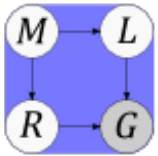
21st Century Physical Sciences
The Age of Algorithms?

Stephen Roberts

Machine Learning Research Group,
University of Oxford

UNIVERSITY OF OXFORD



ATI summit 13/1/16

The Algorithm

An **algorithm** is a self-contained step-by-step set of operations to be performed.

Algorithms exist that perform calculation, data processing, and automated reasoning.

The Algorithm

An **algorithm** is a self-contained step-by-step set of operations to be performed.

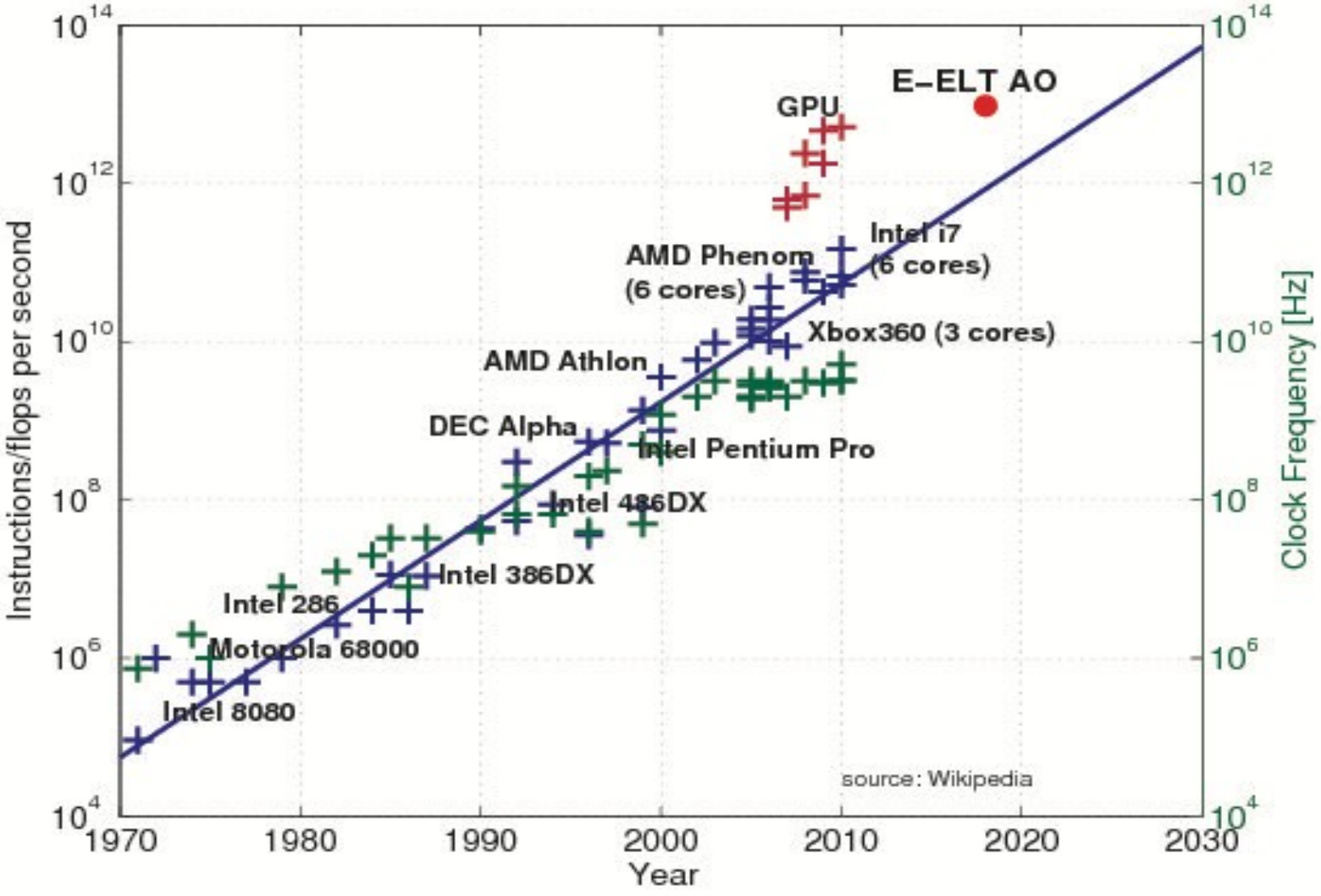
Algorithms exist that perform calculation, **data processing, and automated reasoning.**

The scale of all things data

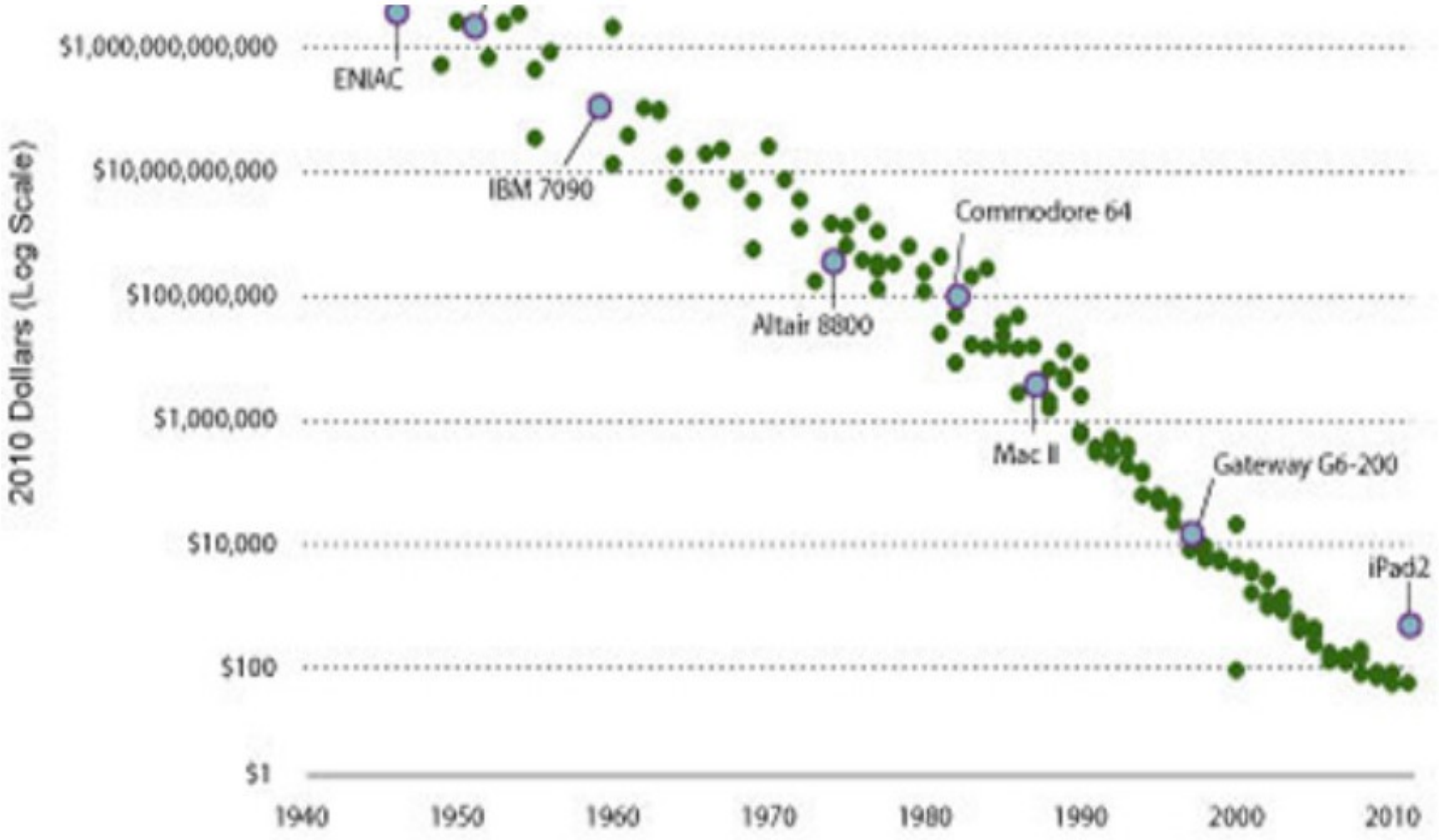
All printed material in the world	200 petabytes (2×10^{17} bytes)
All words ever spoken by human beings	5 exabytes (5×10^{18} bytes)
Predicted internet traffic in 2015	960 exabytes (1×10^{21} bytes)

[UC Berkeley School of Information, 2003; Cisco Visual Networking Index, 2011]

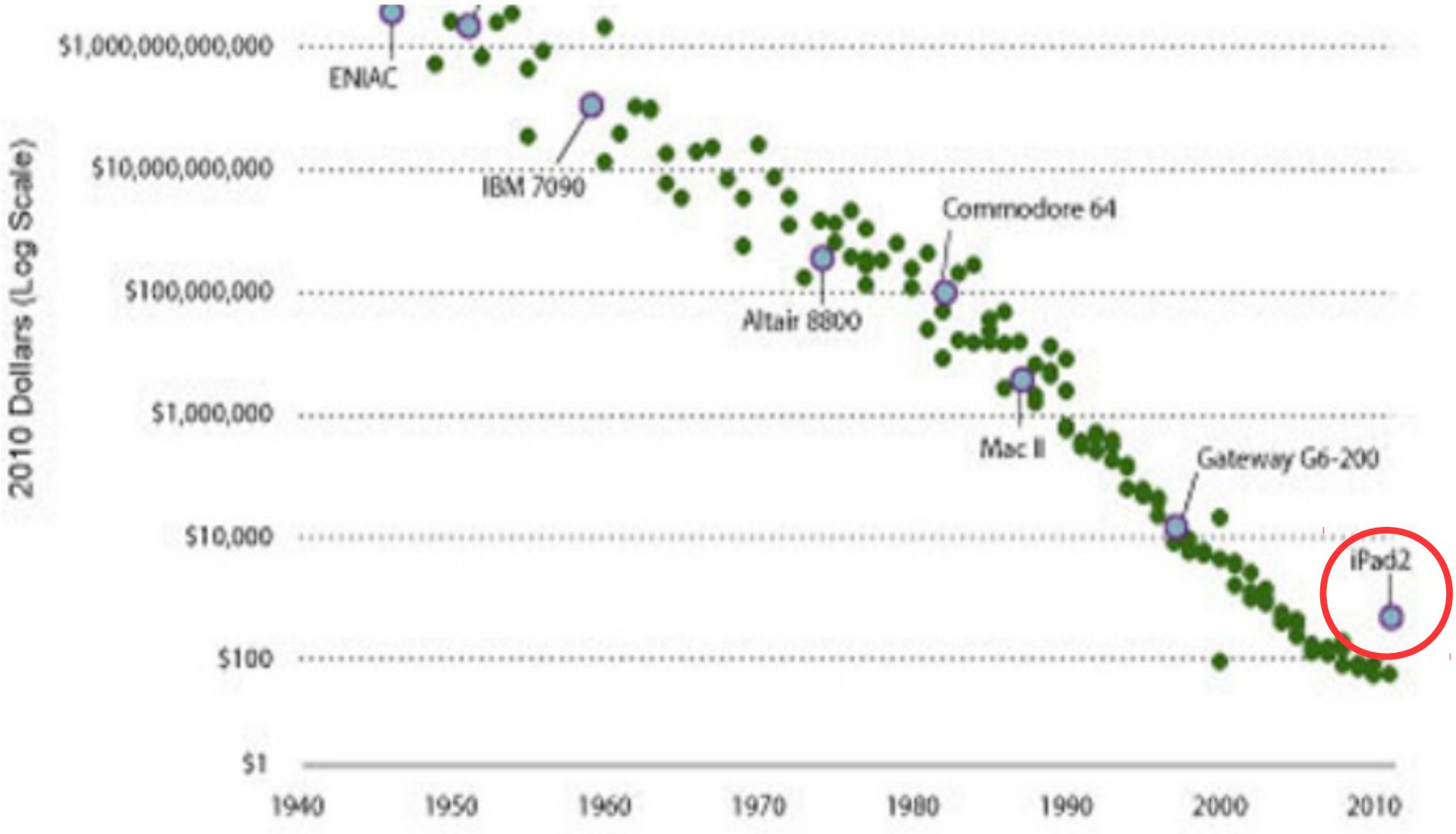
The rise of the flop



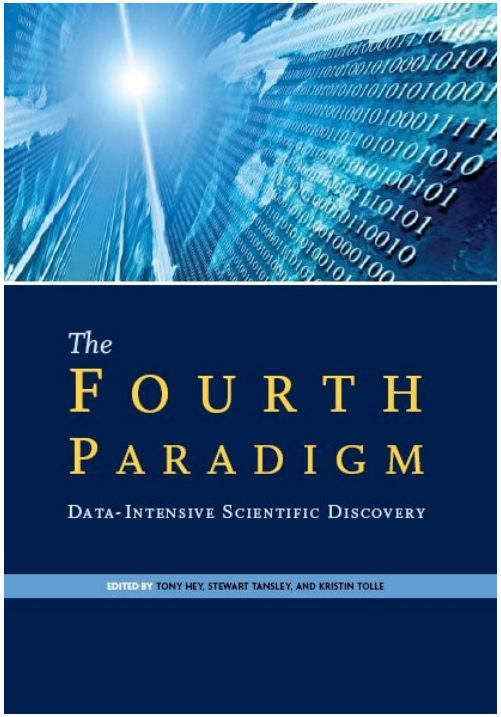
The fall in price



The fall in price



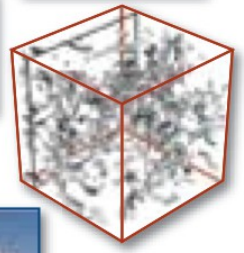
Data analytics as the 4th Paradigm

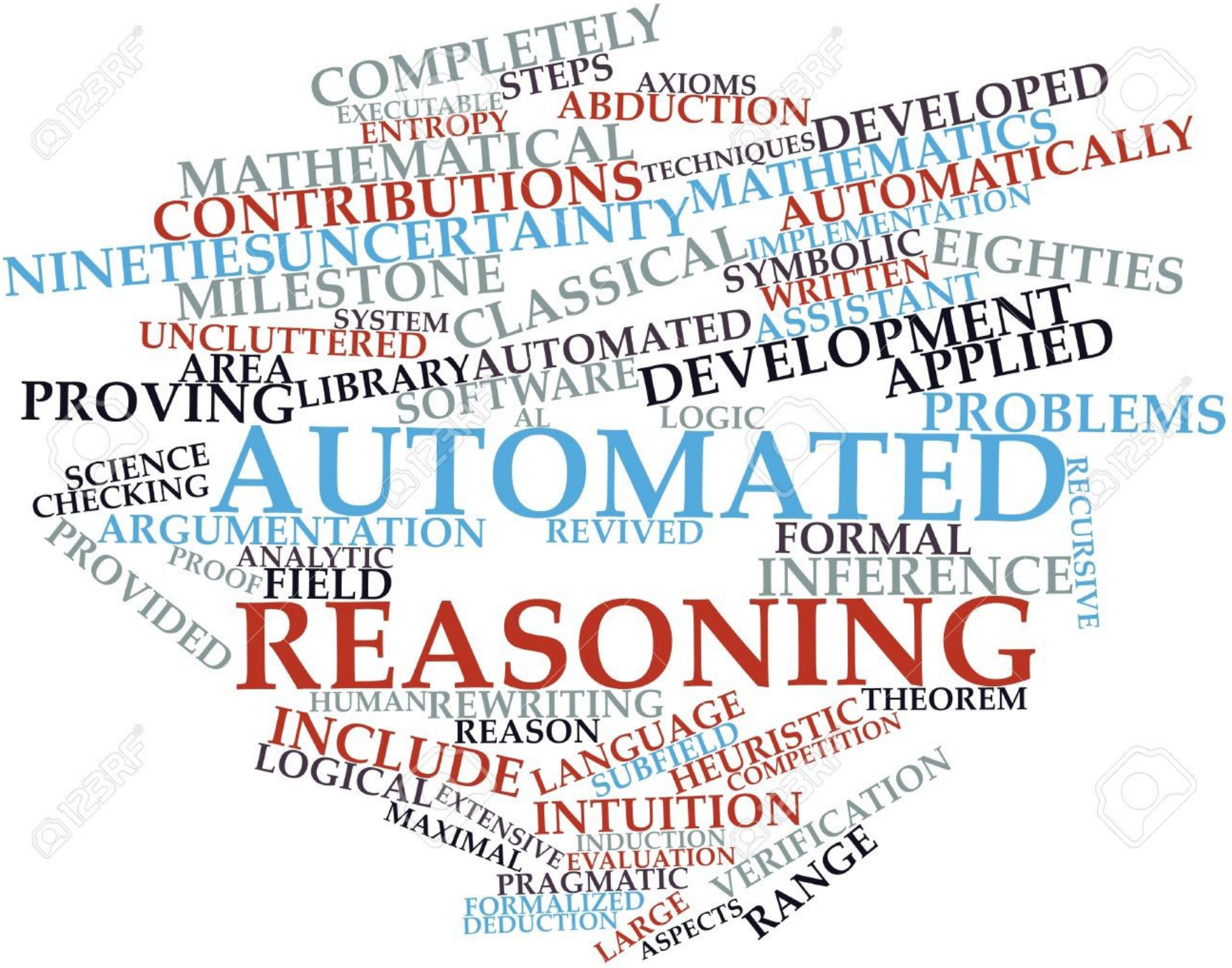


Science Paradigms

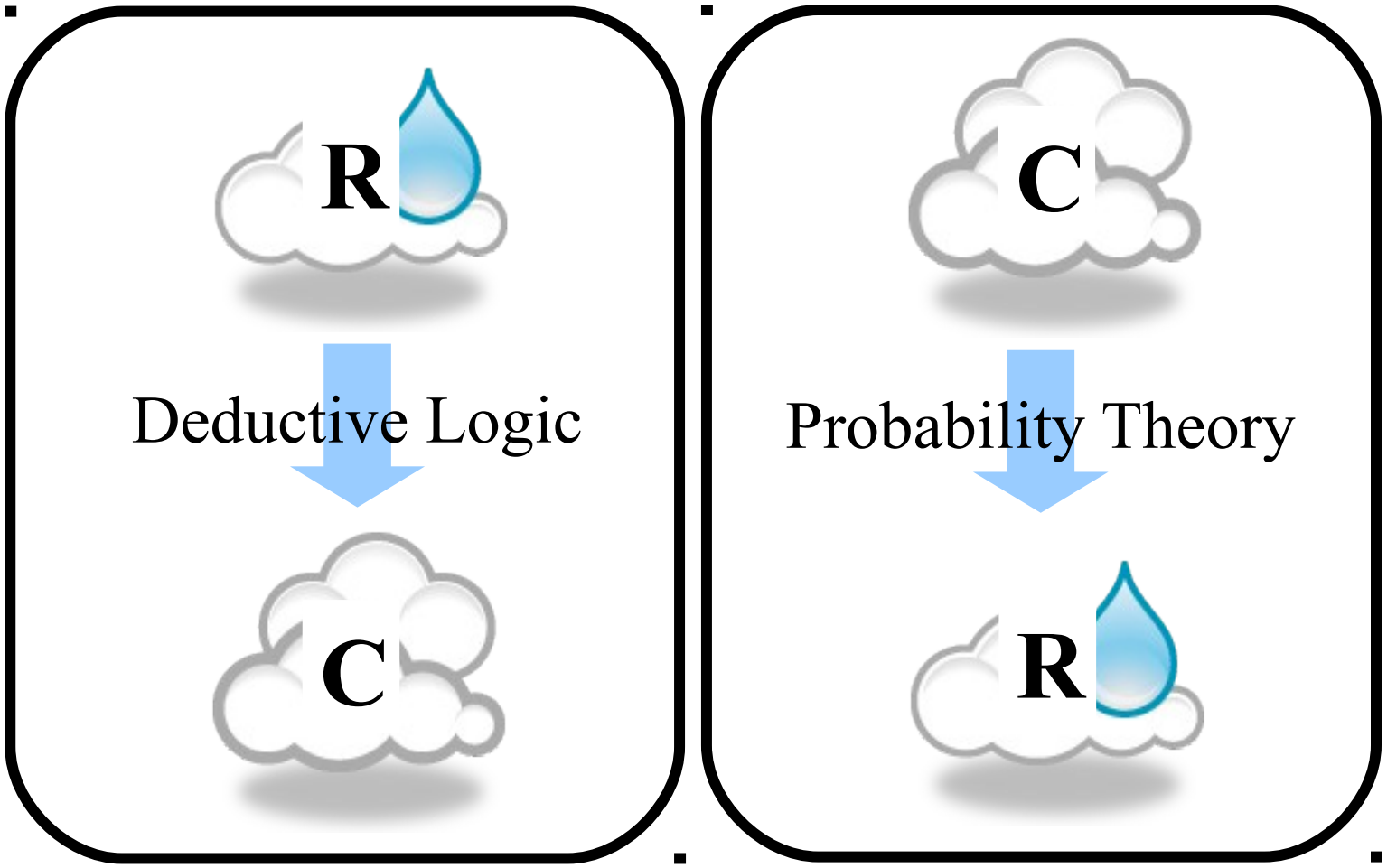
- Thousand years ago: science was **empirical**
- Last few hundred years: **theoretical** branch
- Last few decades: a **computational** branch
- Today: **data exploration** (eScience)
  unify theory, experiment, and simulation

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{4\pi G p}{3} - \mathrm{K}\frac{c^2}{a^2}$$




COMPLETELY
STEPS
AXIOMS
EXECUTABLE
ABDUCTION
ENTROPY
DEVELOPED
MATHEMATICAL
TECHNIQUES
MATHEMATICS
CONTRIBUTIONS
AUTOMATICALLY
NINETIES
UNCERTAINTY
IMPLEMENTATION
SYMBOLIC
MILESTONE
CLASSICAL
EIGHTIES
WRITTEN
ASSISTANT
SYSTEM
UNCLUTTERED
AUTOMATED
DEVELOPMENT
AREA
LIBRARY
SOFTWARE
APPLIED
PROVING
AL
LOGIC
PROBLEMS
AUTOMATED
SCIENCE
CHECKING
RECURSIVE
ARGUMENTATION
REVIVED
FORMAL
PROVIDED
PROOF
ANALYTIC
INFERENCE
FIELD
REASONING
HUMAN
REWRITING
THEOREM
INCLUDE
LANGUAGE
REASON
SUBFIELD
HEURISTIC
COMPETITION
LOGICAL
EXTENSIVE
INTUITION
MAXIMAL
INDUCTION
VERIFICATION
EVALUATION
PRAGMATIC
FORMALIZED
RANGE
DEDUCTION
LARGE
ASPECTS

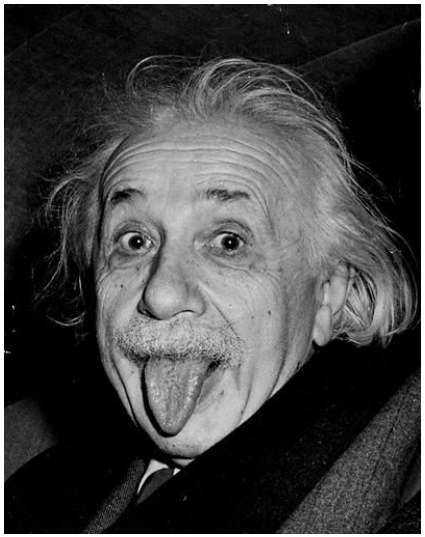
Probability theory represents an extension of traditional logic, *allowing us to reason in the face of uncertainty*



Occam's Razor

Numquam ponenda est pluralitas sine necessitate-
"Plurality must never be posited without necessity"

"Everything should be kept as simple as possible, but no simpler."



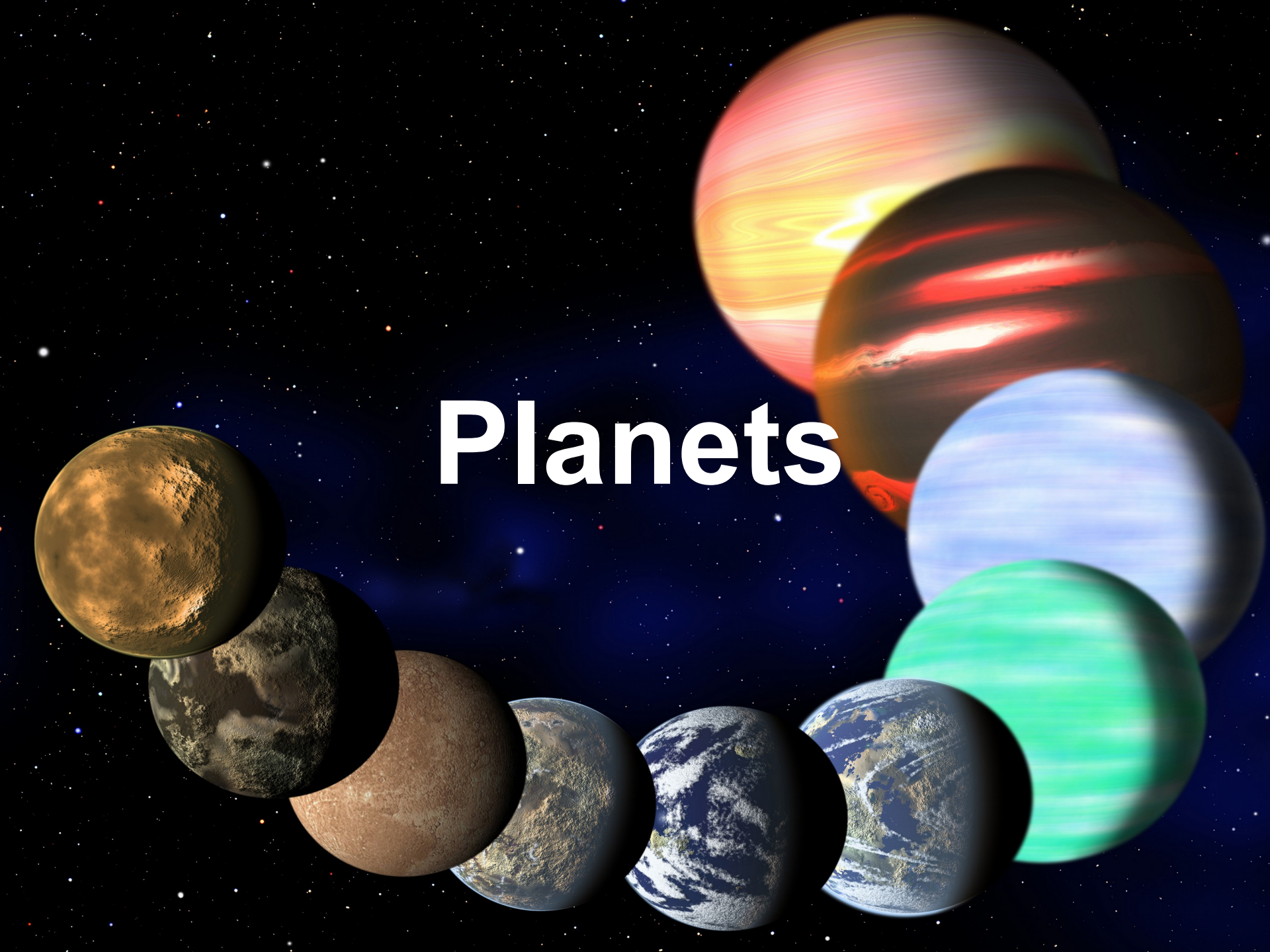
Probabilities over functions not samples

Why is this important?

Enables us to

- Extract *latent* representations from data
- Automate the process of selecting the right complexity of model
- Select *where, when & what* we observe to maximise informativeness (Bayesian experimental design)
- Aggregate and fuse disparate knowledge sources

- Extracting latent trends
- Combining knowledge
- Finding anomalies and changes


Planets

Kepler space telescope

(2009 – 2014, then...)

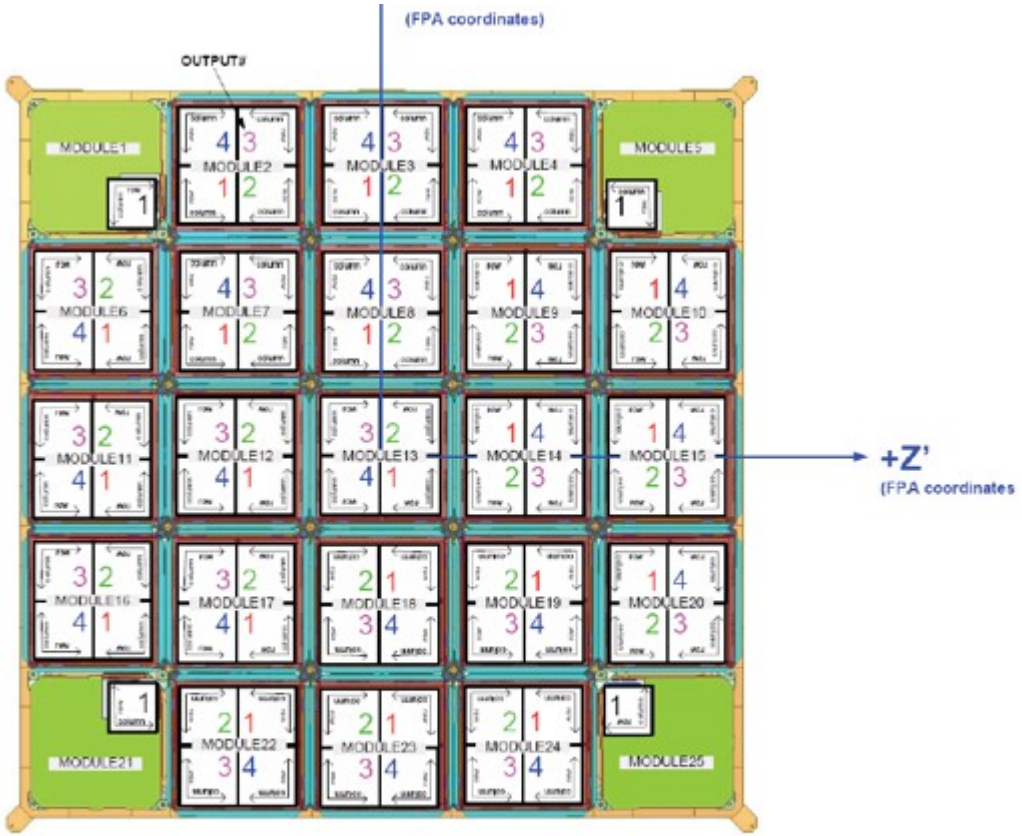
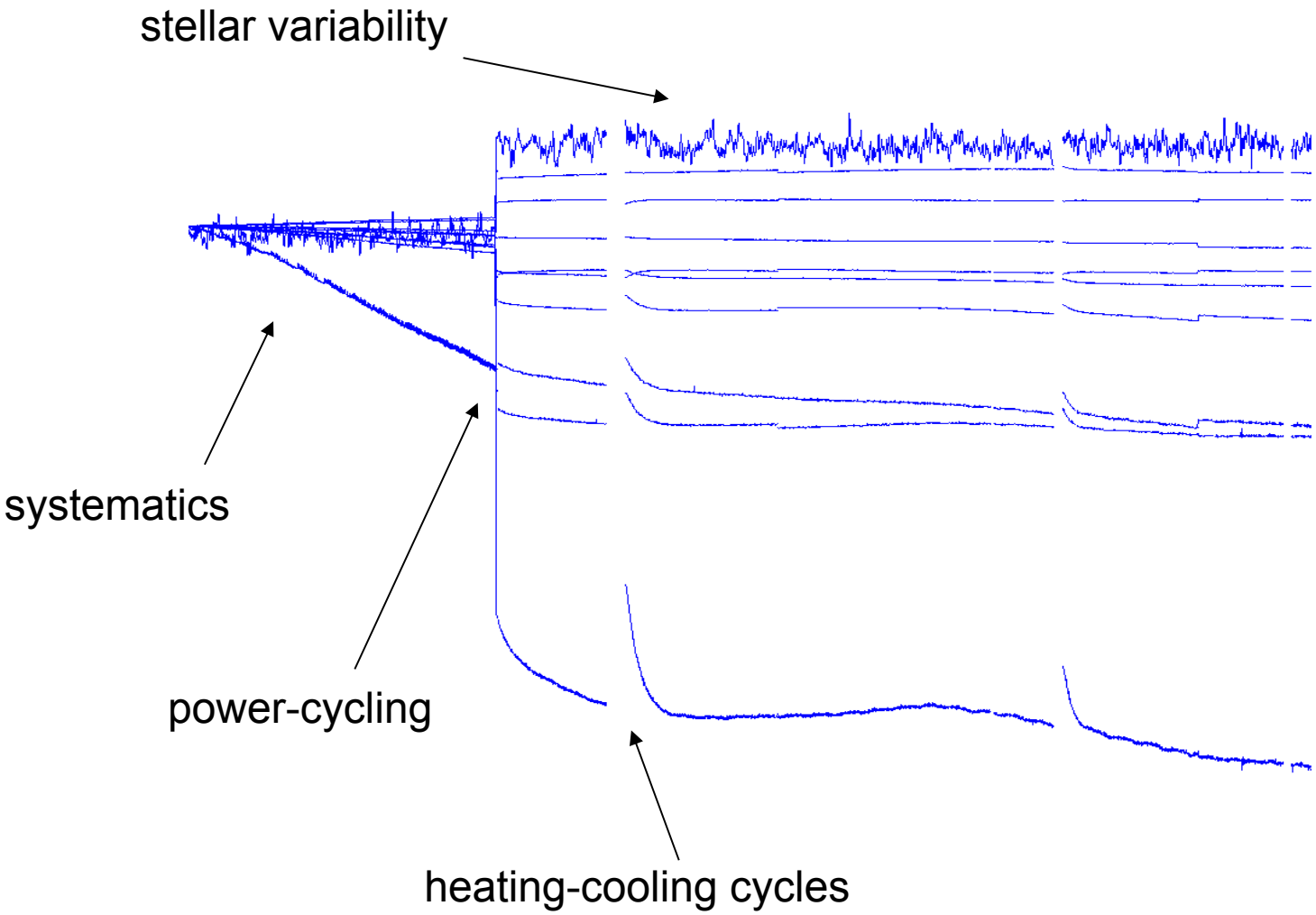
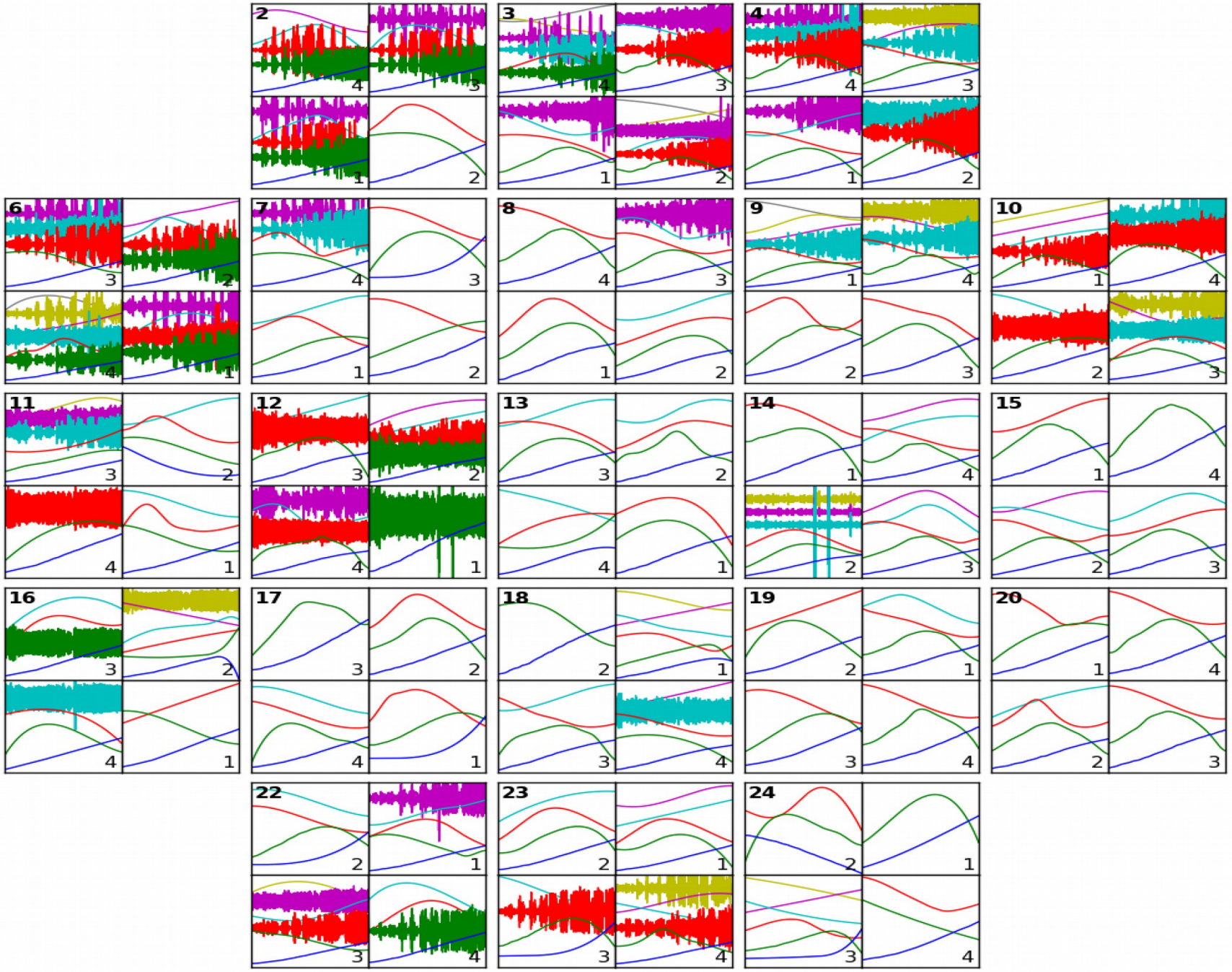


Figure 24: Focal plane layout, labeling modules and outputs (1-4), and the directions of rows and columns. Note that the focal plane is symmetric under 90 degree rotations, with the exception of the central module, module 13. Modules 1, 5, 21, and 25 are FGS modules.

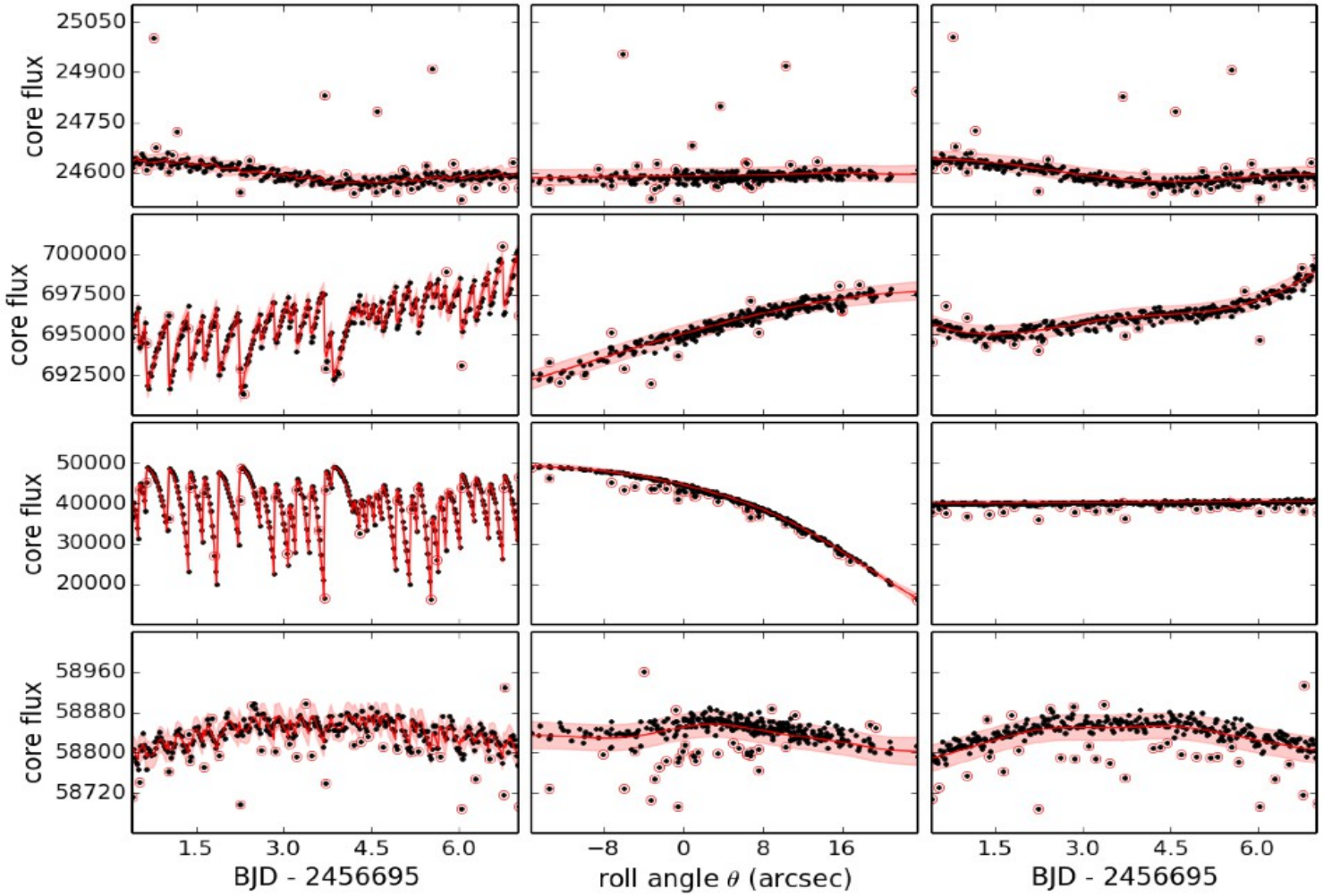
Kepler data



(Joint work with Suzanne Aigrain)


2
3
4
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
22
23
24

After the failure of reaction wheels, Kepler has moved to the K-2 mission- even larger systematics – requiring star by star analysis




People

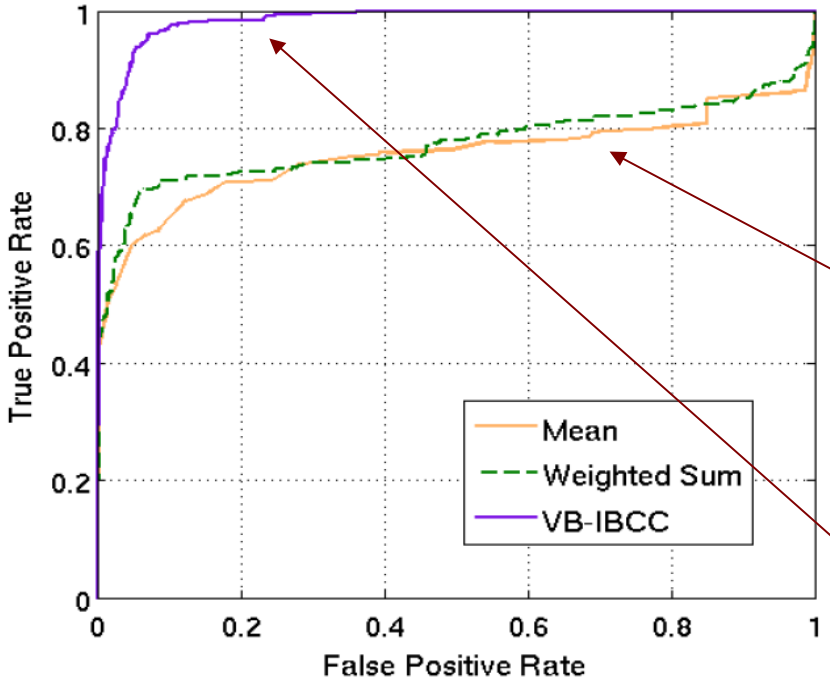
Going back to the roots of computing

Zooniverse


Gmail - Inbox - sjrob.ne
Pattern Analysis and M
Zooniverse - Projects
www.zooniverse.org/projects
Google
Gmail: Email fro...
Pattern Analysis ...
Home — Depart...
Other Bookmarks
ZOONIVERSE
REAL SCIENCE ONLINE
HOME
PROJECTS
ABOUT
EDUCATION
BLOGS
RESEARCHERS
CONTACT
Ancient Lives
Help transcribe ancient papyri, written over 1,000 years ago by the citizens of Oxyrhynchu, the City of the Sharp-Nosed. Everyday life and writing, that we need your help to uncover.
JOIN IN
ANCIENT LIVES
Ice Hunters
Help to discover Kuiper Belt Objects with just the right orbit and just the right characteristics to make them eligible for a visit from the New Horizons mission.
JOIN IN
ICEHUNTERS.org
Planet Hunters
Using public data from NASA's Kepler mission, we are looking for planets around other stars.
JOIN IN
planethunters.org
The Milky Way Project
Sorting and measuring our galaxy, the Milky Way; we're asking you to help us find and draw bubbles in beautiful infrared data from the Spitzer Space Telescope.
THE MILKY WAY
Live Projects
ANCIENT LIVES
oldWeather
ICEHUNTERS.org
planethunters.org
THE MILKY WAY PROJECT
MOON ZOO
GALAXY ZOO HUBBLE
SOLAR STORMWATCH
GALAXY ZOO UNDERSTANDING COSMIC MERGERS
GALAXY ZOO THE HUNT FOR SUPERNOVAE
ZOOSHOP
WEAR THE BADGE
SPREAD THE WORD !
GO SHOPPING

Zooniverse : results

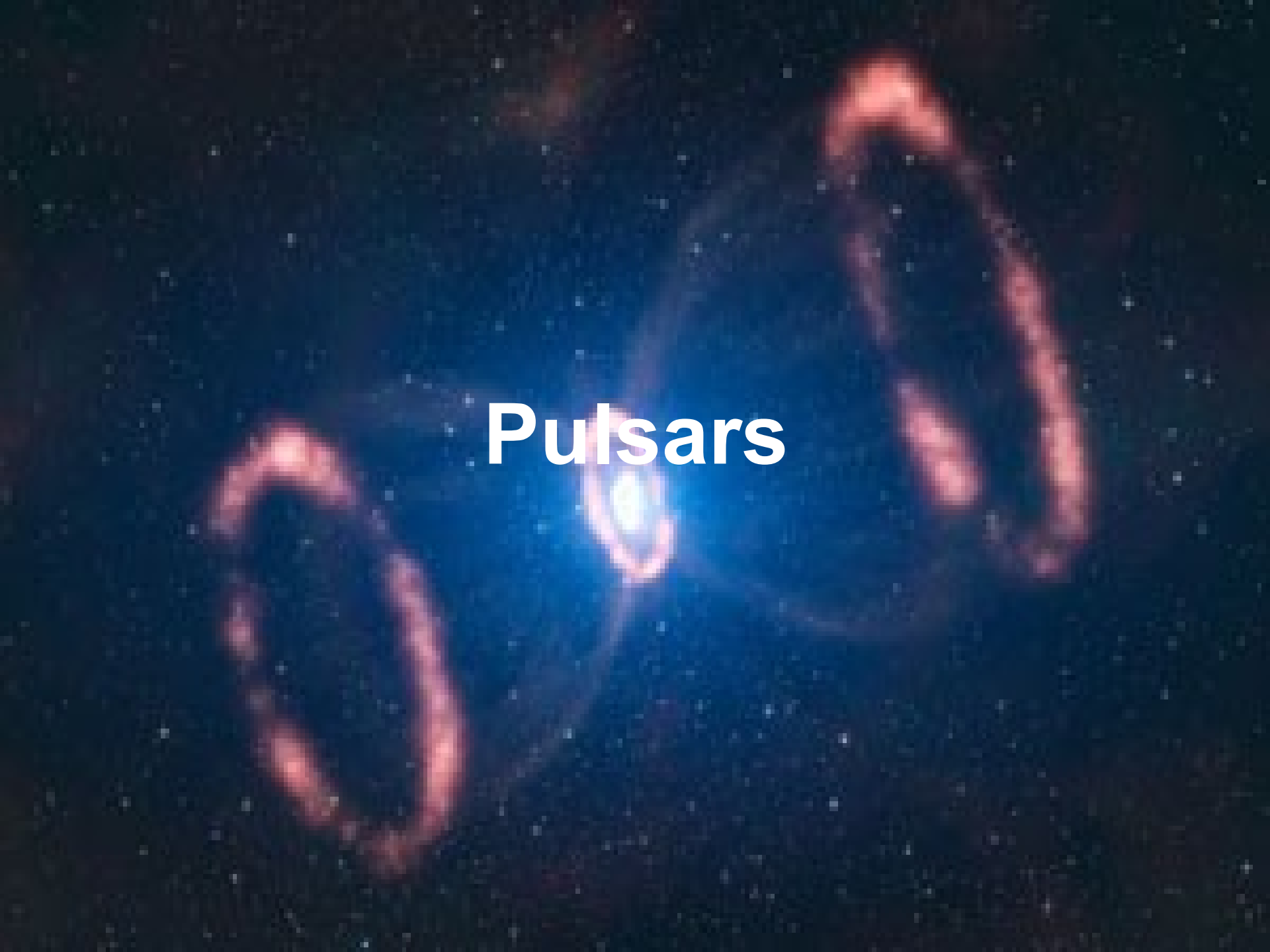
- Bayesian combination of 1000s of weak decision agents (Zooniverse volunteers)



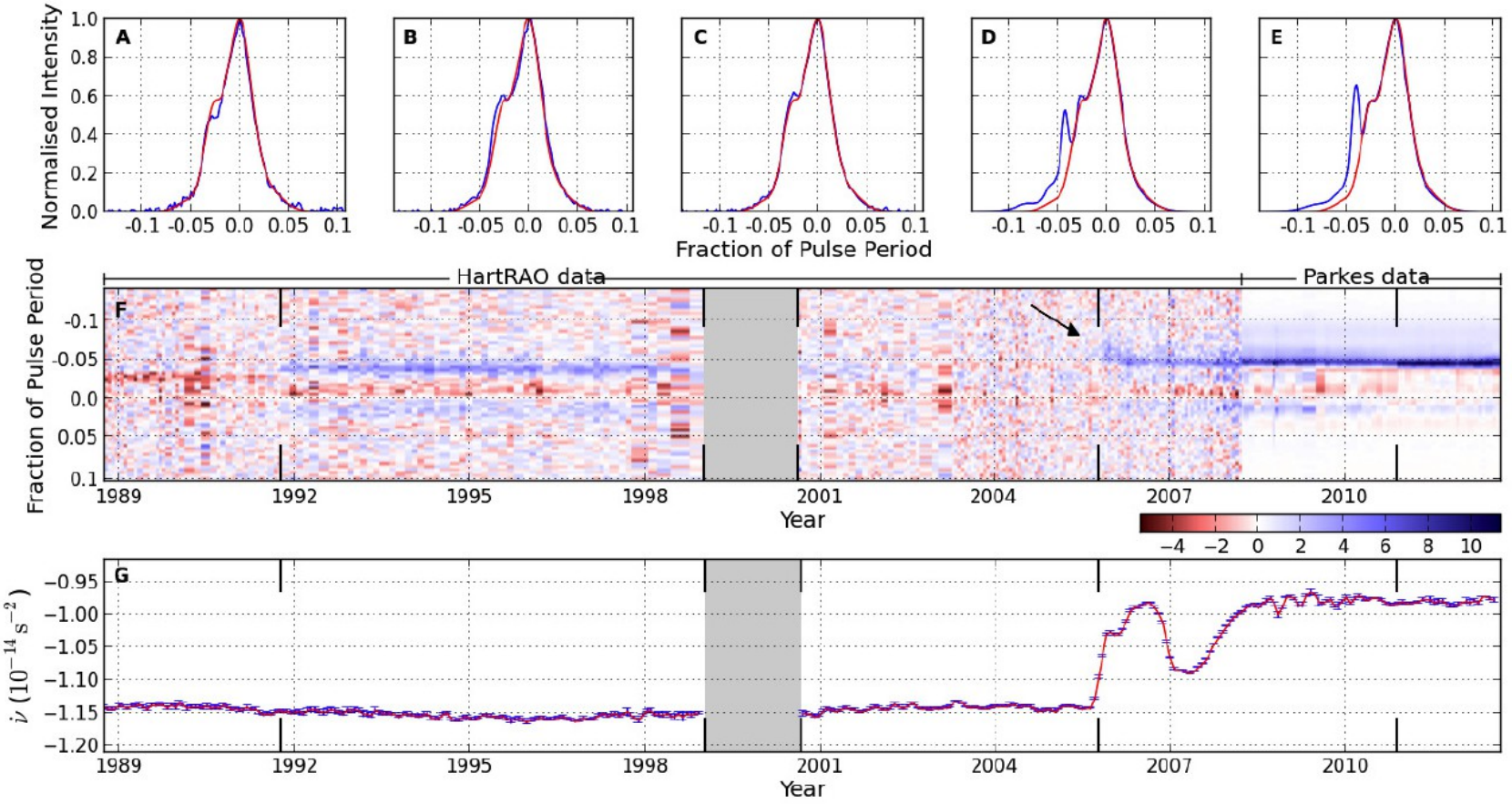
Huge performance improvement over current approach (mean scores)

Method	AUC
Mean of Scores	0.7543
Weighted Sum	0.7722
Simple Majority Voting	0.7809
Weighted Majority Voting	0.7378
Gibbs-IBCC	0.9127
VB-IBCC	0.9840

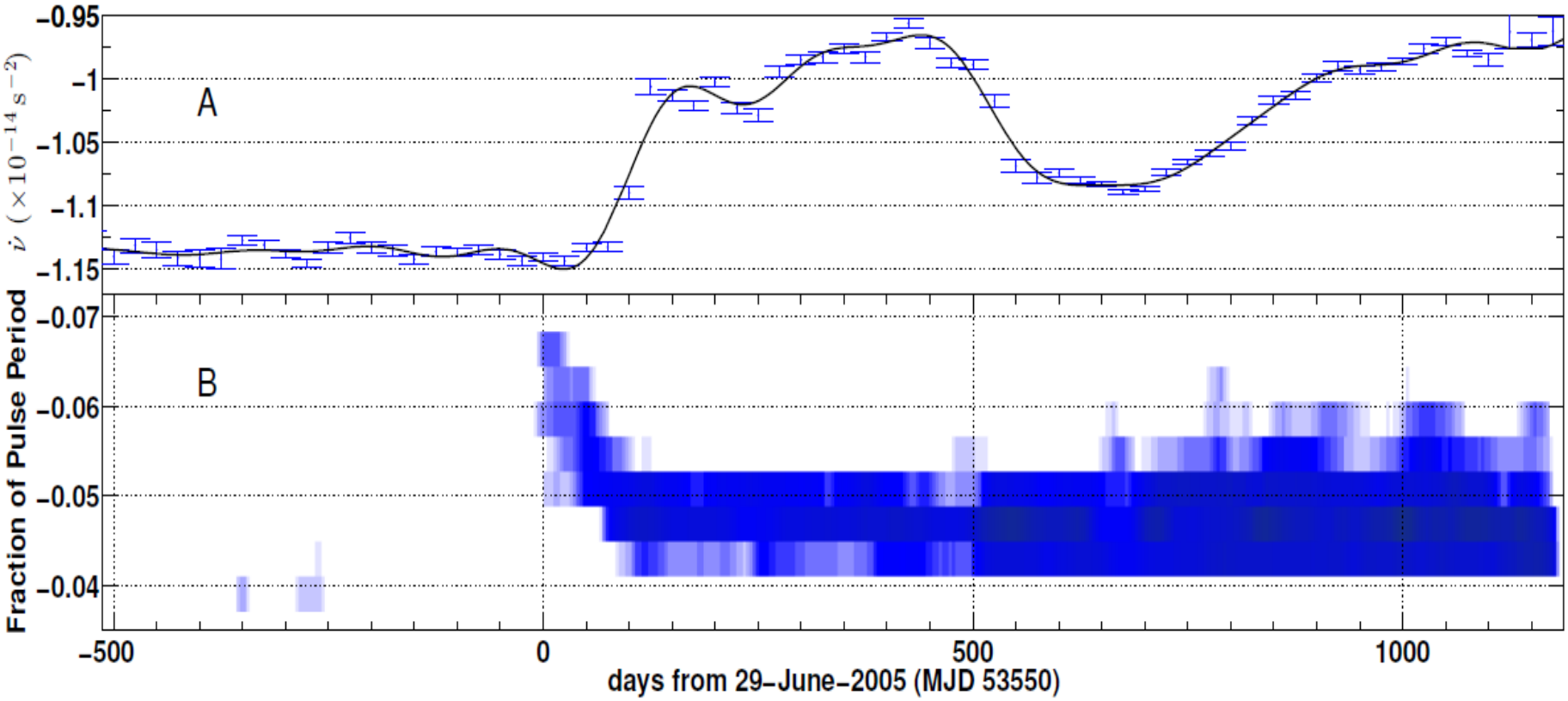
(Joint work with Edwin Simpson, Chris Lintott)


Pulsars

PSR J0738-4042



(Joint work with Aris Karastergiou & Paul Brook)



The Astrophysical Journal Letters, 780:L31 (5pp), 2014 January 10 doi:10.1088/2041-8205/780/2/L31

EVIDENCE OF AN ASTEROID ENCOUNTERING A PULSAR

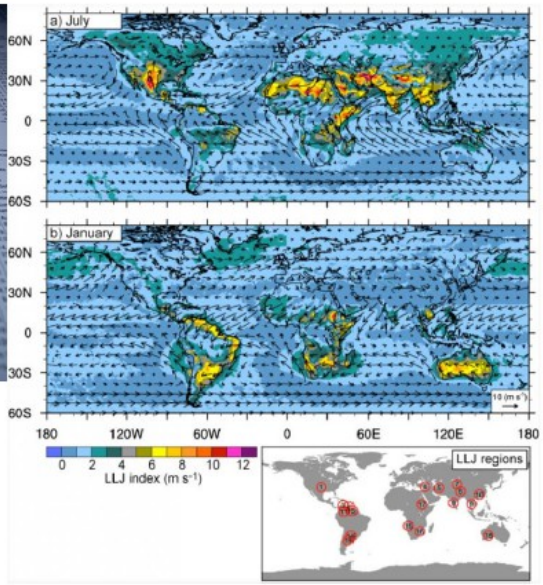
P. R. Brook[1,5], A. Karastergiou[1], S. Buchner [2,6], S. J. Roberts[3], M. J. Keith [4,7], S. Johnston[4], and R. M. Shannon[4]
[1] Astrophysics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UK; paul.brook@astro.ox.ac.uk
[2] Hartebeesthoek Radio Astronomy Observatory, P.O. Box 443, Krugersdorp 1740, South Africa
[3] Information Engineering, University of Oxford, Parks Road, Oxford OX1 3PJ, UK
[4] CSIRO Astronomy and Space Science, Australia Telescope National Facility, P.O. Box 76, Epping, NSW 1710, Australia
Received 2013 October 20; accepted 2013 November 13; published 2013 December 19

So – where are we going?

Data analytics

Beyond-human-scale knowledge extraction

physics, medicine & healthcare, biology, finance & economics, energy, social science, customer-facing e-commerce, digital marketing

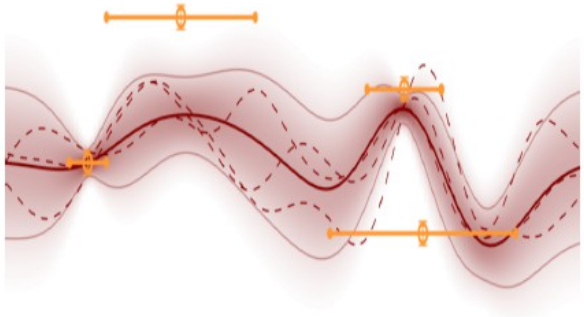


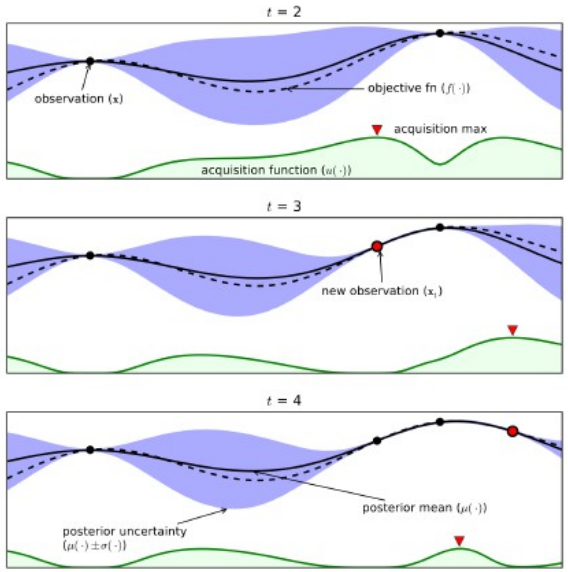
Intelligent systems

New methods of computing under uncertainty

probabilistic programming, probabilistic numerics, scalable intelligent optimisation, automated algorithm creation

```
[assume initial-state-dist (list (/ 1 3) (/ 1 3) (/ 1 3))]
[assume get-state-transition-dist (lambda (s)
  (cond ((= s 0) (list .1 .5 .4)) ((= s 1) (list .2 .2 .6))
        ((= s 2) (list .15 .15 .7)))))]
[assume transition (lambda (prev-state)
  (discrete (get-state-transition-dist prev-state)))]
[assume get-state (mem (lambda (index)
  (if (<= index 0) (discrete initial-state-dist)
                   (transition (get-state (- index 1)))))))]
[assume get-state-observation-mean (lambda (s)
        (cond ((= s 0) -1) ((= s 1) 1) ((= s 2) 0)))]
[observe (normal (get-state-obs-mean (get-state 1)) 1) .9]
[observe (normal (get-state-obs-mean (get-state 2)) 1) .8]
⋮
[observe (normal (get-state-obs-mean (get-state 16)) 1) -1]
[predict (get-state 0)]
[predict (get-state 1)]
⋮
[predict (get-state 16)]
⋮
```



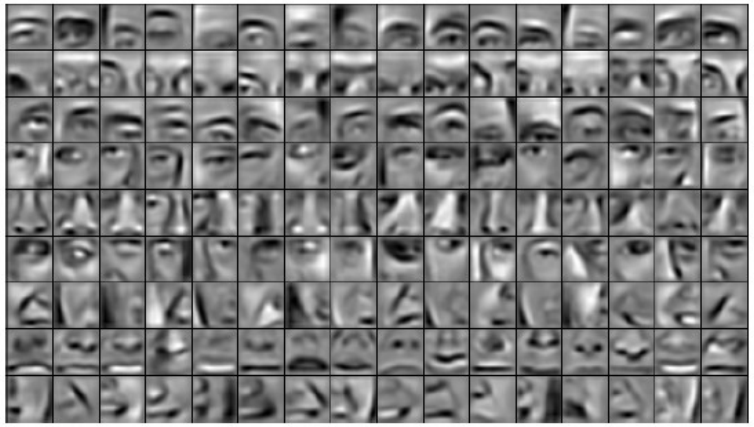


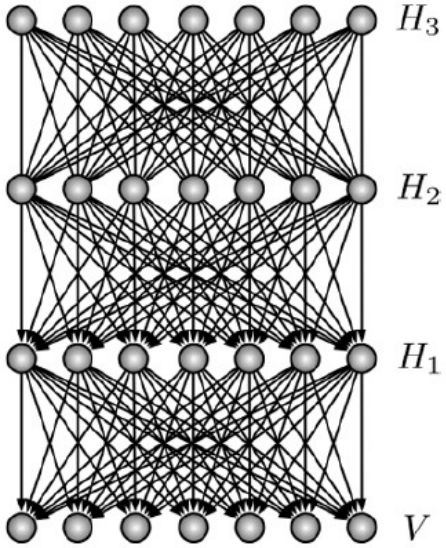
Intelligent systems

Strength in depth - vision, speech and beyond: understanding AI

deep understanding of deep learning, vision & beyond vision, cognitive systems with understanding of the world

Interpreting scenes
find what and where

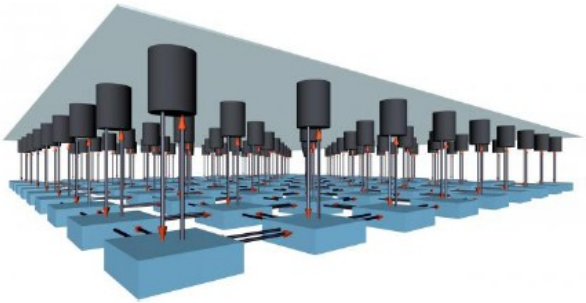


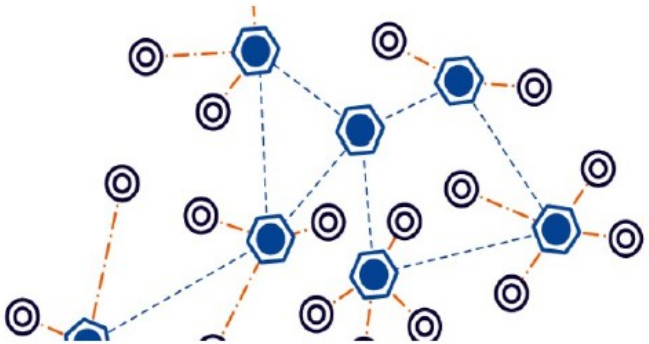


Intelligent systems

Autonomy at scale, intelligent sensing and situation awareness

unifying heterogeneous data at scale, ultra-large machine sensing, autonomous systems, deciding and acting in a complex uncertain world, human-in-the-loop systems

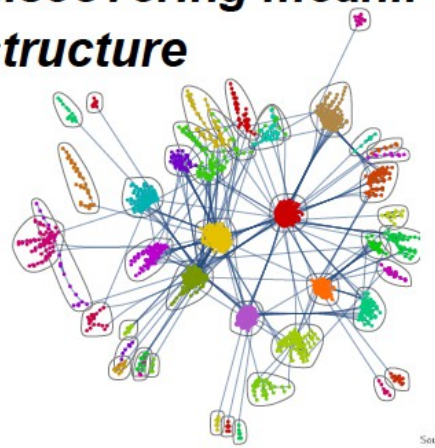


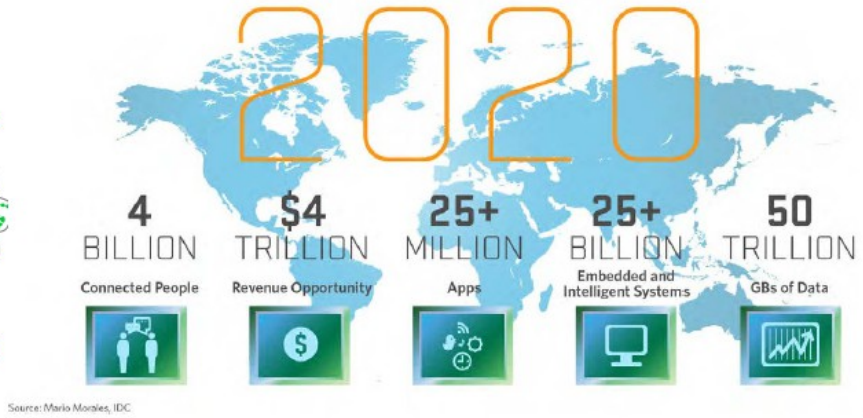


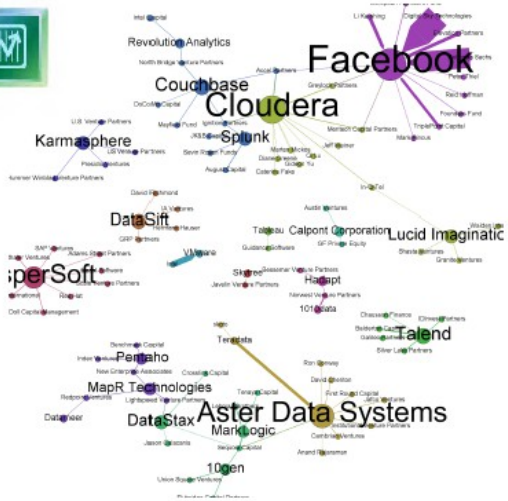
Complex systems

Cyber-physical systems at IoT scale, quantum computation in the wild

understanding and usage of beyond-human-scale dynamic networks, realising the potential of massive-scale connected data, new ways of discovering meaning, topology & homology, understanding multi-scale structure

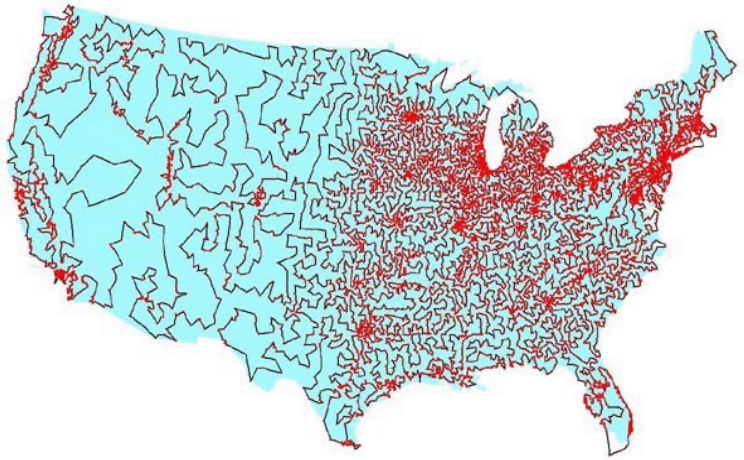


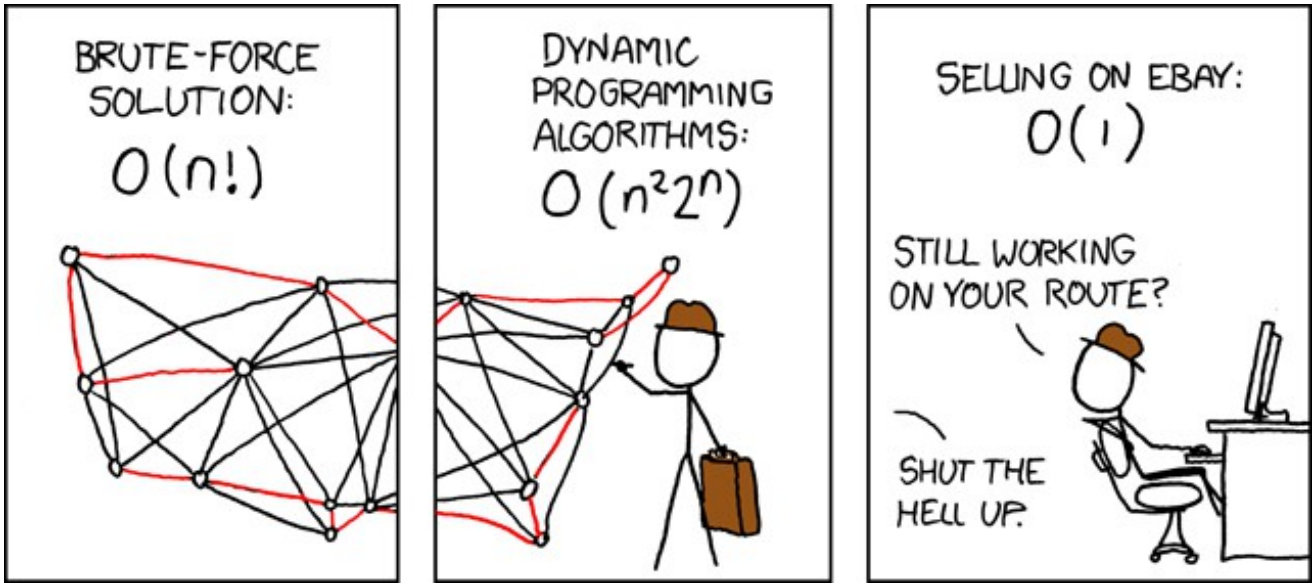




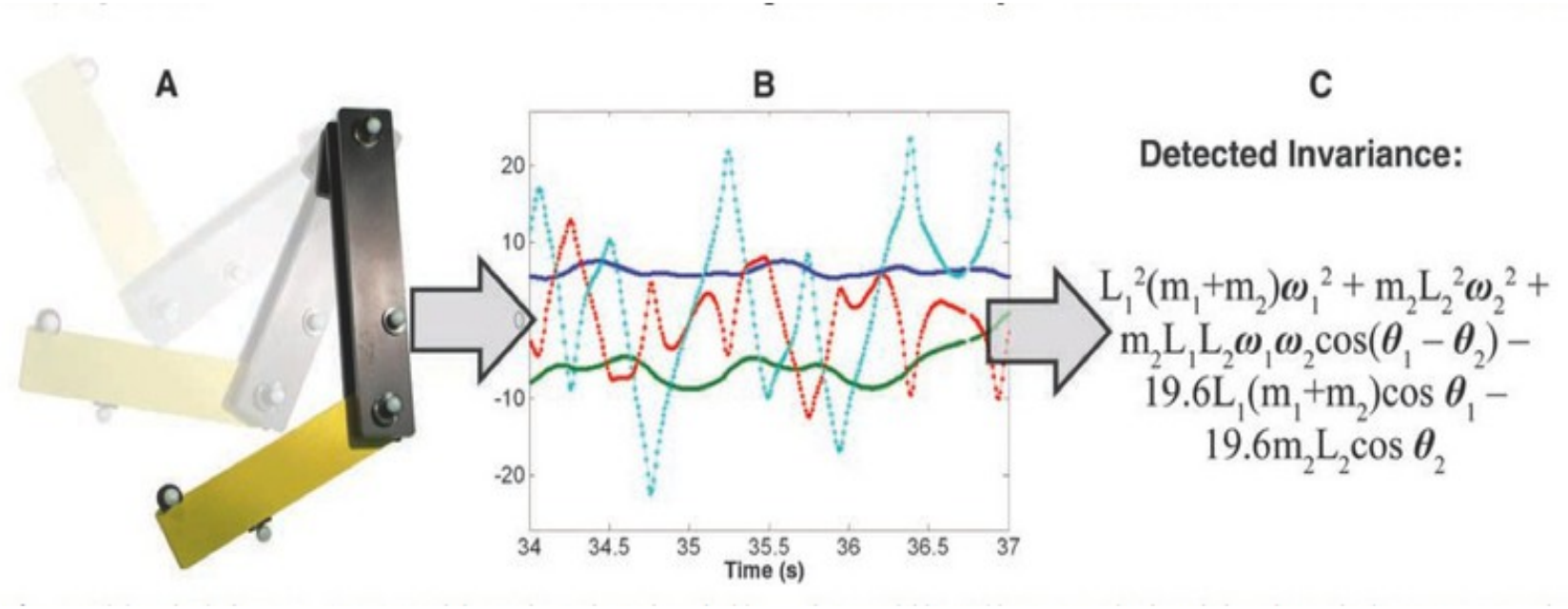
Quantum Computation

A revolution in progress?



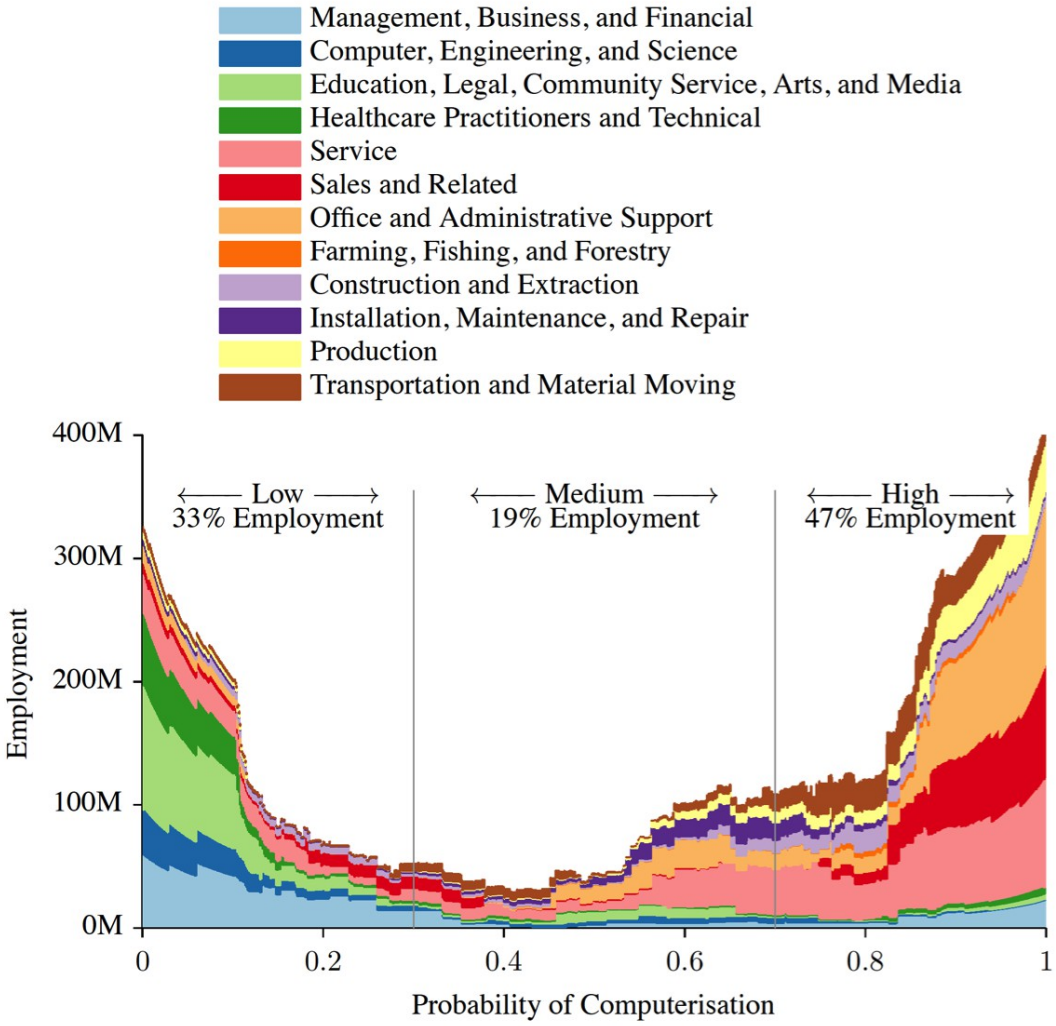


Extracting Physics from data



From Langley et al. 2009

Automation is happening now...



The future of science?

Automated systems can review data, extract meaning & find explanations faster & at scales that we can only dream of

The era of human science has ended.

The era of the Automated Scientist has begun.

With particular thanks to

Mike Osborne, Suzanne Aigrain, Aris Karastergiou, Chris Lintott, Matt Jarvis, Edwin Simpson, Steve Reece

and, of course, **the algorithm...**