

Beyond The Data

- 1. Opening the process of generating science**
- 2. From data centres to computing centres**
- 3. Accompanying the 4th industrial revolution**

Roland Walter

Astronomy Department



**UNIVERSITÉ
DE GENÈVE**

Opening the Process of Generating Science

→ Computing centres

→ Federation of data centres

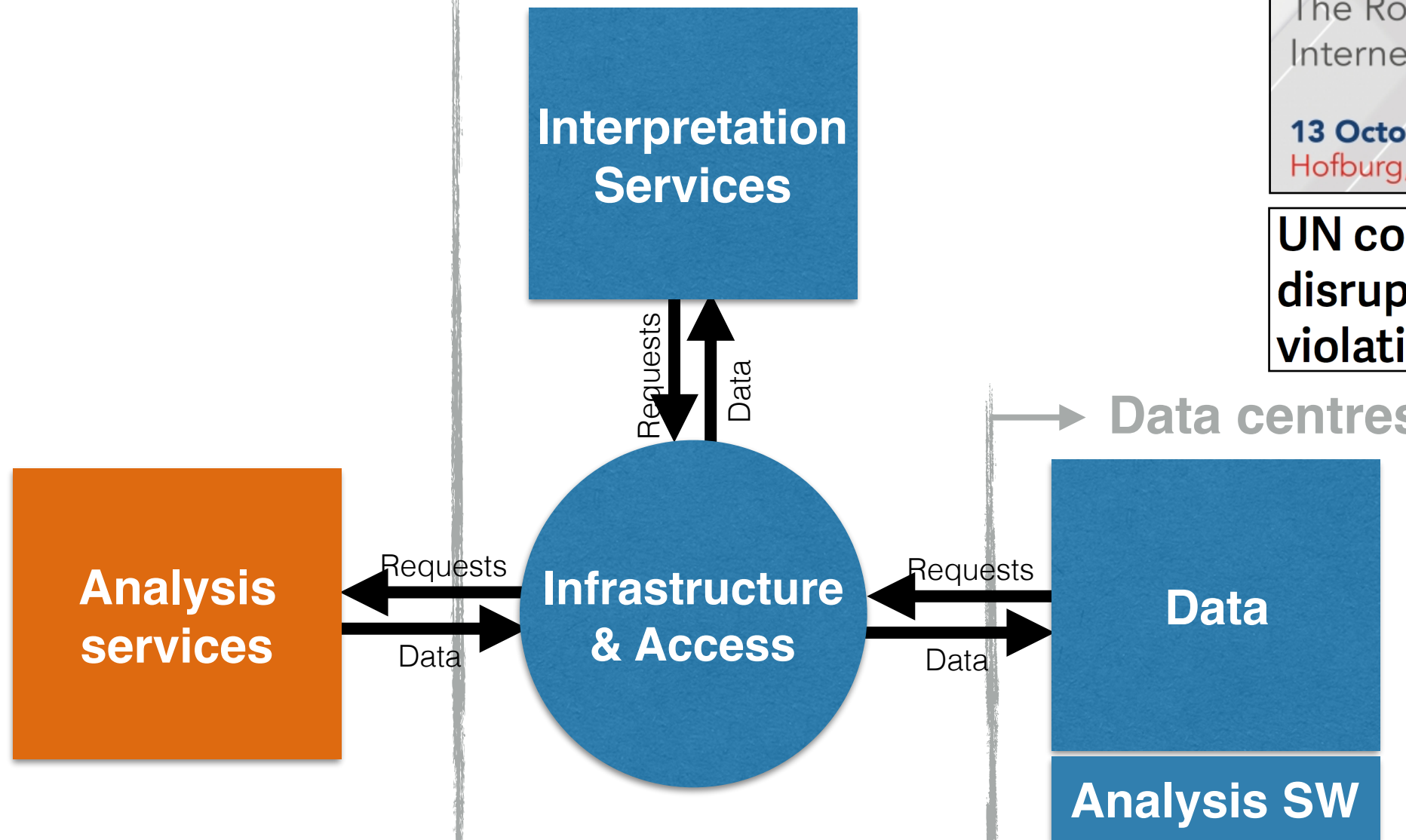
INTERNET FREEDOM CONFERENCE

The Role and Responsibilities of
Internet Intermediaries

13 October 2017

Hofburg, Vienna

**UN condemns internet access
disruption as a human rights
violation**



**The FCC revealed its plan to repeal net neutrality.
It could change how we use the Internet.**

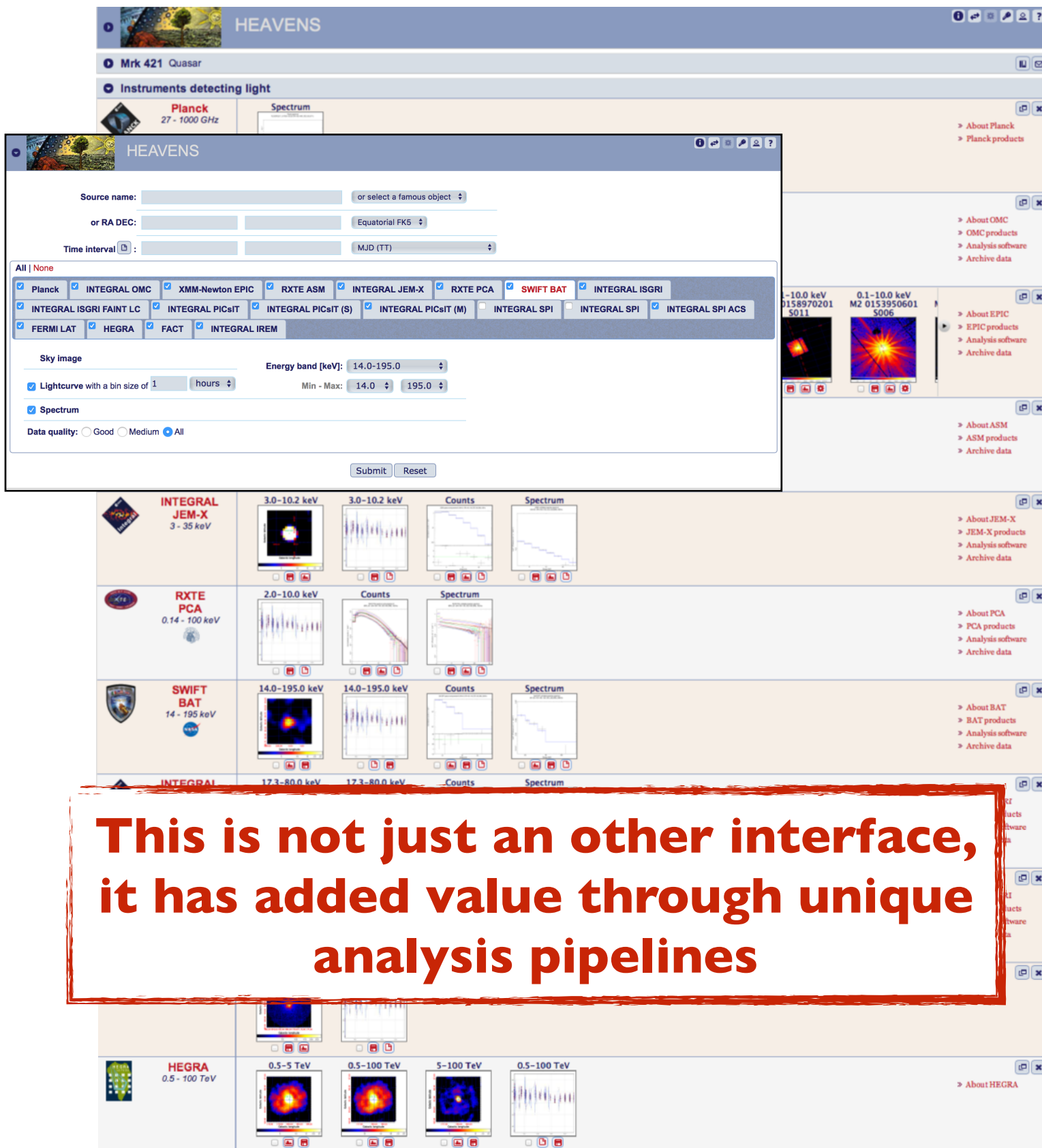
Mike Snider

USA TODAY

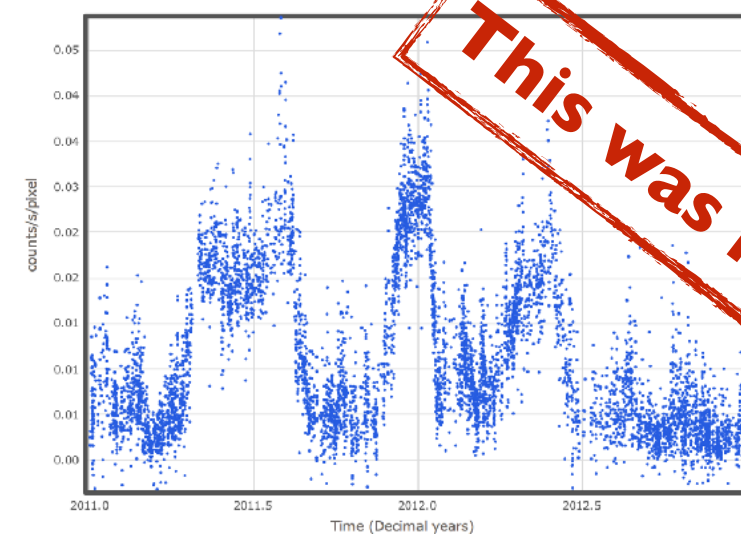
Published 11:41 a.m. ET Nov. 21, 2017

Updated 6:42 p.m. ET Nov. 21, 2017

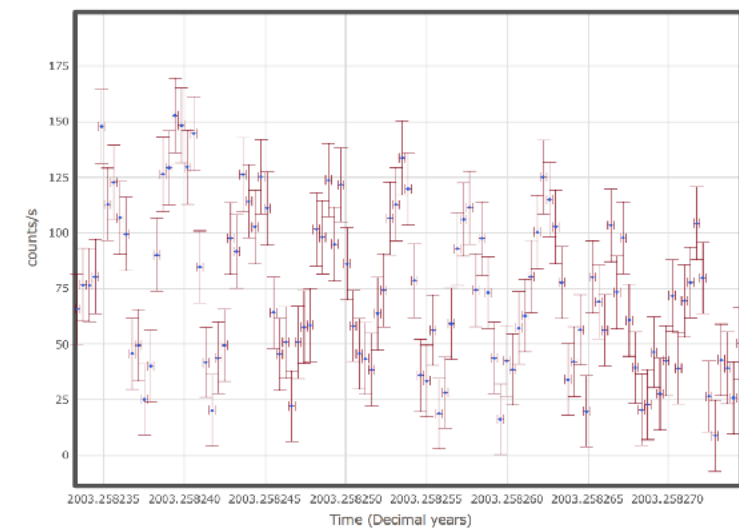
HEAVENS <http://www.isdc.unige.ch/heavens/>



**This is not just an other interface,
it has added value through unique
analysis pipelines**



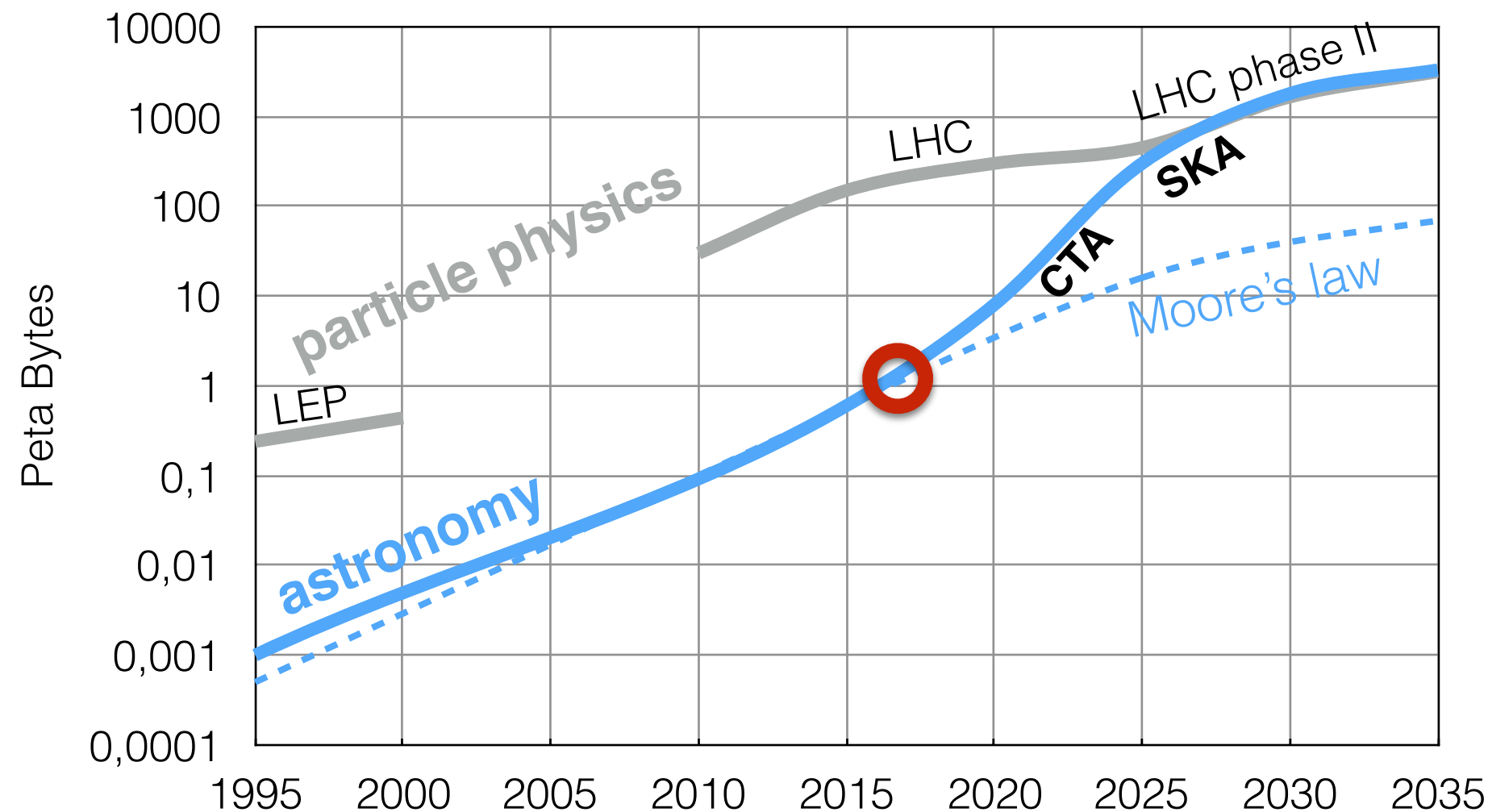
This was in 2005



One can go much further !

Driven by the needs of science,
education and outreach

From Data Centres to Computing Centres



INTEGRAL	0.001 PB/y
2001-	
CTA	12 PB/y
2019-	
LSST	15 PB/y
2022-	
SKA	300 PB/y
2025?-	
LEP	0.04 PB/y
1989-2000	
LHC	30 PB/y
2010-	

———— Distributed data centres —————→

←———— Centralised (“cloud”) computing centres

Distributed **data** centres are not cost effective.
In the next years **services** are likely to be integrated in large **computing** centres supporting experiments generating Big Data.

From Data Centres to Computing Centres



Large computing centres exist today

Cray XC 50 at the Swiss National Supercomputing Centre (Lugano)

361760 cores equivalent

6500 Tesla GPU

6 PB scratch disks

200 PB tape storage

100 Gbps internet connection

Only European machine in Top10

#3 in Top500

#6 in TopGreen

>100x typical University computing

Evolution of ~~Data~~ Knowledge Management

- **People want results and data**
- **Generating results shall not require specific knowledge**
- **Results shall be tailored to the user's (scientists, school, public) needs**
- **Results shall be related to the physics, not to observations/instruments**

Services need to add scientific value to the data

Evolution of Knowledge Management

Volume
Computing

Knowledge

Interpretation

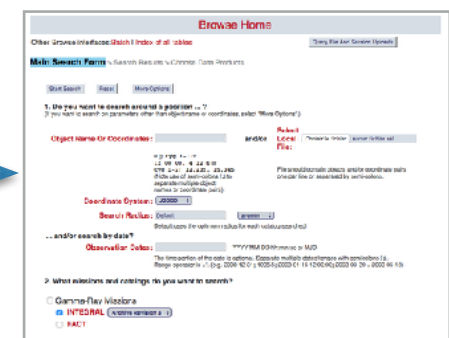
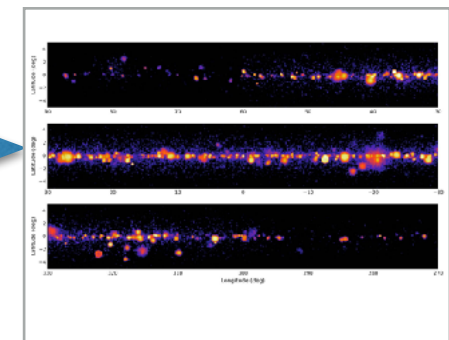
Analysis

Data & Information

2025

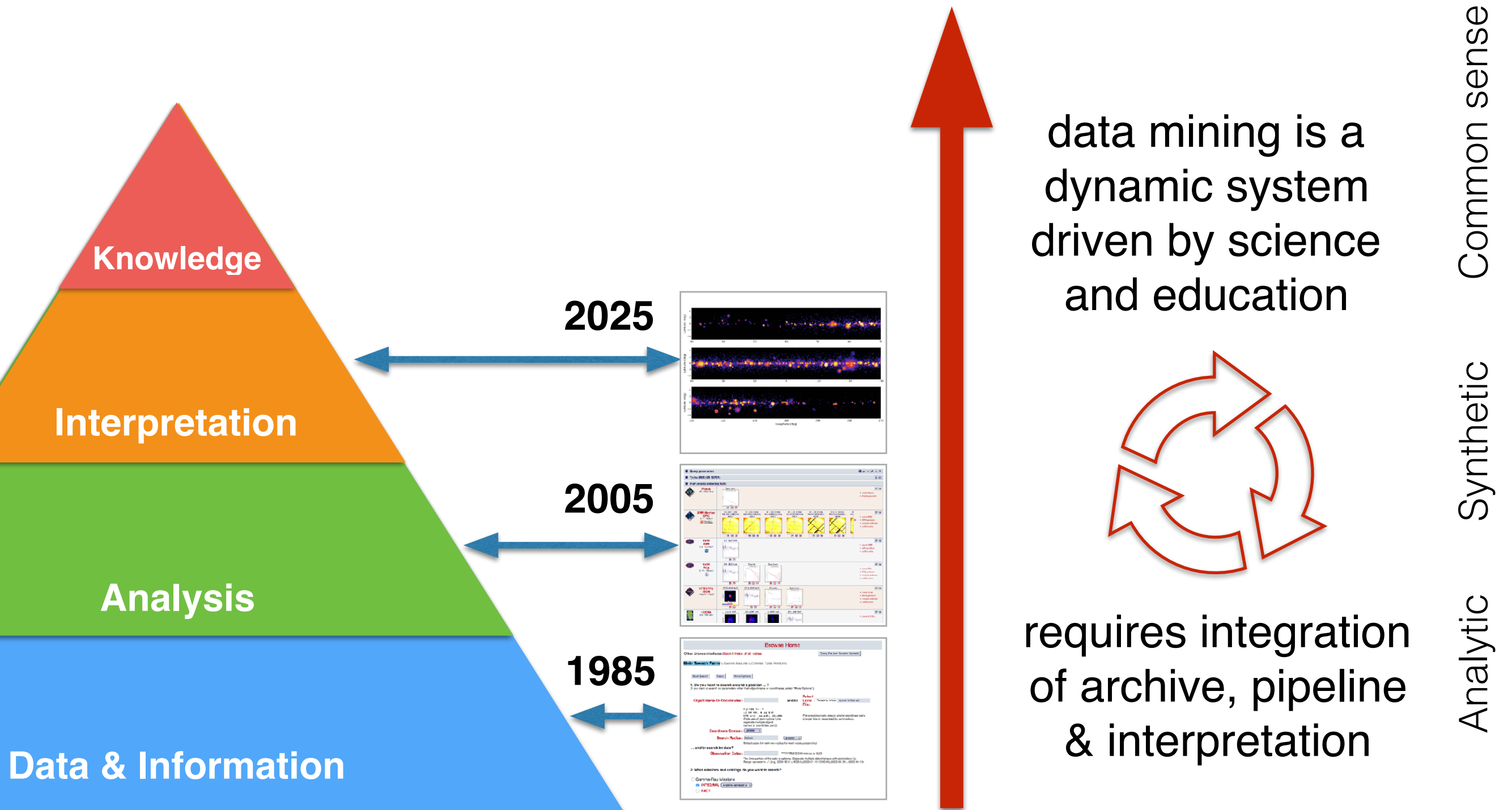
2005

1985

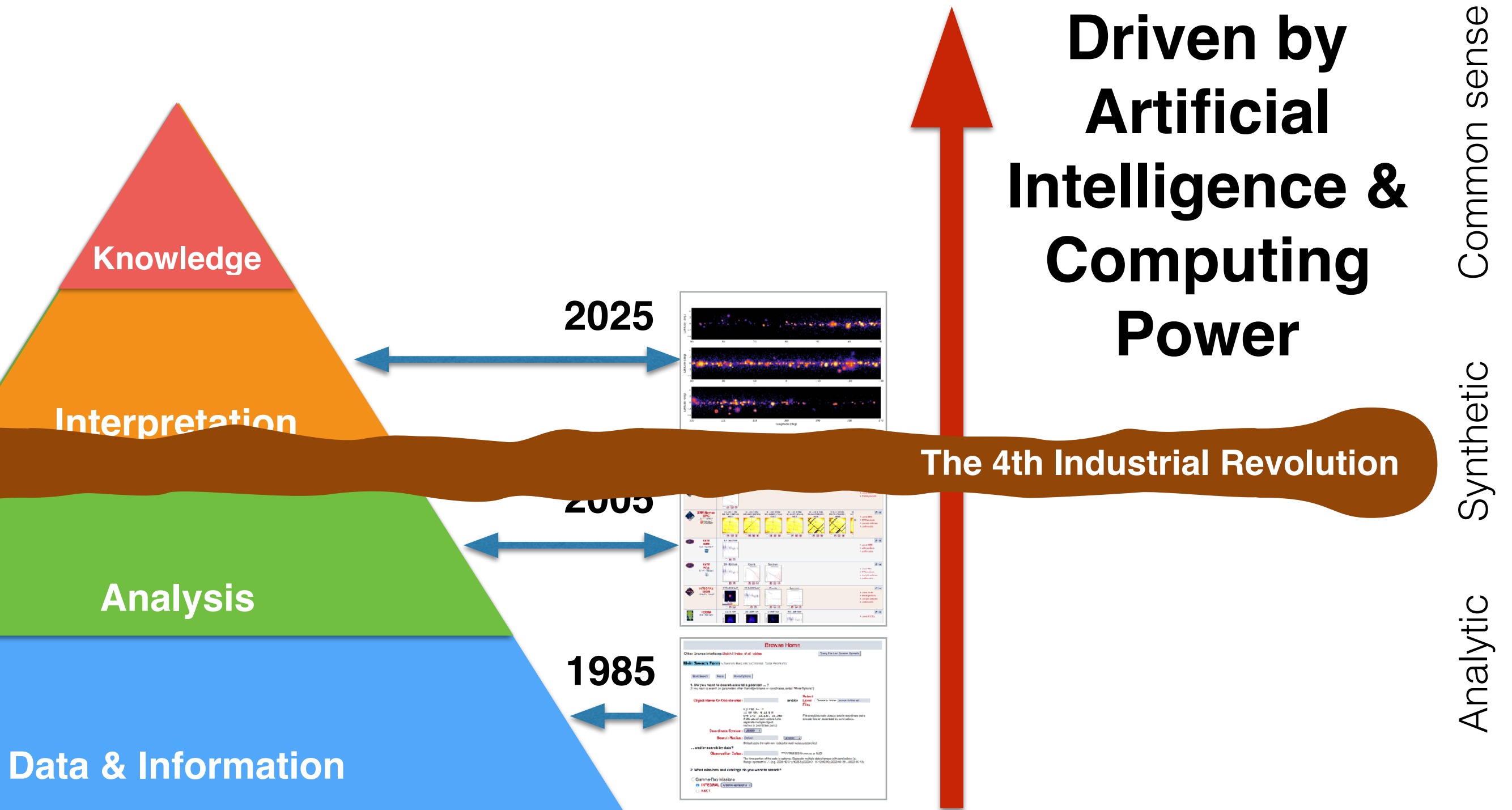


Common sense
Synthetic
Analytic

Opening the Process of Generating Science



Opening the Process of Generating Science



The 4th Industrial Revolution

- AI speaks to us via our smartphones
 - AI searches and finds for us
 - AI recognises us
 - AI monitors our health and helps decide on medical treatments
 - AI assists lawyers and the military
 - AI will soon drive our cars and work at our place
- **AI transforms science and education and comes at the rescue interpreting data flows exceeding human insight**

Klaus Schwab, World Economic Forum executive chairman:

One ambition of the VO

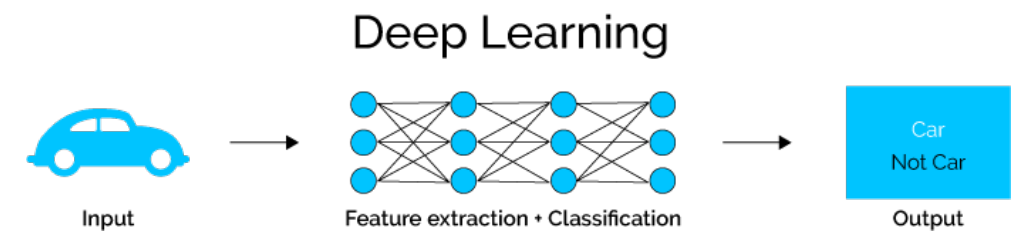
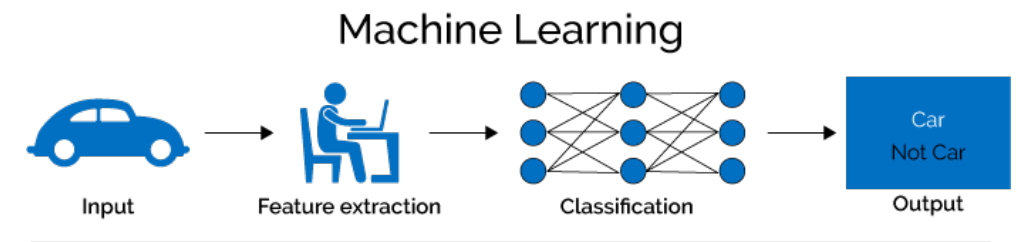
Previous industrial revolutions liberated humankind from animal power, made mass production possible and brought digital capabilities to billions of people. This Fourth Industrial Revolution is, however, fundamentally different. It is characterised by a range of new technologies that are fusing the physical, digital and biological worlds, impacting all disciplines, economies and industries, and even challenging ideas about what it means to be human. The business models of each and every industry will be transformed.

Should this not be the ambition of the OpenUniverse ??

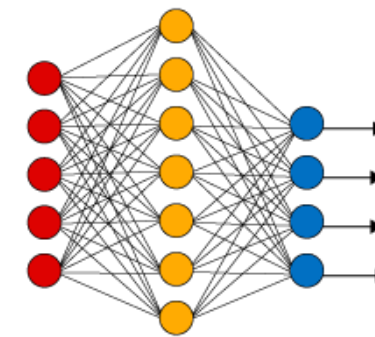
The 4th Industrial Revolution

Deep Learning does not need human knowledge to discover...

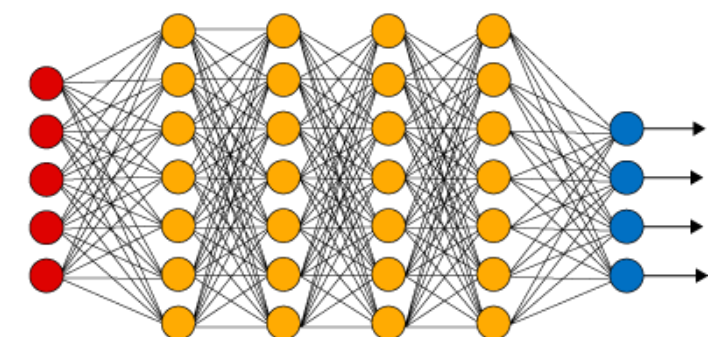
- Features extraction based on training data alone, **no user interaction**
- **Computationally expensive**
- Best architecture not intuitive



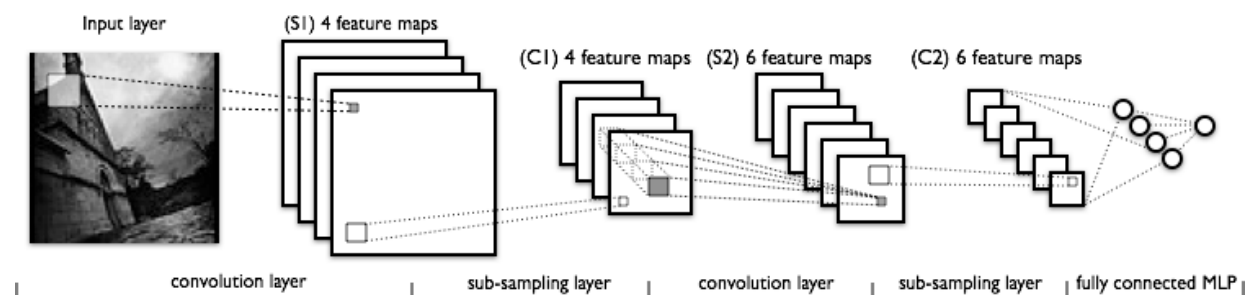
Simple Neural Network



Deep Learning Neural Network



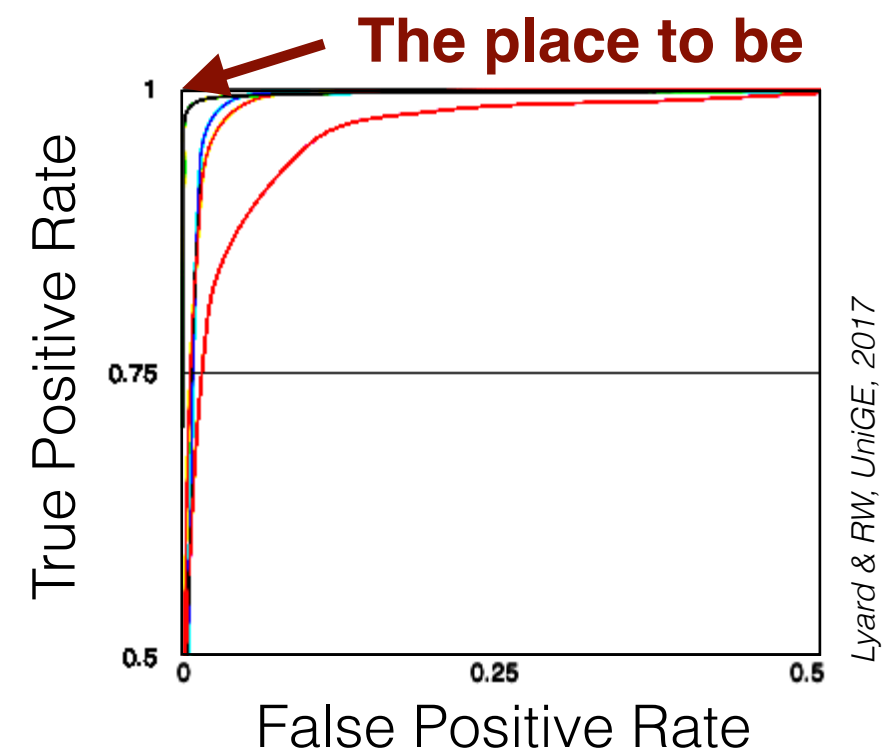
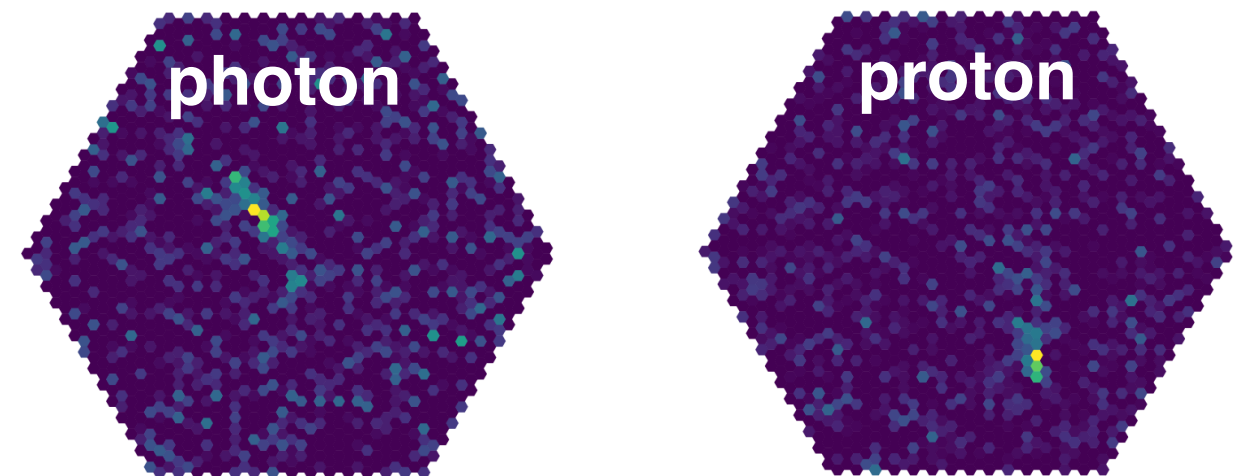
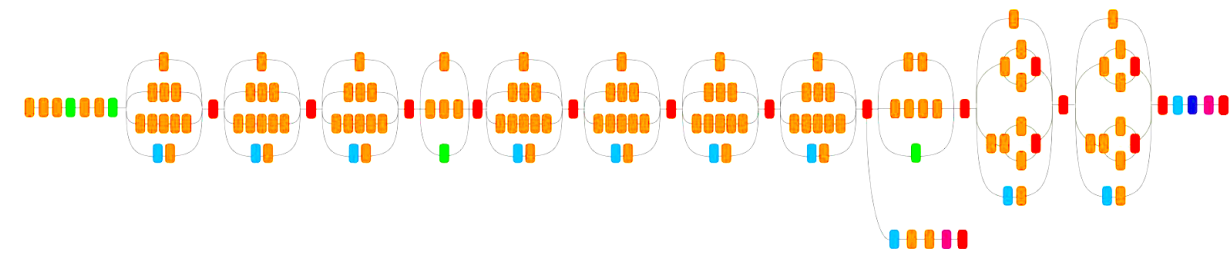
● Input Layer ● Hidden Layer ● Output Layer



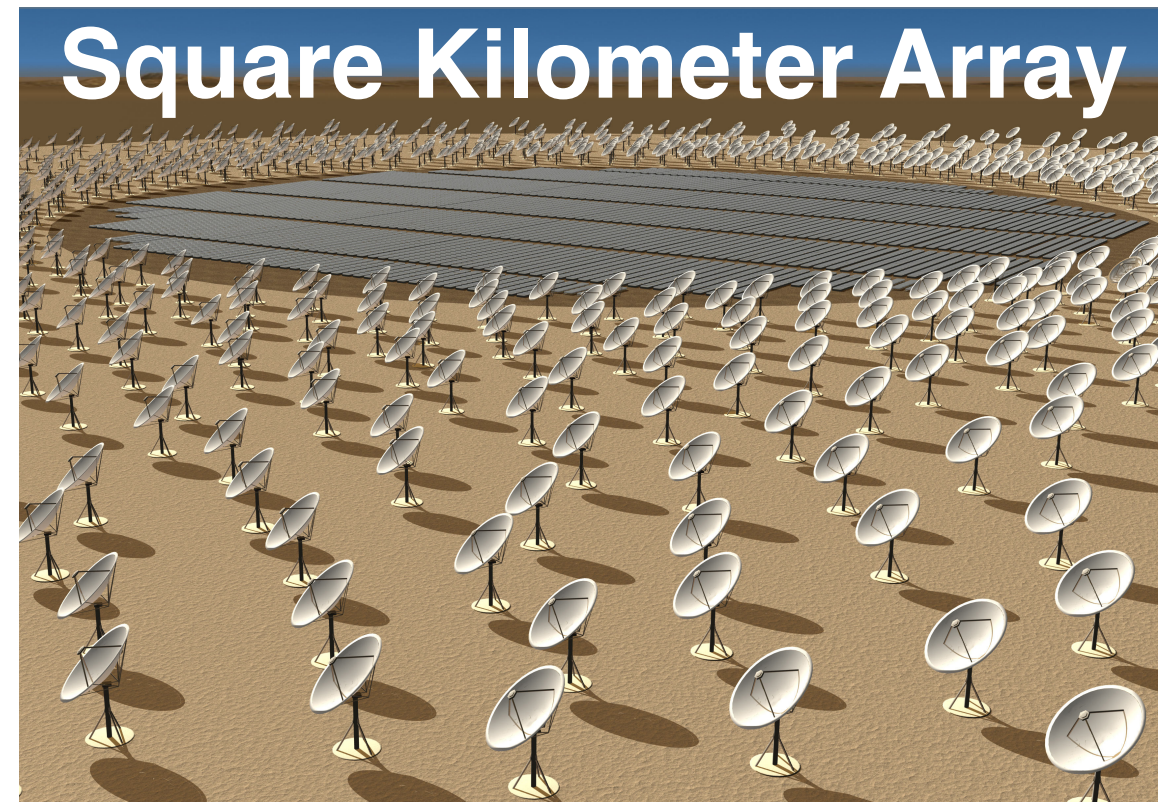
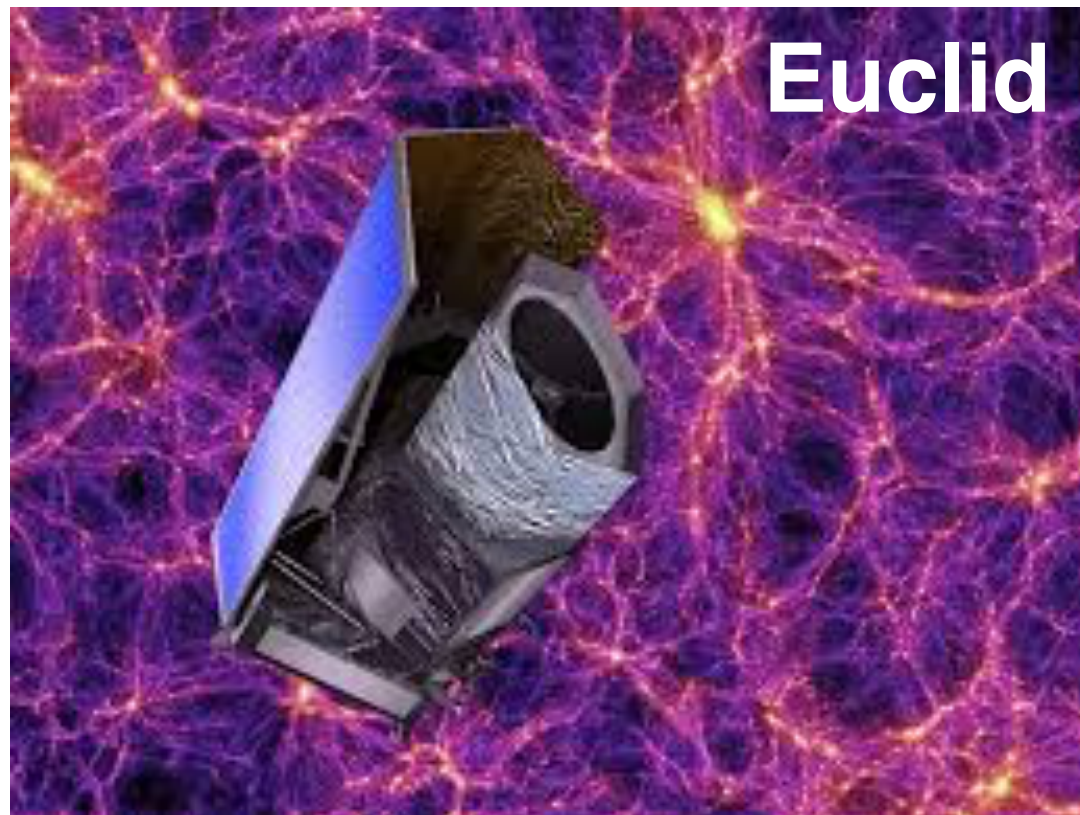
Classification via Deep Learning



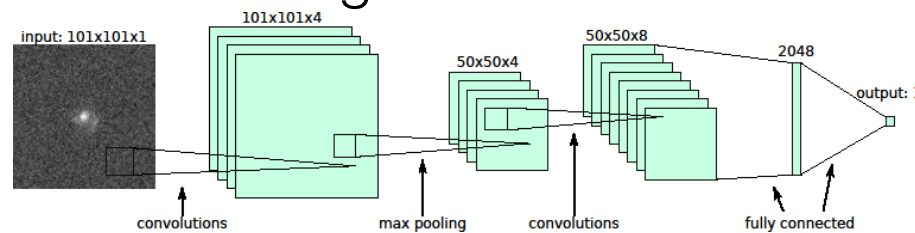
- Use state-of-the-art **InceptionV3** from Google
- Apply to Cherenkov event for classification



Characterising Galaxies via Deep Learning

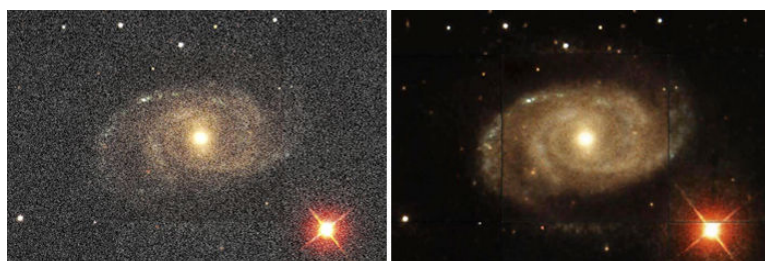


- Detection of gravitational lenses



Kneib et al, EPFL, 2017

- Image denoising



Schawinski et al, ETHZ, 2017

Class	Training Samples		Precision (%)
	Actual	Augmented	
Bent-tailed	177	25488	95
FR I	125	36000	91
FR II	227	32688	75
Average			88

Aniyan et al, SKA Cape Town, 2017

Feeding Old Data to AI ?

- **Instrument Idiosyncrasie detection and deconvolution**
- **Global search**
- **Searching features in the noise**
- **Denoising data**
- **Finding mistakes !**
- (**we are only at the very beginning of AI...**)

Legacy, a 30 years tribute



This mission brought not only new capabilities that resulted in unexpected discoveries, but also a pioneering approach to operations and archiving that changed X-ray astronomy...

Nicholas White

The world is moving in a new direction, should the OpenUniverse targets 30 years old ideas ?

Where is the future ?

- Increase transparency
- Resurface data
- Broaden the user base
- “Digital divide”

Improving the outcomes of the 3rd industrial revolution

**We have mostly discussed
30 years old ideas**

And / Or

Accompanying the 4th industrial revolution

- Universal access to modern investigations and computing power
- Free internet & computing...
- Resolving the “Artificial Intelligence divide”
- Make data accessible to modern analysis techniques

Any interest ?

A Great challenge for the UN

