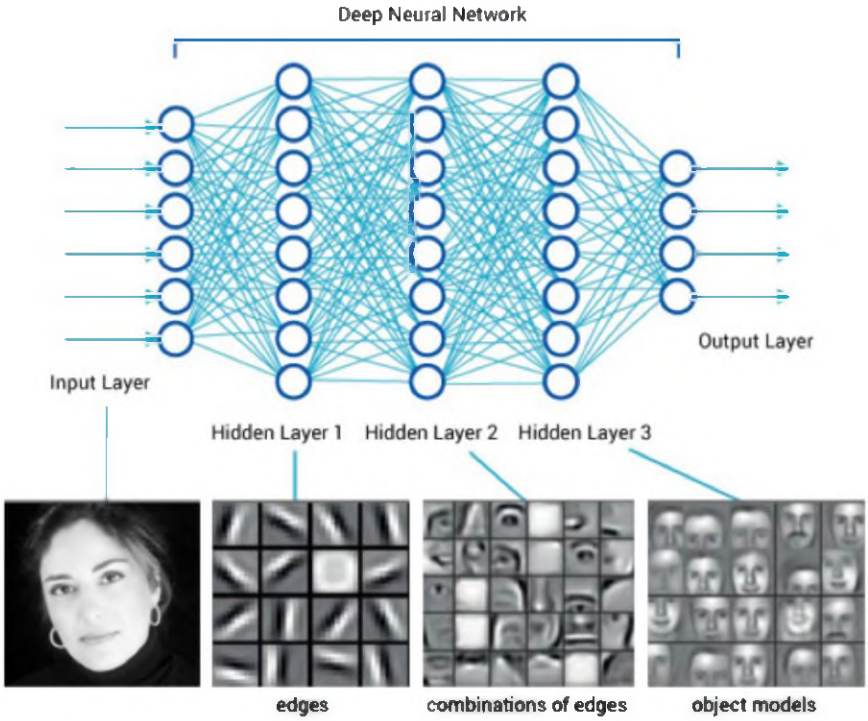
Discrimination/bias in (big) data analytics



Understanding the significance of a mistake

	Really an +	Really a -
Predicted as +	Correct	False +
Predicted as -	False -	Correct

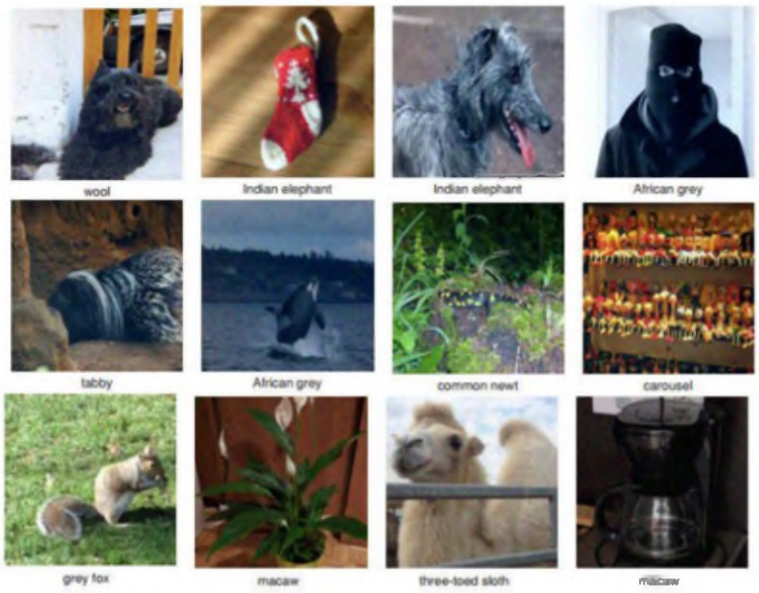
Deep Learning



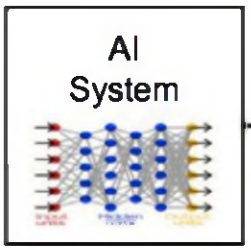
Complex machine learning and data mining methods are difficult, if not, impossible to explain!

Image courtesy of arXiv:1707.08945 [cs.CR]

This is not a 45 MPH sign. You're supposed to stop, not speed up!



Explainable AI



- We are entering a new age of AI applications
- Machine learning is the core technology
- Machine learning models are opaque, non-intuitive, and difficult for people to understand

- Why did you do that?
- Why not something else?
- When do you succeed?
- When do you fail?
- When can I trust you?
- How do I correct an error?

http://www.darpa.mil/program/explainable-artificial-intelligence

Will current AI be illegal in the EU in 2018?

by Colm Gorey

26 OCT 2016 66 SHARES

Series of colourful algorithms on a screen. Image: Mclek/Shutterstock

Silicon Valley and the EU have never really seen eye to eye, but a recent regulation – that means citizens have the right to demand an explanation to an algorithmic decision by 2018 – could drive the two to legal war.

Article 22

Automated individual decision-making, including profiling

1. The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her.

2. Paragraph 1 shall not apply if the decision:

(a) is necessary for entering into, or performance of, a contract between the data subject and a data controller;

(b) is authorised by Union or Member State law to which the controller is subject and which also lays down suitable measures to safeguard the data subject's rights and freedoms and legitimate interests; or

(c) is based on the data subject's explicit consent.

3. In the cases referred to in points (a) and (c) of paragraph 2, the data controller shall implement suitable measures to safeguard the data subject's rights and freedoms and legitimate interests, at least the right to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision.

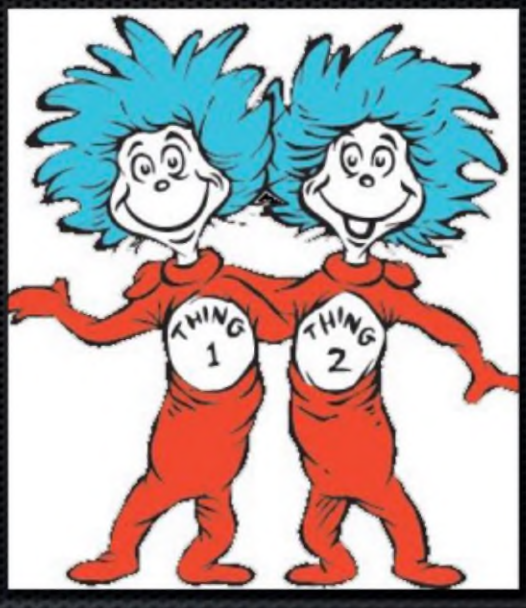
4. Decisions referred to in paragraph 2 shall not be based on special categories of personal data referred to in Article 9(1), unless point (a) or (g) of Article 9(2) applies and suitable measures to safeguard the data subject's rights and freedoms and legitimate interests are in place.

...there's more in Recital 71

- In order to ensure fair and transparent processing ..., the controller should ***use appropriate mathematical or statistical procedures for the profiling***, implement technical and organisational measures appropriate to ensure, in particular, that ... ***prevents, inter alia, discriminatory effects on natural persons*** on the basis of racial or ethnic origin, political opinion, religion or beliefs, trade union membership, genetic or health status or sexual orientation, or that result in measures having such an effect.

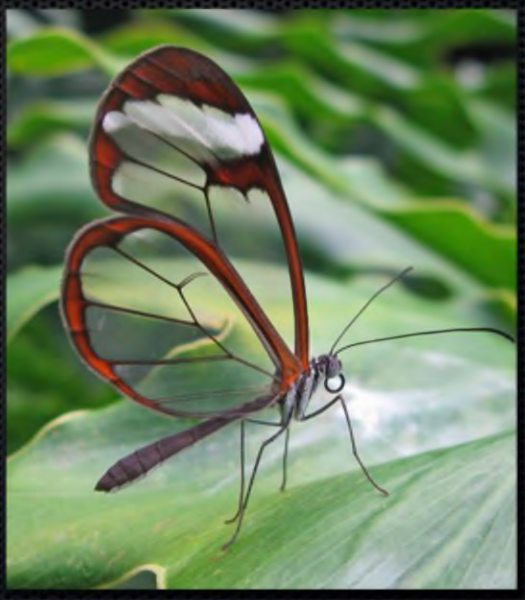
Two major GDPR challenges that are often overlooked...

- The GDPR's position on ***discrimination***
- The notion of a '***right to explanation***'



Barriers to Transparency (Burrell)

- Intentional concealment.
- Gaps in technical literacy.
- ***Mismatch between machine learning/data mining capabilities and human-scale reasoning and interpretation demands.***



GDPR & The Digital Age of Consent

Work with Mary Aiken

We need to talk about the Irish 'digital age of consent'

Mooted change of parental consent limit from 16 to 13 should make us concerned

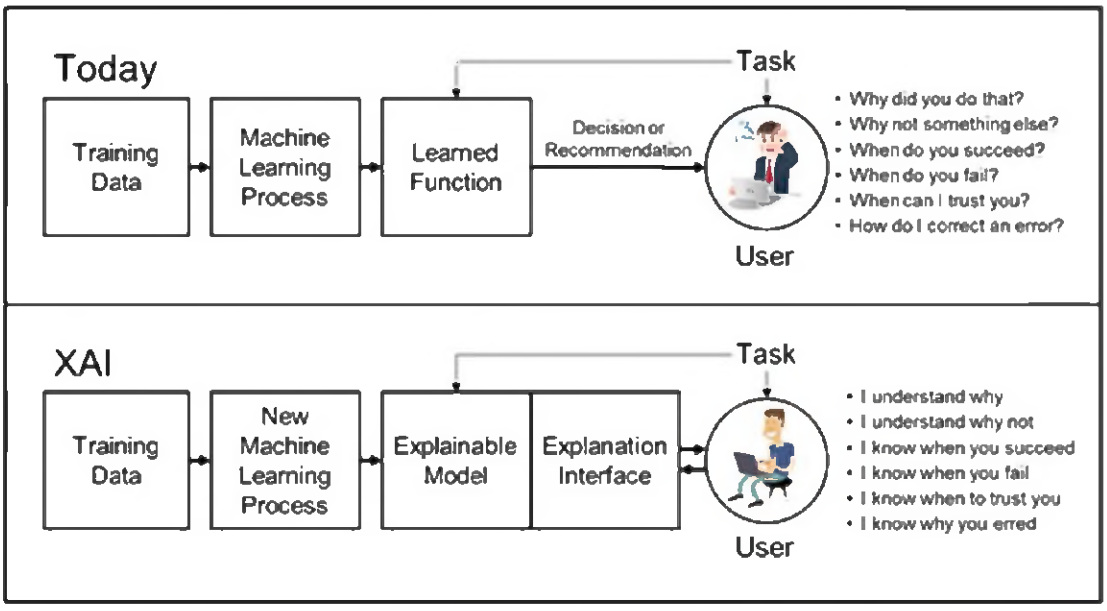
Thu, Jul 13, 2017, 05:38

Mary Aiken, Barry O'Sullivan

2

The default digital age of consent in the EU is 16. However, member states have the option of lowering that age of consent to the age of 13. Photograph: Andrew Gombert/EPA

Explainable AI

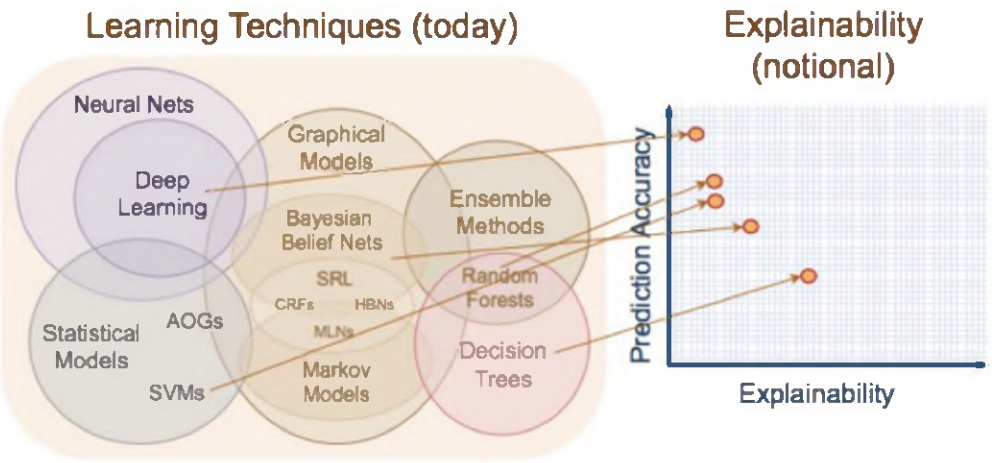


http://www.darpa.mil/program/explainable-artificial-intelligence

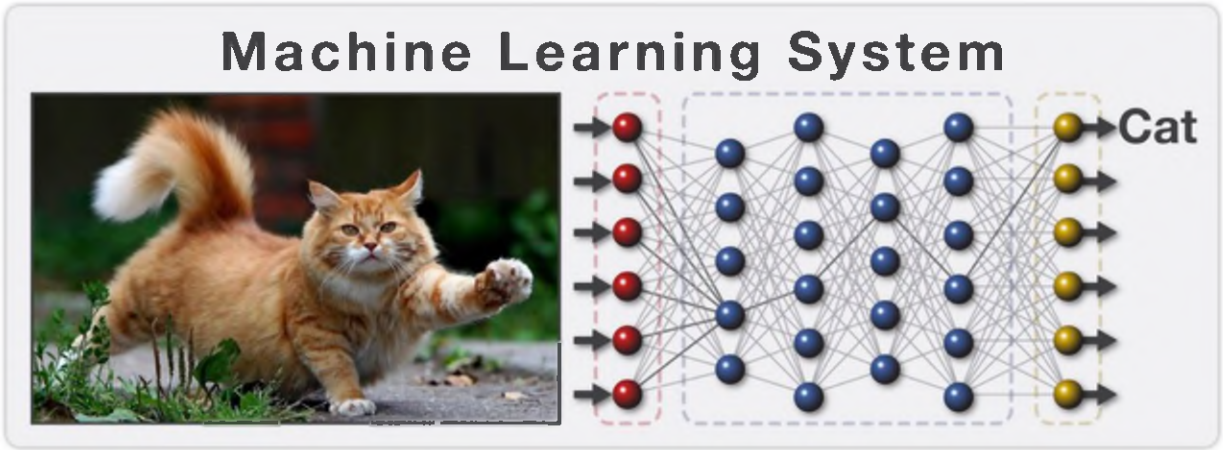
Explainable AI

New Approach

Create a suite of machine learning techniques that produce more explainable models, while maintaining a high level of learning performance



http://www.darpa.mil/program/explainable-artificial-intelligence



This is a cat.

Current Explanation

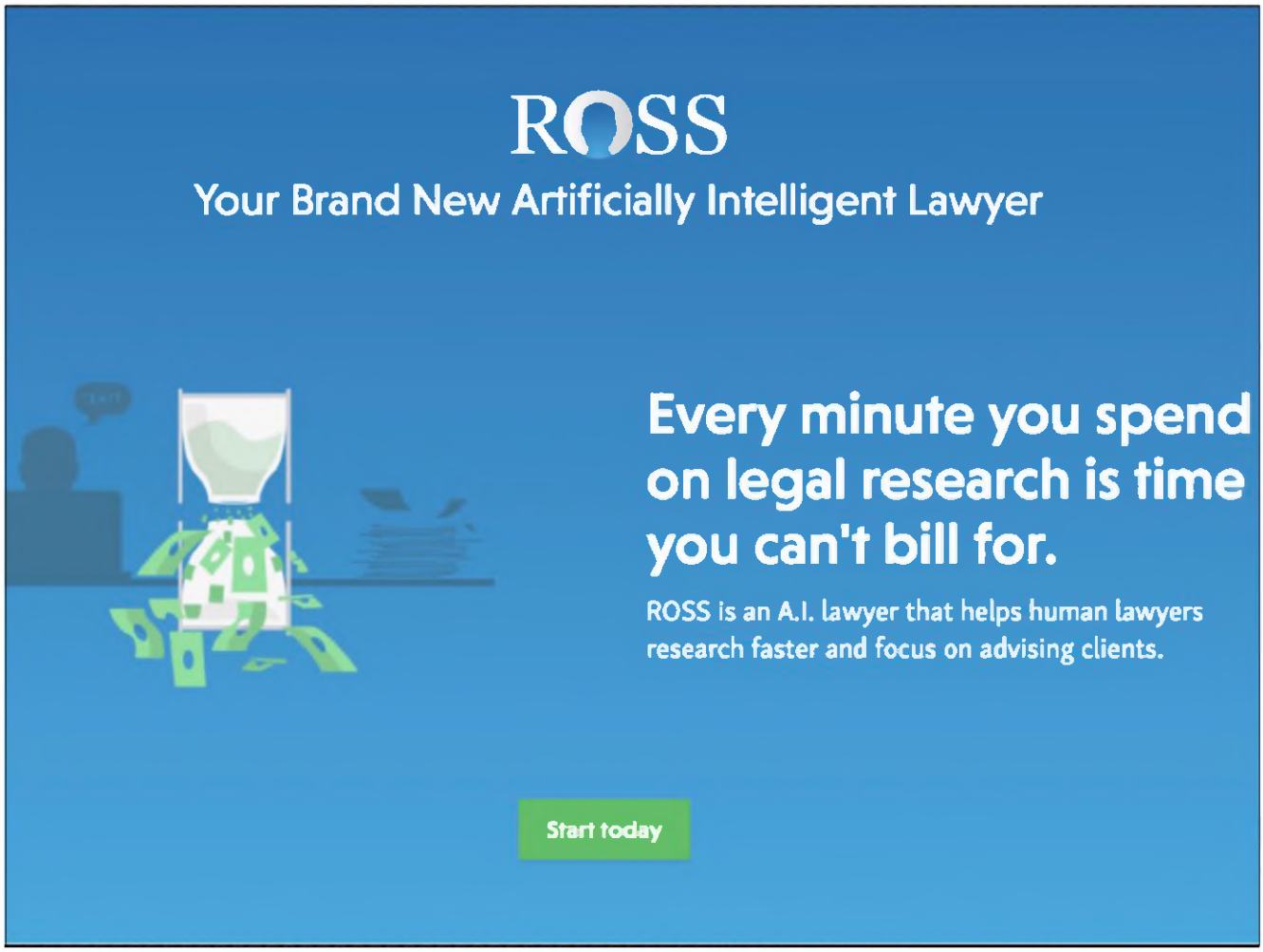
This is a cat:
- It has fur, whiskers, and claws.
- It has this feature:

XAI Explanation

http://www.darpa.mil/program/explainable-artificial-intelligence

Ethics & Artificial Intelligence


Responsibility
Duty
Accountability


ROSS
Your Brand New Artificially Intelligent Lawyer
Every minute you spend on legal research is time you can't bill for.
ROSS is an A.I. lawyer that helps human lawyers research faster and focus on advising clients.
Start today

The world's first artificially intelligent lawyer was just hired at a law firm

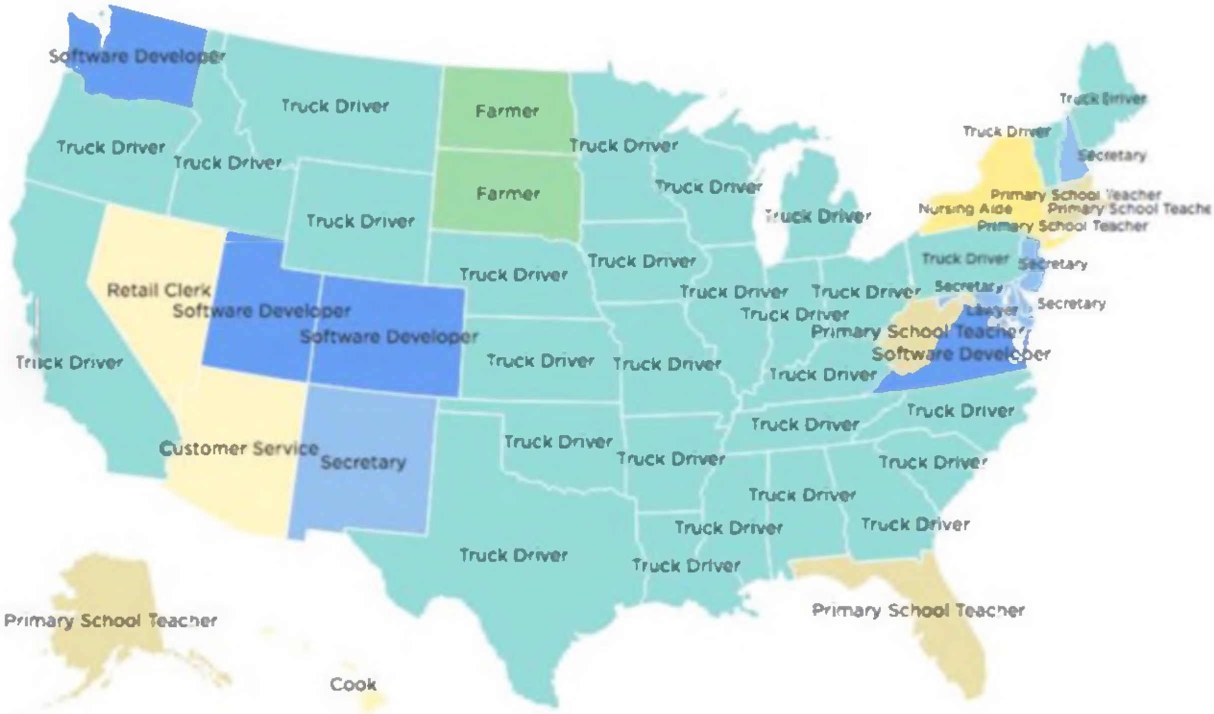
Chris Weller, Tech Insider
May 16, 2016, 3:24 PM 6,351

Lawyers often get a bad reputation for being slimy and conniving (deservedly or not), but ROSS has neither of those qualities.

Ask ROSS to look up an obscure court ruling from 13 years ago, and ROSS will not only search for the case in an instant — without contest or complaint — but it'll offer opinions in plain language about the old ruling's relevance to the case at hand.

A well-dressed humanoid not named Ross. Skye Gould/Tech Insider

Autonomous vehicles exemplify a current AI concern: job loss



Harnessing the Data Revolution for Sustainable Development - Jacob Lew in Addis Ababa

- Data enables **evidence-based policymaking**.
- Data can **accelerate progress on inclusion**, by shining a light on the situations of women, youth, and other disadvantaged groups.
- Data gives us a **tool to monitor and support** the achievement of the SDGs.
- Data **empowers citizens** and it enables a more informed dialogue within and between countries
- Data also plays a **key role in the private sector as a public good** that can help firms make informed investments and take smart

ROBOT-TO-WORKER RATIOS ARE RISING RAPIDLY IN FACTORIES AROUND THE WORLD.

Korea
4.78 ROBOTS

Per 100 Workers

Japan
3.14 ROBOTS

Per 100 Workers

Germany
2.92 ROBOTS

Per 100 Workers

United States
1.64 ROBOTS

Per 100 Workers

Global Average
.66 ROBOTS

Per 100 Workers

China
.36 ROBOTS

Per 100 Workers

https://futurism.com/images/universal-basic-income-answer-automation/

THE DEVELOPING WORLD IS MOST AT RISK.

ETHIOPIA
88%
of jobs at risk

NEPAL
80%
of jobs at risk

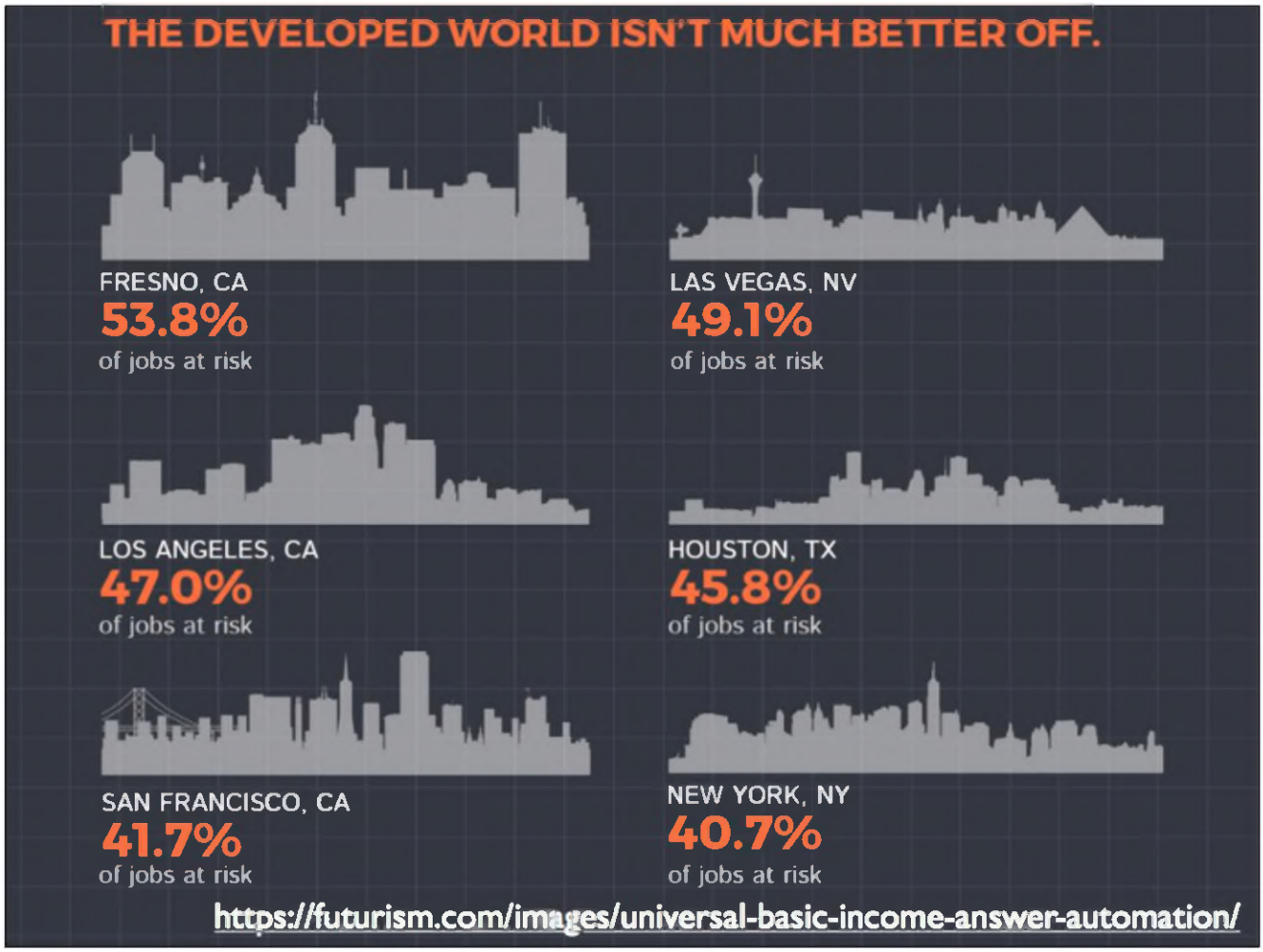
CHINA
77%
of jobs at risk

EL SALVADOR
75%
of jobs at risk

INDIA
69%
of jobs at risk

GLOBAL AVERAGE
57%
of jobs at risk

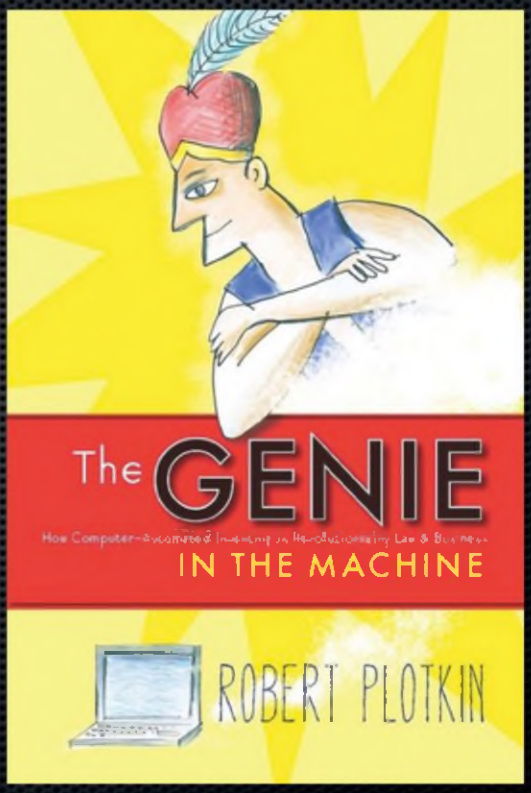
https://futurism.com/images/universal-basic-income-answer-automation/


THE DEVELOPED WORLD ISN'T MUCH BETTER OFF.
FRESNO, CA
53.8%
of jobs at risk
LAS VEGAS, NV
49.1%
of jobs at risk
LOS ANGELES, CA
47.0%
of jobs at risk
HOUSTON, TX
45.8%
of jobs at risk
SAN FRANCISCO, CA
41.7%
of jobs at risk
NEW YORK, NY
40.7%
of jobs at risk
https://futurism.com/images/universal-basic-income-answer-automation/


UNIVERSAL BASIC
INCOME

Can AI Invent?

- Can AI be 'creative'?
  - Genetic Programming
  - Music Performance
- Many AIs are competitive or exceed human performance
- Should AIs be named on patents?
- What are the legal consequences?



Top Ethical Challenges in AI

Source: https://www.weforum.org/agenda/2016/10/top-10-ethical-issues-in-artificial-intelligence

- **Unemployment**. What happens after the end of jobs?
- **Inequality**. How do we distribute the wealth created by machines?
- **Humanity**. How do machines affect our behaviour and interaction?
- **Artificial stupidity**. How can we guard against mistakes?
- **Racist robots**. How do we eliminate AI bias?
- **Security**. How do we keep AI safe from adversaries?
- **Evil genies**. How do we protect against unintended consequences?
- **Singularity**. How do we stay in control of a complex intelligent system?
- **Robot rights**. How do we define the humane treatment of AI?

An Overview of Artificial Intelligence

Professor Barry O'Sullivan, PhD(NUI), FEurAI, MRIA

European Artificial Intelligence Association &

Department of Computer Science, University College Cork