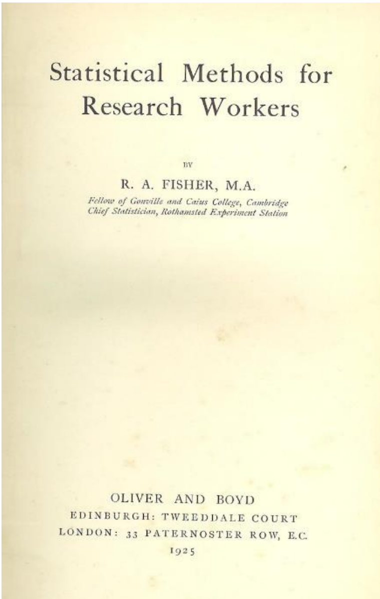
The Alan Turing Institute

The data science revolution in scientific research

Jonathan Rowe
Programme Director, Data Science for Science



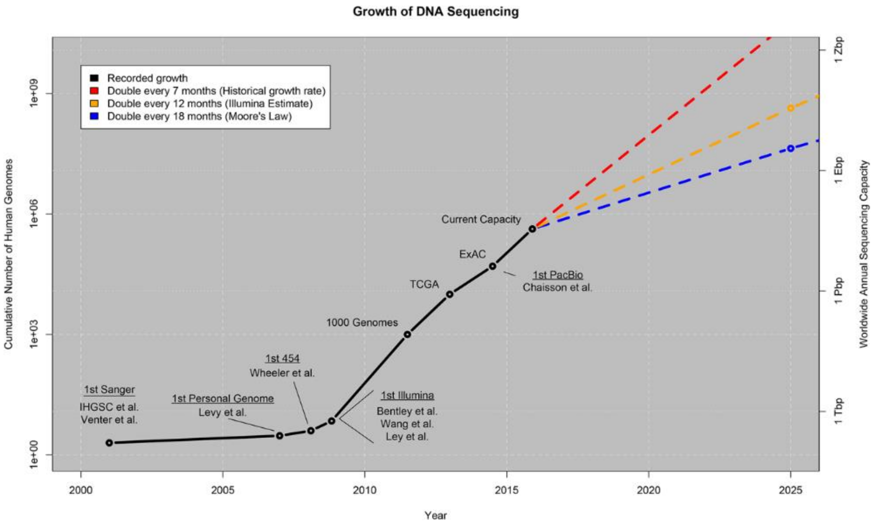
Statistical Methods for
Research Workers

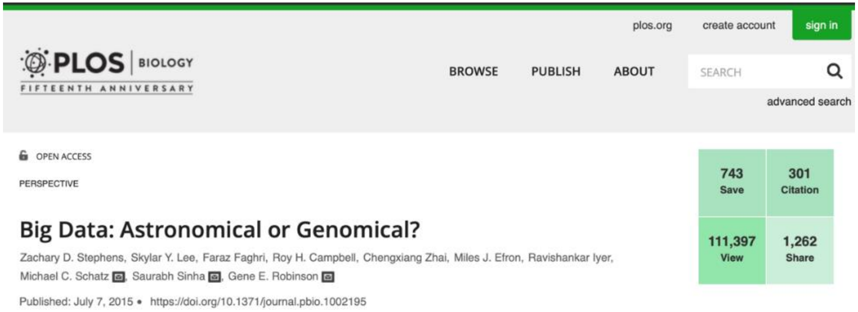
BY

R. A. FISHER, M.A.

Fellow of Gonville and Caius College, Cambridge
Chief Statistician, Rothamsted Experiment Station

OLIVER AND BOYD
EDINBURGH: TWEEDDALE COURT
LONDON: 33 PATERNOSTER ROW, E.C.
1925





plos.org create account sign in

PLOS | BIOLOGY
FIFTEENTH ANNIVERSARY

BROWSE PUBLISH ABOUT SEARCH

advanced search

OPEN ACCESS

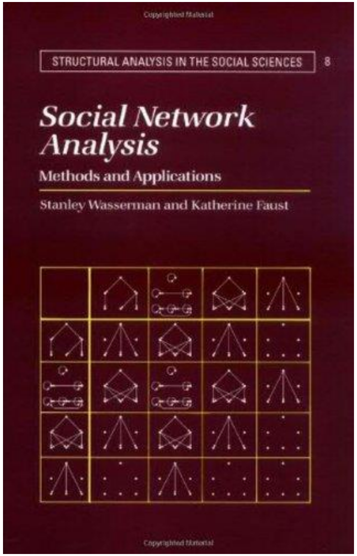
PERSPECTIVE

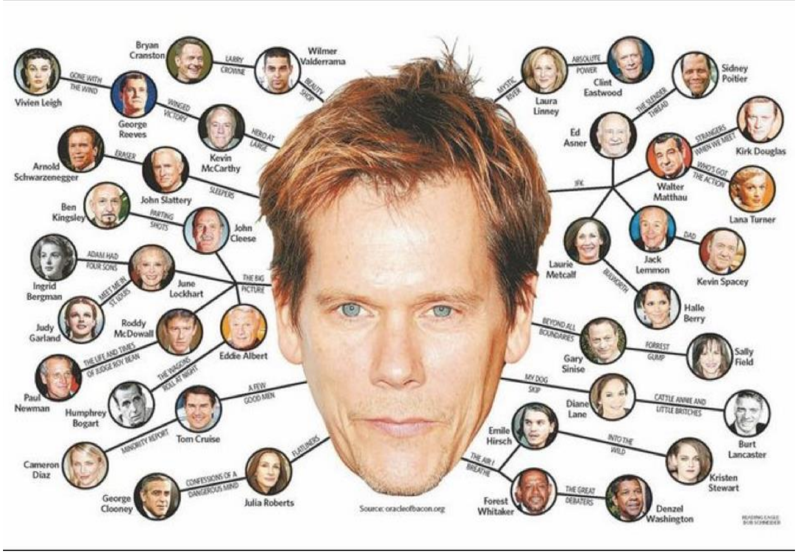
Big Data: Astronomical or Genomical?

Zachary D. Stephens, Skylar Y. Lee, Faraz Faghri, Roy H. Campbell, Chengxiang Zhai, Miles J. Efron, Ravishankar Iyer, Michael C. Schatz, Saurabh Sinha, Gene E. Robinson

Published: July 7, 2015 • https://doi.org/10.1371/journal.pbio.1002195

743 Save	301 Citation
111,397 View	1,262 Share


STRUCTURAL ANALYSIS IN THE SOCIAL SCIENCES 8
Social Network Analysis
Methods and Applications
Stanley Wasserman and Katherine Faust


Vivien Leigh
George Reeves
Bryan Cranston
Wilmer Valderrama
Kevin McCarthy
Arnold Schwarzenegger
John Slattery
Ben Kingsley
John Cleese
Ingrid Bergman
June Lockhart
Judy Garland
Roddy McDowall
Eddie Albert
Paul Newman
Humphrey Bogart
Tom Cruise
Cameron Diaz
George Clooney
Julia Roberts
Laura Linney
Clint Eastwood
Sidney Poitier
Ed Asner
Kirk Douglas
Walter Matthau
Lana Turner
Laurie Metcalf
Jack Lemmon
Kevin Spacey
Halle Berry
Gary Sinise
Sally Field
Diane Lane
Burt Lancaster
Emile Hirsch
Kristen Stewart
Forest Whitaker
Denzel Washington
Source: oracleofbacon.org

CENTRAL INTELLIGENCE AGENCY

THE WORK OF A NATION.
THE CENTER OF INTELLIGENCE.

Report Threats | رعربيعربي | Contact

عربي 中文 English Français Русский Español More ▸

HOME | ABOUT CIA | CAREERS & INTERNSHIPS | OFFICES OF CIA | NEWS & INFORMATION | LIBRARY | KIDS' ZONE

Library

Library

- Publications
- Center for the Study of Intelligence
- **Freedom of Information Act Electronic Reading Room**
  - How to File a FOIA Request
  - How to Check on the Status of Your FOIA Request
  - How to File a Privacy Act Request
  - Frequently Requested Records
  - Annual Reports
  - Historical Collections
  - CREST: 25-Year

CIA Home » Library » Electronic Reading Room

Search Query for FOIA ERR:

Search

FOIA Category Search | Advanced Search | Search Help

AN ASSESSMENT OF THE EVIDENCE FOR PSYCHIC FUNCTIONING

Document Type: CREST
Collection: STARGATE
Document Number (FOIA) /ESDN (CREST): CIA-RDP96-00791R000200070001-9
Release Decision: RIFPUB
Original Classification: K
Document Page Count: 32
Document Creation Date: November 4, 2016
Document Release Date: May 22, 2000
Sequence Number: 1
Case Number:
Content Type: RP
File:

Attachment	Size

PLOS | ONE

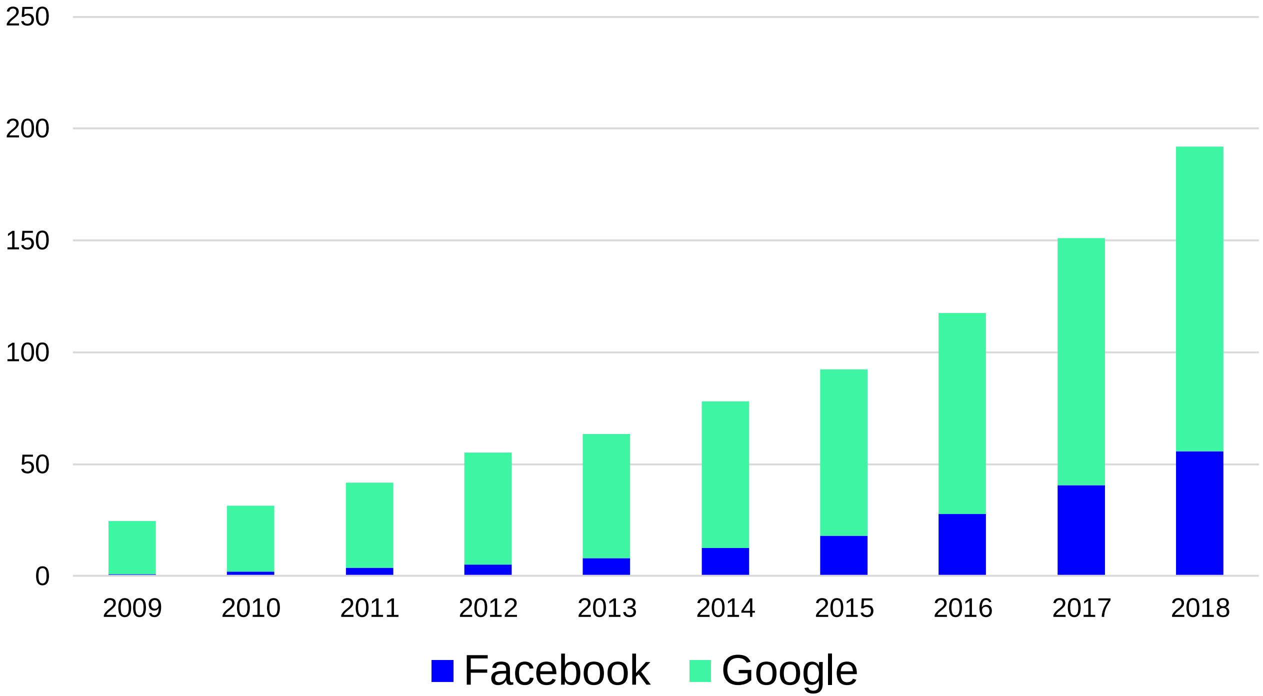
RESEARCH ARTICLE

Testing for Questionable Research Practices in a Meta-Analysis: An Example from Experimental Parapsychology

Dick J. Bierman[1,2]*, James P. Spottiswoode[3], Aron Bijl[4]

1 Experimental Psychology, University of Groningen, Groningen, The Netherlands, **2** Brain & Cognition, University of Amsterdam, Amsterdam, The Netherlands, **3** LFR, Palo Alto, California, United States of America, **4** Brain & Cognition, University of Amsterdam, Amsterdam, The Netherlands

Annual revenue (mostly advertising) in US$b



ImageNet Classification with Deep Convolutional Neural Networks

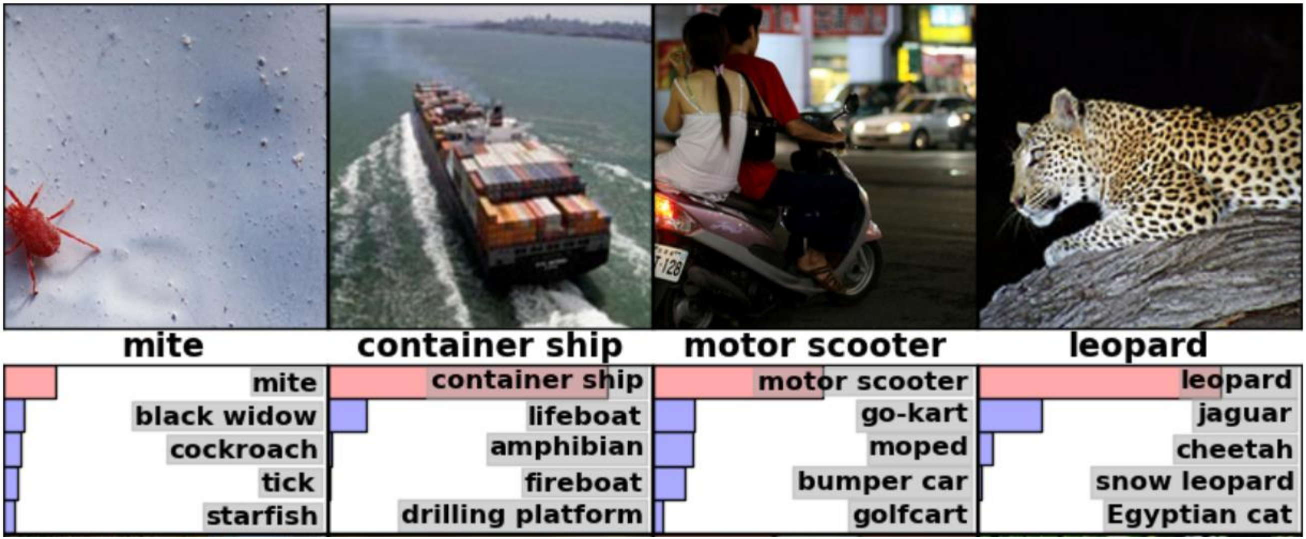
Alex Krizhevsky
University of Toronto
kriz@cs.utoronto.ca

Ilya Sutskever
University of Toronto
ilya@cs.utoronto.ca

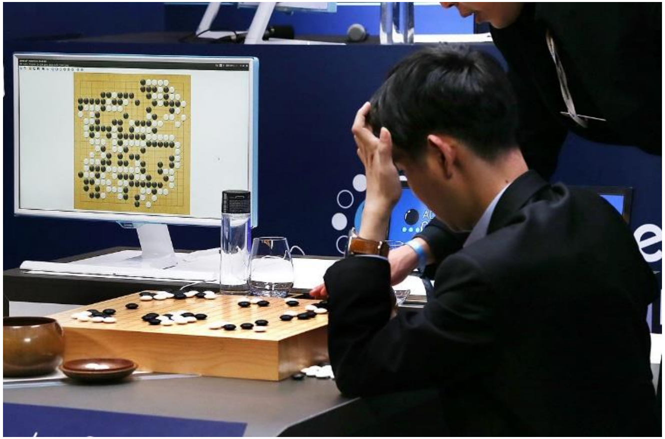
Geoffrey E. Hinton
University of Toronto
hinton@cs.utoronto.ca

Abstract

We trained a large, deep convolutional neural network to classify the 1.2 million high-resolution images in the ImageNet LSVRC-2010 contest into the 1000 different classes. On the test data, we achieved top-1 and top-5 error rates of 37.5% and 17.0% which is considerably better than the previous state-of-the-art. The neural network, which has 60 million parameters and 650,000 neurons, consists of five convolutional layers, some of which are followed by max-pooling layers, and three fully-connected layers with a final 1000-way softmax. To make training faster, we used non-saturating neurons and a very efficient GPU implementation of the convolution operation. To reduce overfitting in the fully-connected layers we employed a recently-developed regularization method called "dropout" that proved to be very effective. We also entered a variant of this model in the ILSVRC-2012 competition and achieved a winning top-5 test error rate of 15.3%, compared to 26.2% achieved by the second-best entry.


mite
container ship
motor scooter
leopard
mite
black widow
cockroach
tick
starfish
container ship
lifeboat
amphibian
fireboat
drilling platform
motor scooter
go-kart
moped
bumper car
golfcart
leopard
jaguar
cheetah
snow leopard
Egyptian cat

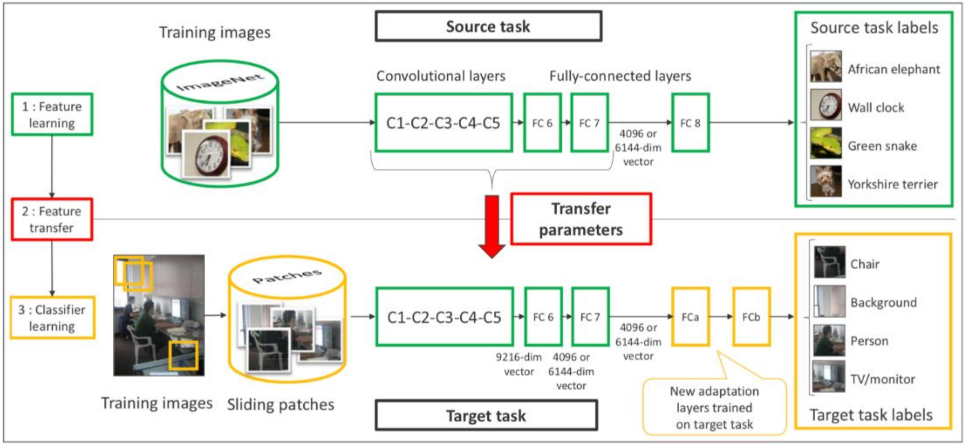

truck 0.959
person 0.968
tie 0.89
car 0.967
car 0.983
car 0.957
truck 0.940
truck 0.779
bicycle 0.761



A general reinforcement learning algorithm that masters chess, shogi and Go through self-play

David Silver,[1,2*] Thomas Hubert,[1*] Julian Schrittwieser,[1*]
Ioannis Antonoglou,[1,2] Matthew Lai,[1] Arthur Guez,[1] Marc Lanctot,[1]
Laurent Sifre,[1] Dharshan Kumaran,[1,2] Thore Graepel,[1,2]
Timothy Lillicrap,[1] Karen Simonyan,[1] Demis Hassabis[1]

[1]DeepMind, 6 Pancras Square, London N1C 4AG.
[2]University College London, Gower Street, London WC1E 6BT.



Learning and Transferring Mid-Level Image Representations using Convolutional Neural Networks

Maxime Oquab[1,*] Leon Bottou[2] Ivan Laptev[1,*] Josef Sivic[1,*]
[1]INRIA, Paris, France [2]MSR, New York, USA

nature > scientific reports > articles > article

SCIENTIFIC REPORTS

Article | OPEN | Published: 16 October 2017

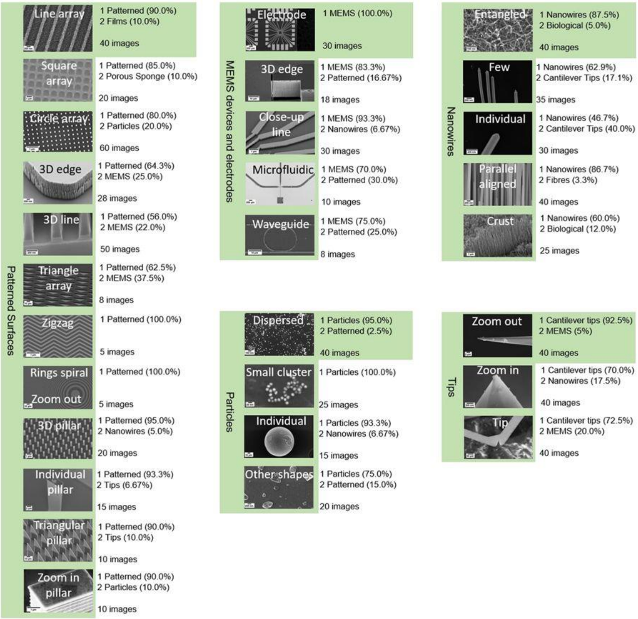
Neural Network for Nanoscience Scanning Electron Microscope Image Recognition

Mohammad Hadi Modarres, Rossella Aversa, Stefano Cozzini, Regina Ciancio, Angelo Leto & Giuseppe Piero Brandino

Scientific Reports **7**, Article number: 13282 (2017) | Download Citation

Abstract

In this paper we applied transfer learning techniques for image recognition, automatic categorization, and labeling of nanoscience images obtained by scanning electron microscope (SEM). Roughly 20,000 SEM images were manually classified into 10 categories to form a labeled training set, which can be used as a reference set for future applications of deep learning enhanced algorithms in the nanoscience domain. The categories chosen spanned the range of 0-Dimensional (0D) objects such as particles, 1D nanowires and fibres, 2D films and coated surfaces, and 3D patterned surfaces such as pillars. The training set was used to retrain on the SEM dataset and to compare many



Generative Adversarial Nets

Ian J. Goodfellow,* Jean Pouget-Abadie,† Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair,‡ Aaron Courville, Yoshua Bengio§
Département d'informatique et de recherche opérationnelle
Université de Montréal
Montréal, QC H3C 3J7

Published as a conference paper at ICLR 2018

Progressive Growing of GANs for Improved Quality, Stability, and Variation

Tero Karras NVIDIA
Timo Aila NVIDIA
Samuli Laine NVIDIA
Jaakko Lehtinen NVIDIA and Aalto University
{tkarras,taila,slaine,jlehtinen}@nvidia.com

Open Access

Accelerating Science with Generative Adversarial Networks: An Application to 3D Particle Showers in Multilayer Calorimeters

Michela Paganini, Luke de Oliveira, and Benjamin Nachman
Phys. Rev. Lett. **120**, 042003 – Published 26 January 2018

Article | References | Citing Articles (11) | PDF | HTML | Export Citation

ABSTRACT

Physicists at the Large Hadron Collider (LHC) rely on detailed simulations of particle collisions to build expectations of what experimental data may look like under different theoretical modeling assumptions. Petabytes of simulated data are needed to develop analysis techniques, though they are expensive to generate using existing algorithms and computing resources. The modeling of detectors and the precise description of particle cascades as they interact with the material in the calorimeter are the most computationally demanding steps in the simulation pipeline. We therefore introduce a deep neural network-based generative model to enable high-fidelity, fast, electromagnetic calorimeter simulation. There are still challenges for achieving precision across the entire phase space, but our current solution can reproduce a variety of particle shower properties while achieving speedup factors of up to $100\,000\times$. This opens the door to a new era of fast simulation that could save significant computing time and disk space, while extending the reach of physics searches and precision measurements at the LHC and beyond.

Active learning machine learns to create new quantum experiments

Alexey A. Melnikov, Hendrik Poulsen Nautrup, Mario Krenn, Vedran Dunjko, Markus Tiersch, Anton Zeilinger, and Hans J. Briegel

PNAS February 6, 2018 115 (6) 1221-1226; published ahead of print January 18, 2018 https://doi.org/10.1073/pnas.1714936115

Contributed by Anton Zeilinger, November 14, 2017 (sent for review August 24, 2017; reviewed by Jacob D. Biamonte and Jonathan P. Dowling)

Article | Figures & SI | Info & Metrics | PDF

Significance

Quantum experiments push the envelope of our understanding of fundamental concepts in quantum physics. Modern experiments have exhaustively probed the basic notions of quantum theory. Arguably, further breakthroughs require the tackling of complex quantum phenomena and consequently require complex experiments and involved techniques. The designing of such complex experiments is difficult and often clashes with human intuition. We present an autonomous learning model which learns to design such complex experiments, without relying on previous knowledge or often flawed intuition. Our system not only learns how to design desired experiments more efficiently than the best previous approaches, but in the process also discovers nontrivial experimental techniques. Our work demonstrates that learning machines can offer dramatic advances in how experiments are generated.

Founding the Institute

"We will found The Alan Turing Institute to ensure Britain leads the way again in the use of big data and algorithm research"

George Osborne, Chancellor of the Exchequer
Budget Speech, March 2014

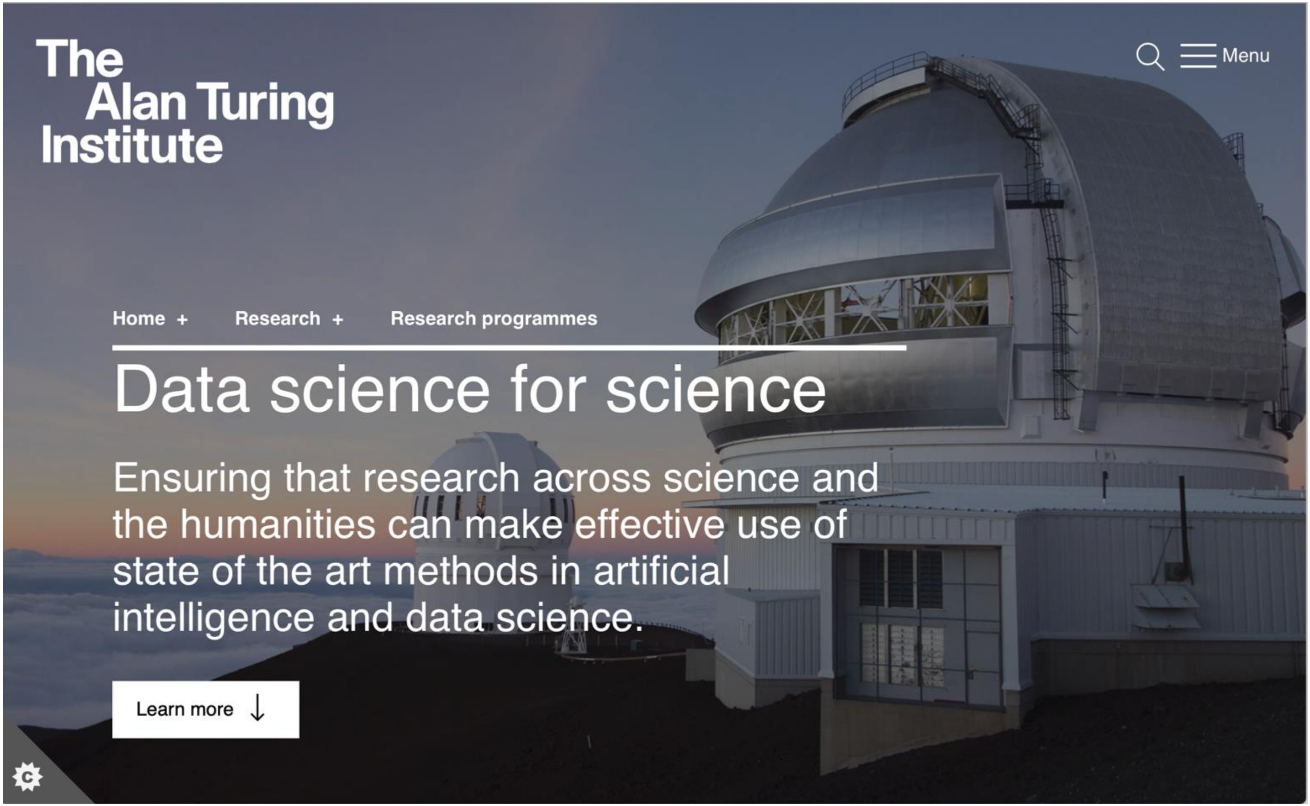
The Alan Turing Institute

EPSRC
Engineering and Physical Sciences Research Council

Network of industry, charity, government partners

Network of university members

Strategic government investment


The Alan Turing Institute
Menu
Home +
Research +
Research programmes
Data science for science
Ensuring that research across science and the humanities can make effective use of state of the art methods in artificial intelligence and data science.
Learn more

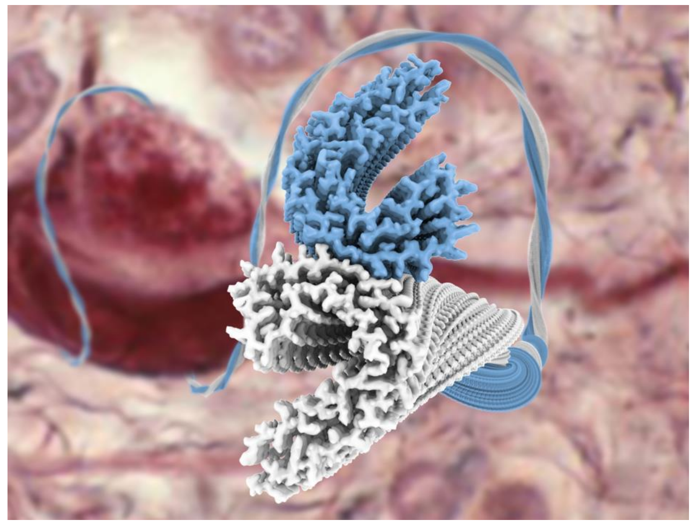
nature
International journal of science

Article | Published: 05 July 2017

Cryo-EM structures of tau filaments from Alzheimer's disease

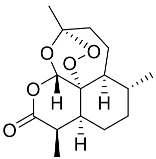
Anthony W. P. Fitzpatrick, Benjamin Falcon, Shaoda He, Alexey G. Murzin, Garib Murshudov, Holly J. Garringer, R. Anthony Crowther, Bernardino Ghetti, Michel Goedert & Sjors H. W. Scheres

Nature **547**, 185–190 (13 July 2017) | Download Citation



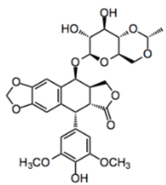
Anti-malarial

Artemisinin
Wormwood
(*artemisia annua*)



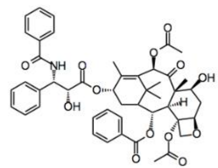
Anti-cancer

Etoposide Mayapple
(*podophyllum)*



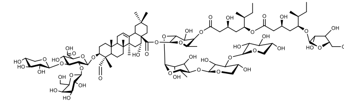
Anti-cancer

Taxol Yew
(*taxus baccata*)



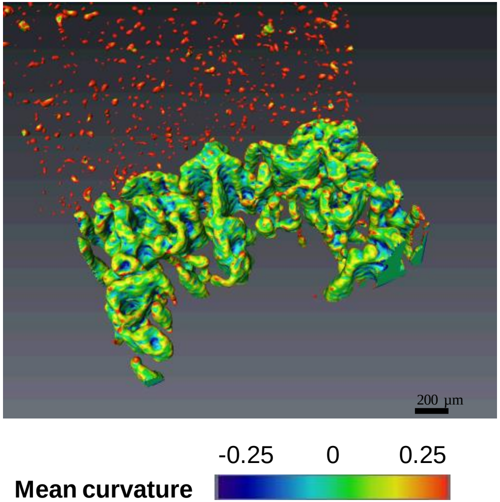
Adjuvant

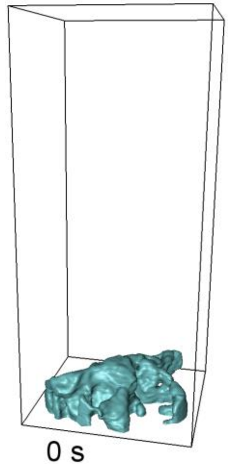
Qs-21 Soapbark
(*quillaja saponaria*)

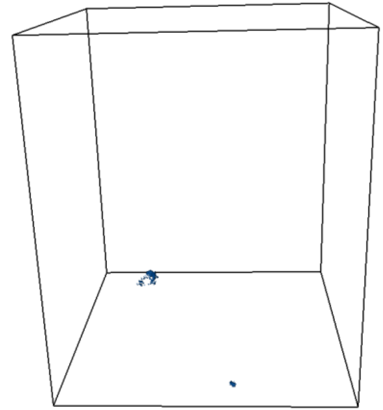


The majority of plant natural products cannot be accessed – and most of them have not yet been discovered

John Innes Centre





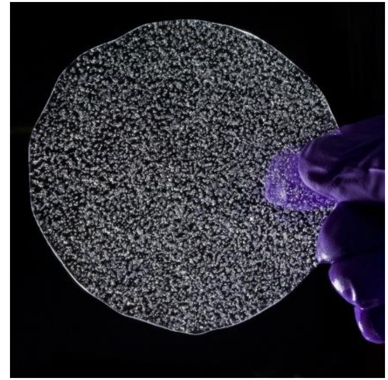


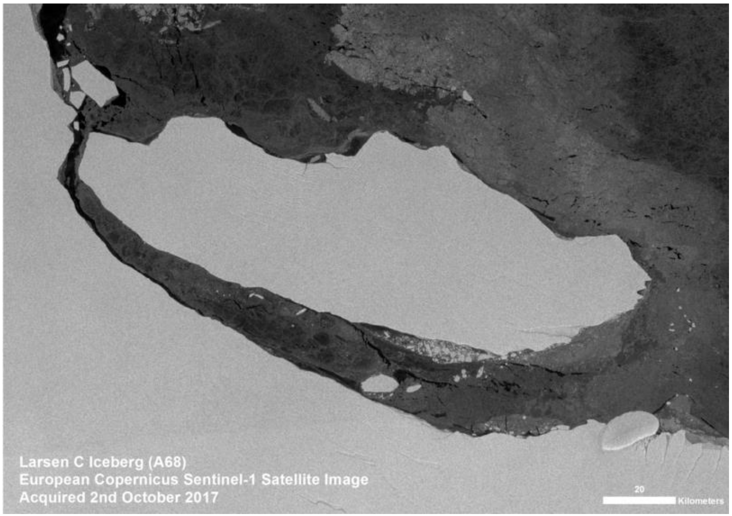
A positive mean curvature means concave towards the liquid

B. Cai et al. Acta Mater. 2014(76): 371-380; B. Cai et al. Acta Mater. 2017(117):160-169;
B. Cai et al. Scripta Mater. submitted.

British
Antarctic Survey

NATURAL ENVIRONMENT RESEARCH COUNCIL





ARTICT: Art through the ICT lens

Francisco de Goya, Doña Isabel de Porcel (NG1473), before 1805

Color Photograph

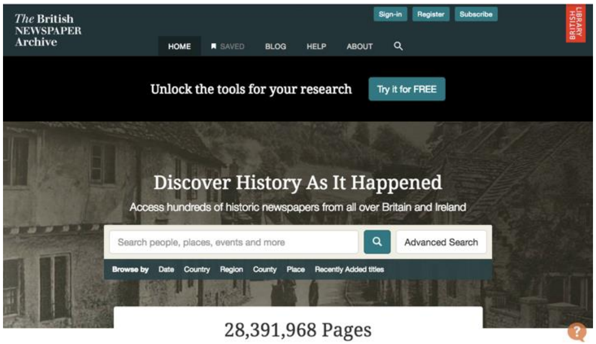
XRR

IRR

XRF

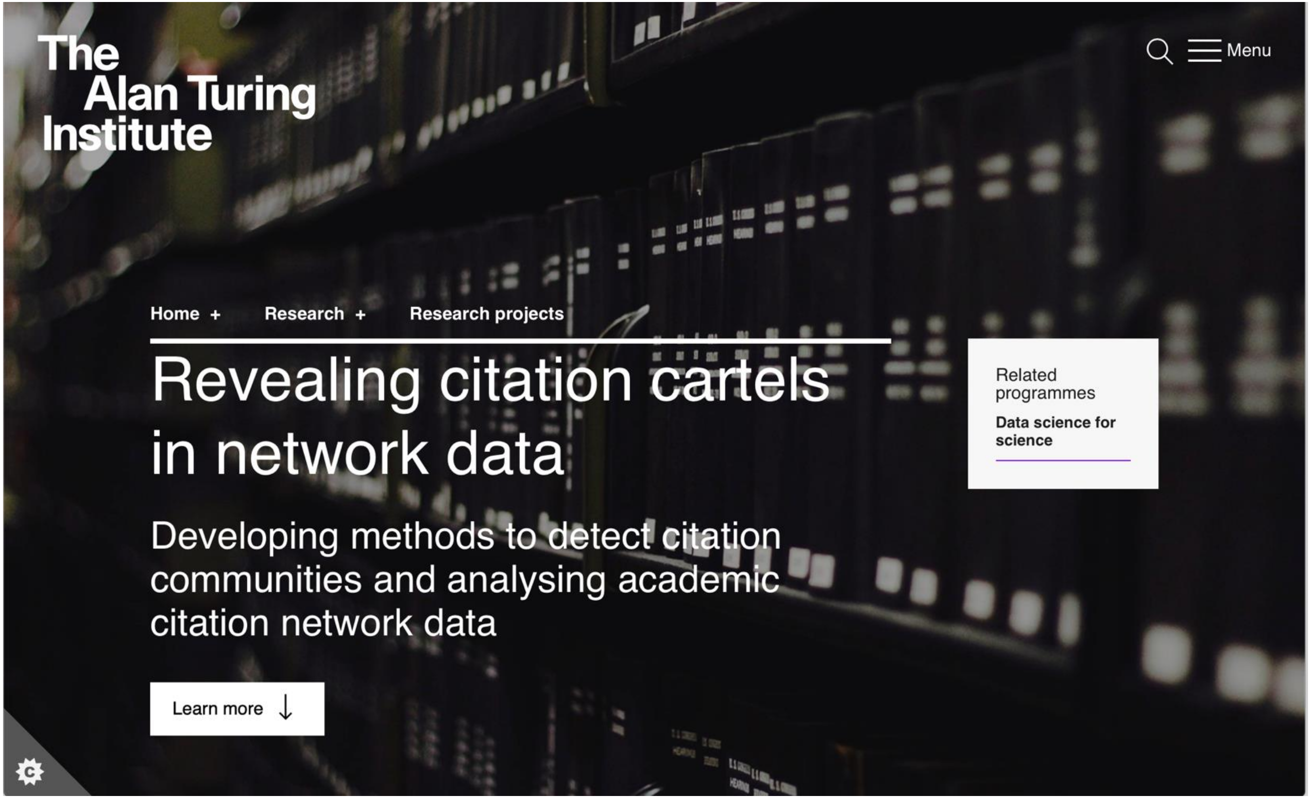
EPSRC Engineering and Physical Sciences Research Council · THE NATIONAL GALLERY · UCL · Imperial College London · Duke UNIVERSITY · BRUKER

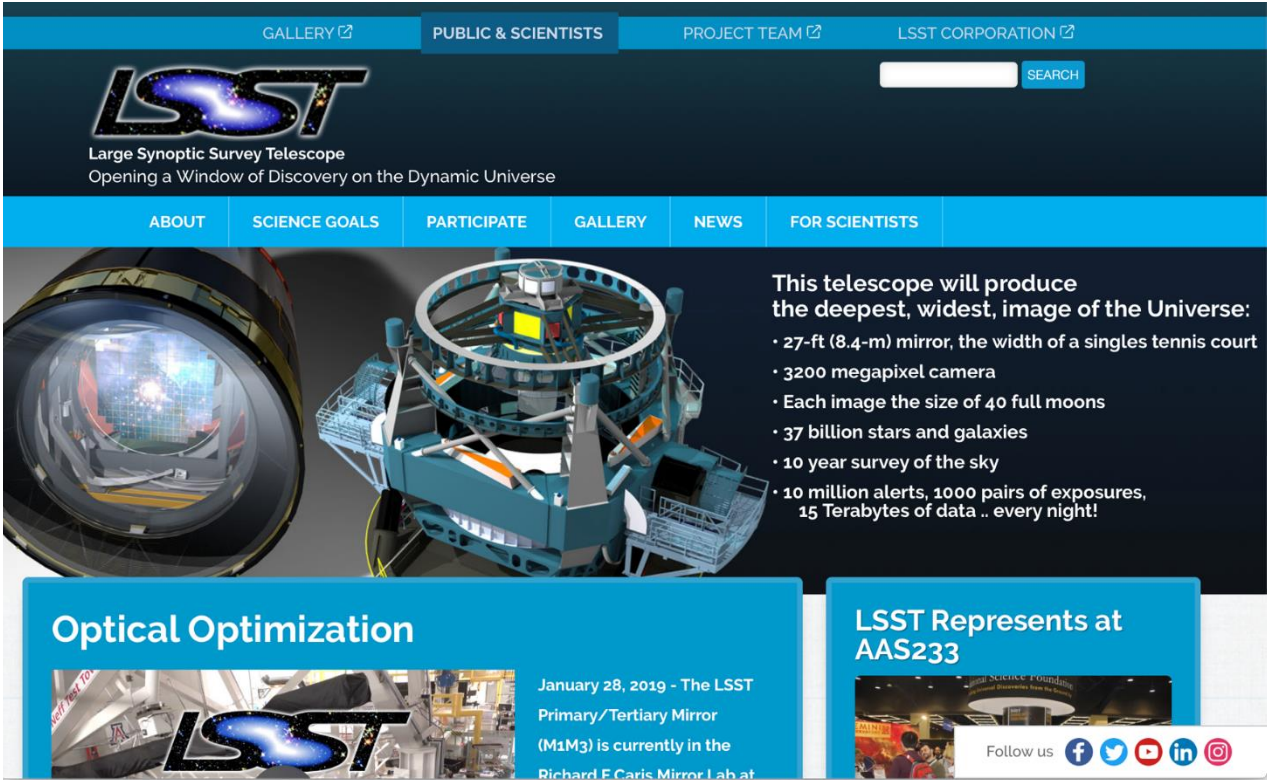
Living with Machines – history from data



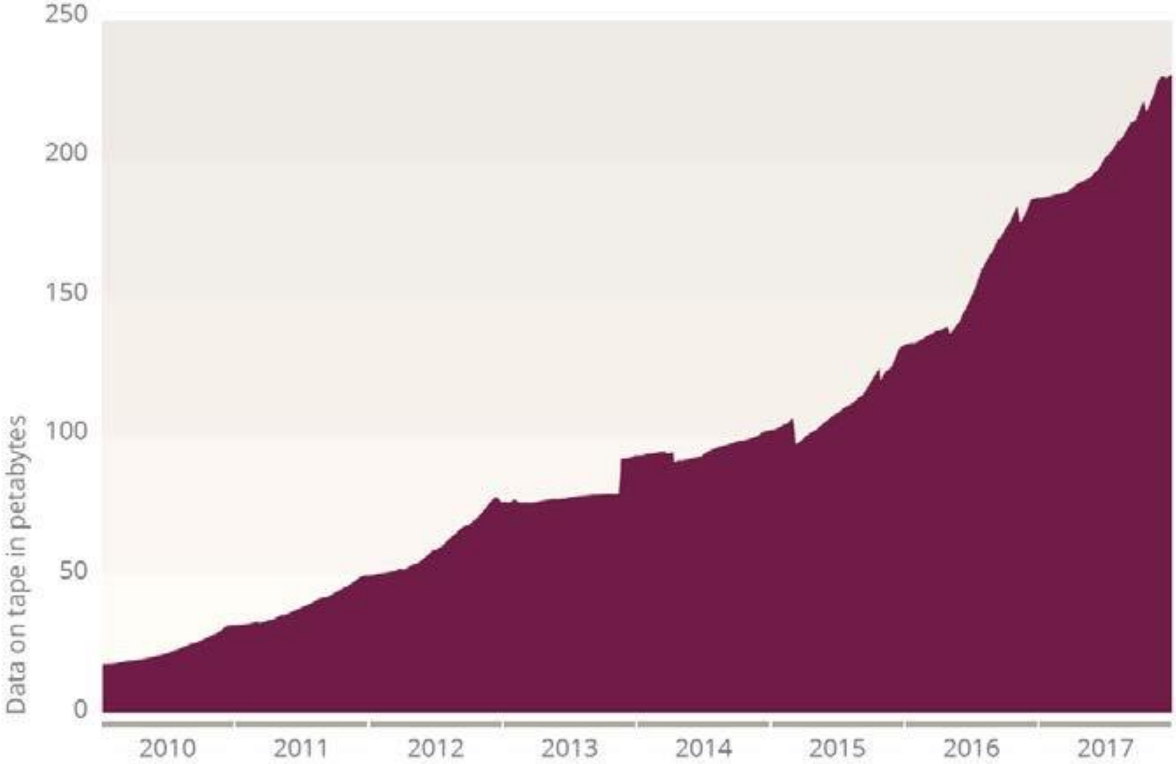
The **British Library's** British Newspaper Archive

Findmypast's event-based records (census, electoral registration, births/ marriages/deaths, trade directories).


The Alan Turing Institute
Menu
Home +
Research +
Research projects
Revealing citation cartels in network data
Developing methods to detect citation communities and analysing academic citation network data
Learn more
Related programmes
Data science for science


GALLERY
PUBLIC & SCIENTISTS
PROJECT TEAM
LSST CORPORATION
SEARCH
LSST
Large Synoptic Survey Telescope
Opening a Window of Discovery on the Dynamic Universe
ABOUT
SCIENCE GOALS
PARTICIPATE
GALLERY
NEWS
FOR SCIENTISTS
This telescope will produce the deepest, widest, image of the Universe:
• 27-ft (8.4-m) mirror, the width of a singles tennis court
• 3200 megapixel camera
• Each image the size of 40 full moons
• 37 billion stars and galaxies
• 10 year survey of the sky
• 10 million alerts, 1000 pairs of exposures, 15 Terabytes of data .. every night!
Optical Optimization
January 28, 2019 - The LSST Primary/Tertiary Mirror (M1M3) is currently in the
LSST Represents at AAS233
Follow us

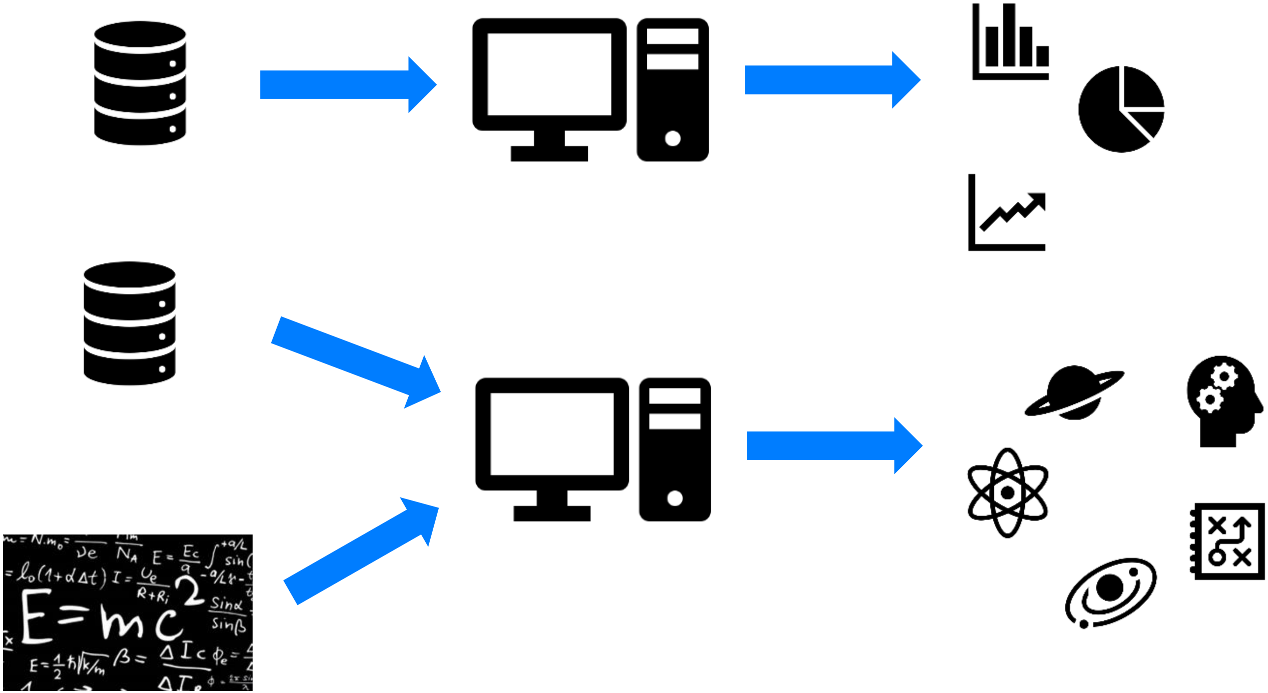
EVOLUTION OF THE TOTAL AMOUNT OF DATA STORED ON TAPE AT CERN IN PB

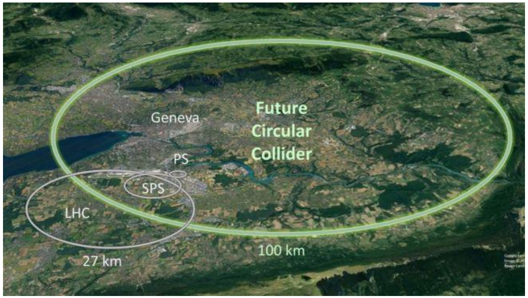


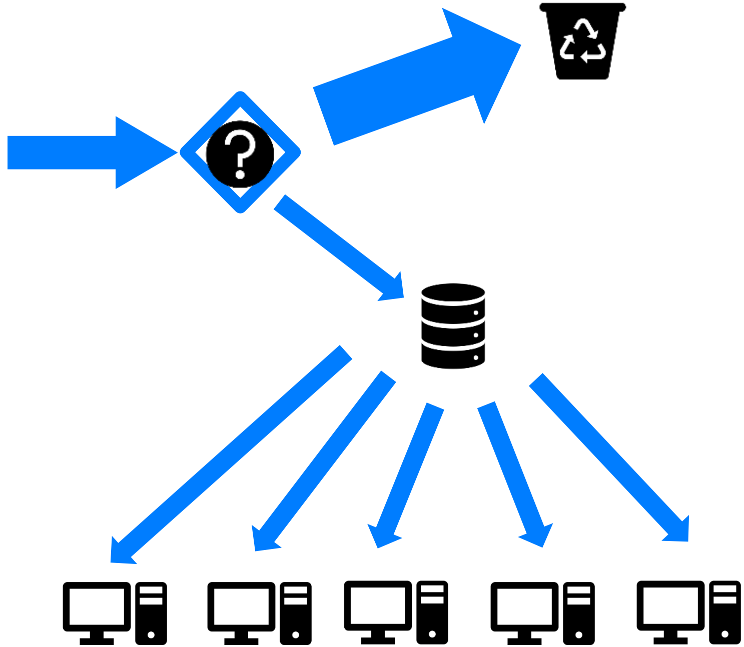
"A just machine to make big decisions
Programmed by fellows with compassion and vision
We'll be clean when their work is done
We'll be eternally free, yes and eternally young.

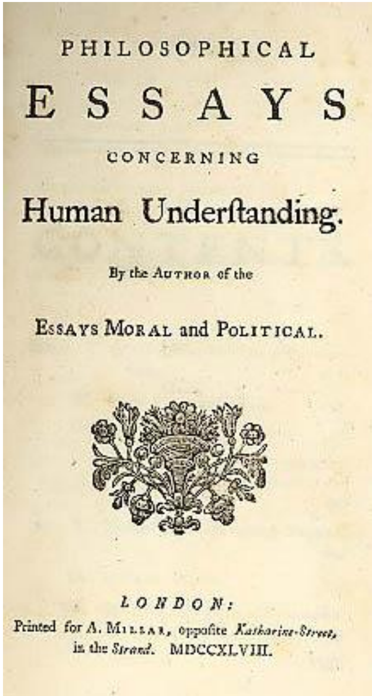
What a beautiful world this will be
What a glorious time to be free."

Donald Fagen
I.G.Y. (International Geophysical Year)


E=mc2


Geneva
Future Circular Collider
PS
SPS
LHC
27 km
100 km





PHILOSOPHICAL

ESSAYS

CONCERNING

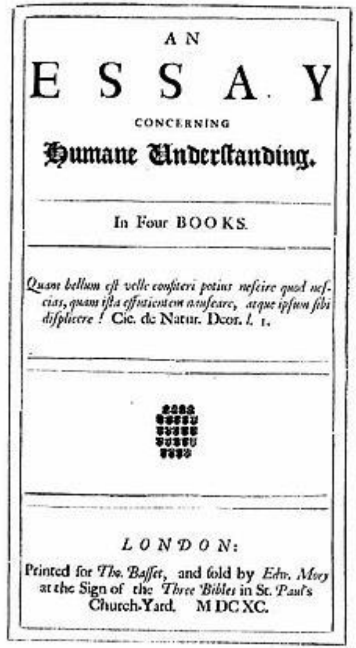
Human Underſtanding.

By the AUTHOR of the

ESSAYS MORAL and POLITICAL.

LONDON:

Printed for A. MILLAR, oppoſite *Katharine-Street*, in the *Strand*. MDCCXLVIII.



AN

ESSAY

CONCERNING

Humane Underſtanding.

In Four BOOKS.

Quam bellum eſt velle confiteri potius neſcire quod neſcias, quam iſta effutientem nauſeare, atque ipſum ſibi diſplicere! Cic. de Natur. Deor. *l.* 1.

LONDON:

Printed for *Tho. Baſſet*, and ſold by *Edw. Mory* at the Sign of the *Three Bibles* in St. *Paul's* Church-Yard. MDCXC.

Critik

der

reinen Vernunft

von

Immanuel Kant

Profeſſor in Königsberg.

Riga,

verlegts Johann Friedrich Hartknoch

1781.

RESEARCH | *REPORTS*

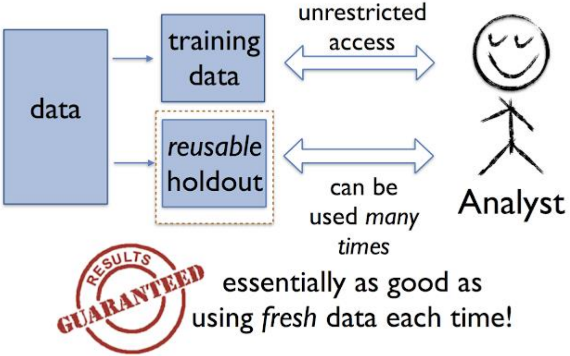
STATISTICS

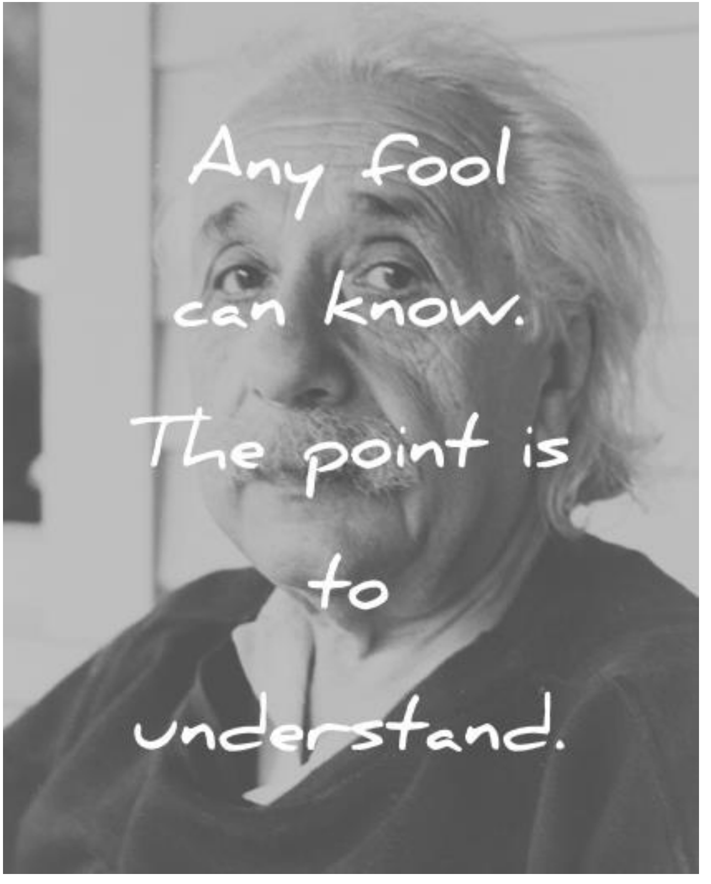
The reusable holdout: Preserving validity in adaptive data analysis

Cynthia Dwork,[1*] Vitaly Feldman,[2*] Moritz Hardt,[3*] Toniann Pitassi,[4*] Omer Reingold,[5*] Aaron Roth[6*]

Misapplication of statistical data analysis is a common cause of spurious discoveries in scientific research. Existing approaches to ensuring the validity of inferences drawn from data assume a fixed procedure to be performed, selected before the data are examined. In common practice, however, data analysis is an intrinsically adaptive process, with new analyses generated on the basis of data exploration, as well as the results of previous analyses on the same data. We demonstrate a new approach for addressing the challenges of adaptivity based on insights from privacy-preserving data analysis. As an application, we show how to safely reuse a holdout data set many times to validate the results of adaptively chosen analyses.

Reusable holdout method




Any fool
can know.
The point is
to
understand.

The Alan Turing Institute

turing.ac.uk
@turinginst