

# Data Science for Physics and Astronomy

*Adrian Bevan (QMUL)*

*Benjamin Joachimi (UCL)*

*Darren Price (Manchester)*

*Jessie Wand (ATL)*



# Why are we here?

- 2016 Scoping workshop ,Big Data in the physical sciences‘  
<https://indico.cern.ch/event/449964/overview>

Includes link to resulting white paper

- Data science for science research programme:  
<https://www.turing.ac.uk/research/research-programmes/data-science-science>
- UKRI Strategic Priorities Fund @ ATI:  
<https://www.turing.ac.uk/research/asg>

Includes AI for Science (for National Labs only?)

## **Workshop Goal: collaboration**

- inter-disciplinary networking
- learn about challenges across fields
- identify common problems and knowledge exchange opportunities
- establish scoping projects

## **Workshop Goal: training**

- UK-level training for PhD students?
- training opportunities for senior researchers?
- knowledge transfer between data science and P/A?

## **Workshop Goal: planning for the future**

- How do we organise P/A data science activities in future?
- How will the P/A community link with the ATI?
- prepare for funding opportunities

# The plan

## *Today:*

- Now - Lightning talks
- Around lunch – find working groups
- Afternoon – group work: discuss data science

## *Tomorrow:*

- Morning – updates from groups
- Afterwards – group work: towards workshop goals
- Afternoon – group reports and next steps

# Meeting points of contact

Jessie Wand (ATI)  
[jwand@turing.ac.uk](mailto:jwand@turing.ac.uk)

Benjamin Joachimi (UCL)  
[b.joachimi@ucl.ac.uk](mailto:b.joachimi@ucl.ac.uk)

# Possible discussion seed topics

- Deep learning algorithms, very large datasets
- Dependence on ML algorithm choice and architectures
- Inhomogeneous and imbalanced data
- Time series, high data rates
- Explainable AI and data visualisation
- Inference and regression (“modelling“)
- Classification
- Unsupervised learning and anomaly detection
- Simulation, data augmentation



# Group discussions Monday afternoon

## *Constraints:*

- group size: 3-10
- inter-disciplinary

## *Task:*

- produce list of common data science challenges and opportunities for collaboration
- put notes (or edit directly) before end of day here:

[https://docs.google.com/document/d/1semrmTNFjI6TwEndKm\\_v49XiKRawVotmGkwelcJsvag/edit?usp=sharing](https://docs.google.com/document/d/1semrmTNFjI6TwEndKm_v49XiKRawVotmGkwelcJsvag/edit?usp=sharing)

# Splinter rooms

- Jack Good (opposite end of ATI from Enigma)
- David Blackwell (opposite end of ATI from Enigma)
- Ursula Franklin (off main common area)
- Lovelace and common area (tea/coffee/lunches)
- Enigma