



University
of Dundee



TREEHOOSE

Mid-Sprint Progress

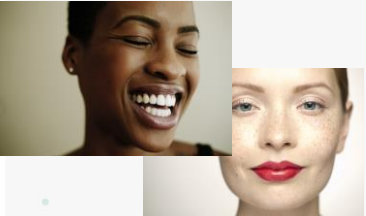
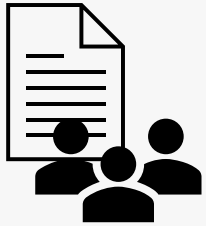
Dr Christian Cole



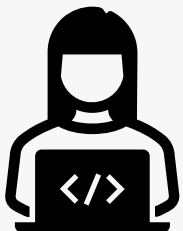
Algorithms



Sensitive Data



Approved User



Trusted Research Environment

powered by
aws

Elastic
CPU/GPU



Enclave



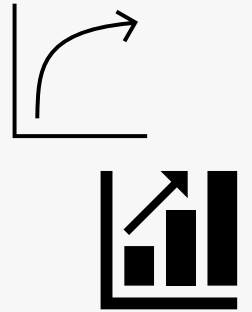
CSV



SQL



Outputs





HIC Cloud TRE

← → ↻ 🏠 🔒 https://d678a16ggft8p.cloudfront.net ☆ 📧 📱 ☰


Health Informatics Centre


HIC Cloud TRE
Trusted Research Environment on AWS (staging/eu-west-2
c16b4cb)


 Username


 Password


Login


 Trusted Research Environment


 Dashboard


 SSH Keys


 Studies

 Workspaces

 **Studies**

 **Find & Select Studies**
Select the desired studies

 **Select Compute**
Select a compute platform

 **Create Workspace**
Create the workspace

Submit Study for Approval

Select one or more studies to proceed to the next step.

Next →

Organization **1**

Open Data

Pending Approval

Rejected Approval


☐ **MSc Precision Medicine and Data Science**

Study: MSc_2021
Projects: researcher-teaching2021


Led by Alex Doney


▶ Files


▶ Permissions


 Trusted Research Environment


Chris Cole Logout


 Dashboard


 SSH Keys


 Studies

 Workspaces

 **Studies**

 **Find & Select Studies**
Select the desired studies

 **Select Compute**
Select a compute platform


 **Create Workspace**
Create the workspace


☐ **Ubuntu Linux-v1**

Ubuntu Linux environment using the MATE Desktop environment.

☒ **Windows-v1.1**

Standard Windows environment.

 Previous

Next 



Trusted Research Environment

Chris Cole

Logout



Dashboard



SSH Keys



Studies



Workspaces



Research Workspaces 3

Create Research Workspace

All

Available

Stopped

Pending

Errored

Terminated

AVAILABLE

MyMScWorkspace

Created 23 minutes ago by Chris Cole 41ae469e-2f5a-4748-b6de-94087f7810e2

Connections

View Detail

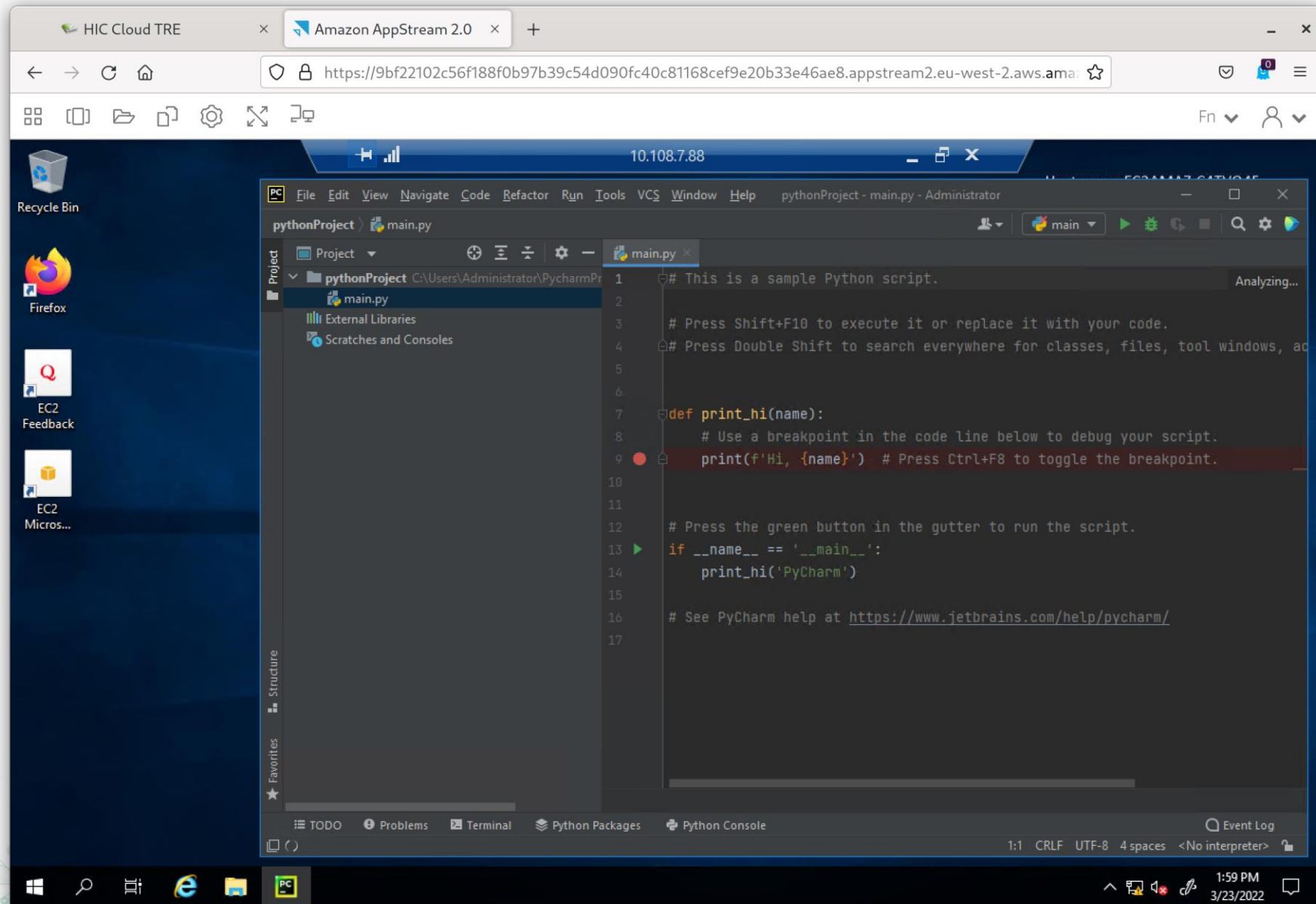
Stop

Terminate

2021-22 MSc course

Owner	Chris Cole
Studies	1
Project	researcher-teaching2021
Restricted CIDR	134.36.217.184/32
Workspace Type	Windows-v1.1

\$0.00
YESTERDAY'S COST



The screenshot displays a remote desktop session of an Ubuntu system. The desktop environment includes a terminal window on the left showing command-line interactions, a JupyterLab interface in the center, and a MATLAB R2021b window on the right. The MATLAB window is active, showing a 3D surface plot titled 'Figure 1'. The plot is a colorful, jagged surface representing a random distribution over a 10x10 grid. The MATLAB Command Window shows the command `>> surf(rand(10))` and the output `ans` with the value `'874227'`. The JupyterLab interface shows a file browser with a list of files and folders, including `anaconda3`, `Desktop`, `Documents`, `Downloads`, `Music`, `Pictures`, and `Public`. The terminal window shows a series of commands and outputs, including `[I 2022-03-16 16:48:04.281750] JupyterLab is running at https://localhost:8888/lab/tree/studies/u-JTHpml4LapCv...` and `[W 2022-03-16 16:48:04.281750] JupyterLab is running at https://localhost:8888/lab/tree/studies/u-JTHpml4LapCv...`.

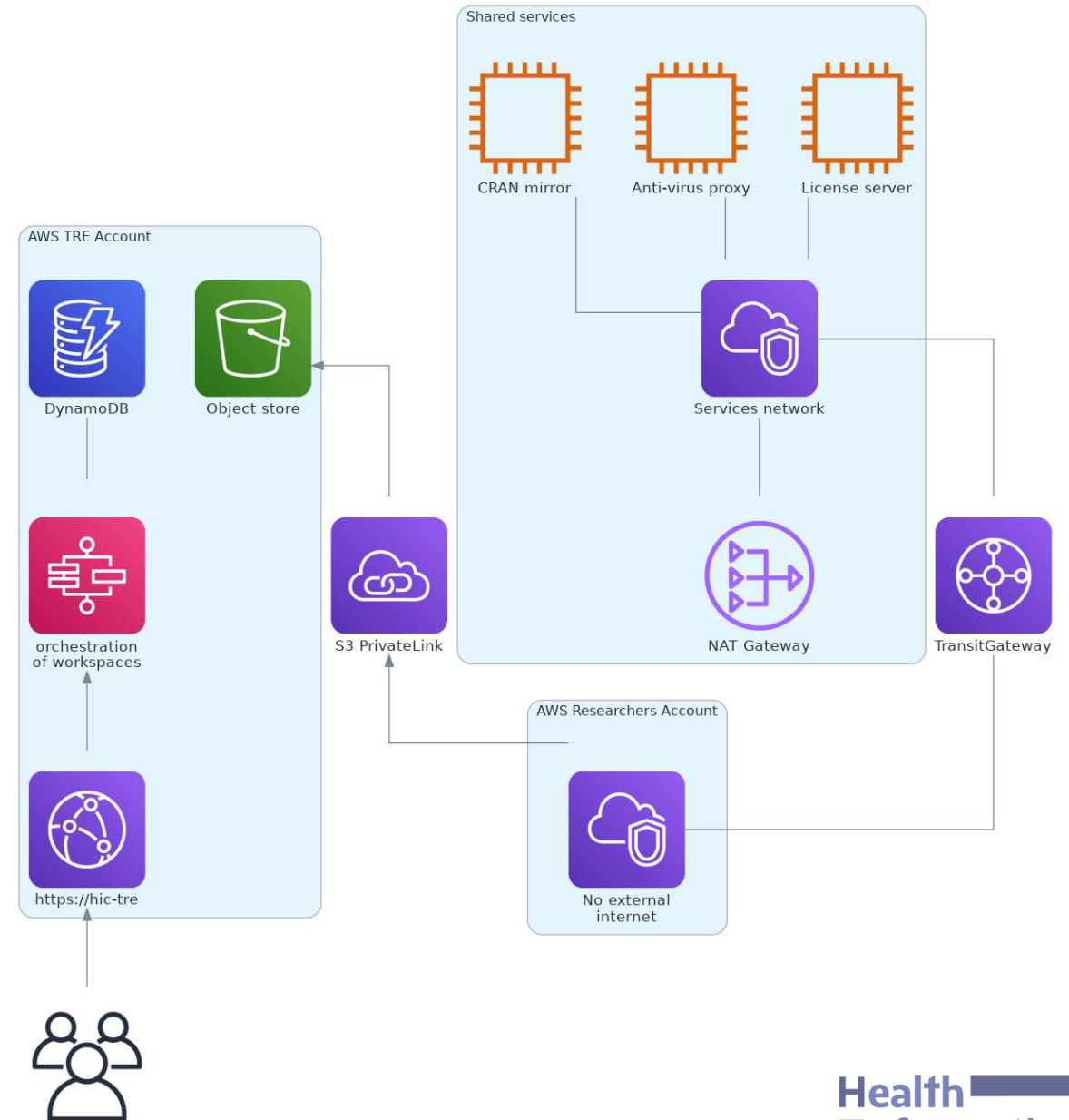
Infrastructure-as-code

- Storage is in one AWS account, separated by project
- Researcher workspaces in a separate AWS account
- Virtual display restricts copy/paste, web-browser is the only interface
 - Complies with Governance Requirements



Licencing

- Open-source: easy!
- Commercial: A problem when software requires a fixed host/MAC identifier, or external activation per instance
- MatLab: we run our own license server in the TRE (bridged to outside network)
- Microsoft Office: not yet optimised



Managing Costs

- Understanding shared overheads: small number of large projects vs large number of small projects
- Many projects have different compute requirements, sustained workloads tend to be short term so reserved instances won't work
- Non constant interactive workloads (e.g. "office hours") so savings plans don't work
- Cloud resilience vs cost: full redundancy is expensive



Challenges

- University Administration not “Sprinting”
 - Procurement with JISC, Finance, etc
- Project Planning with Industrial Partner
 - Aligning mutual needs and timelines
 - Setting up of agile sprints takes time
- University PPIE Resource “re-positioned”
 - Needed to find alternative





UK Research
and Innovation

HDRUK
Health Data Research UK



ADRUK
Data-driven change

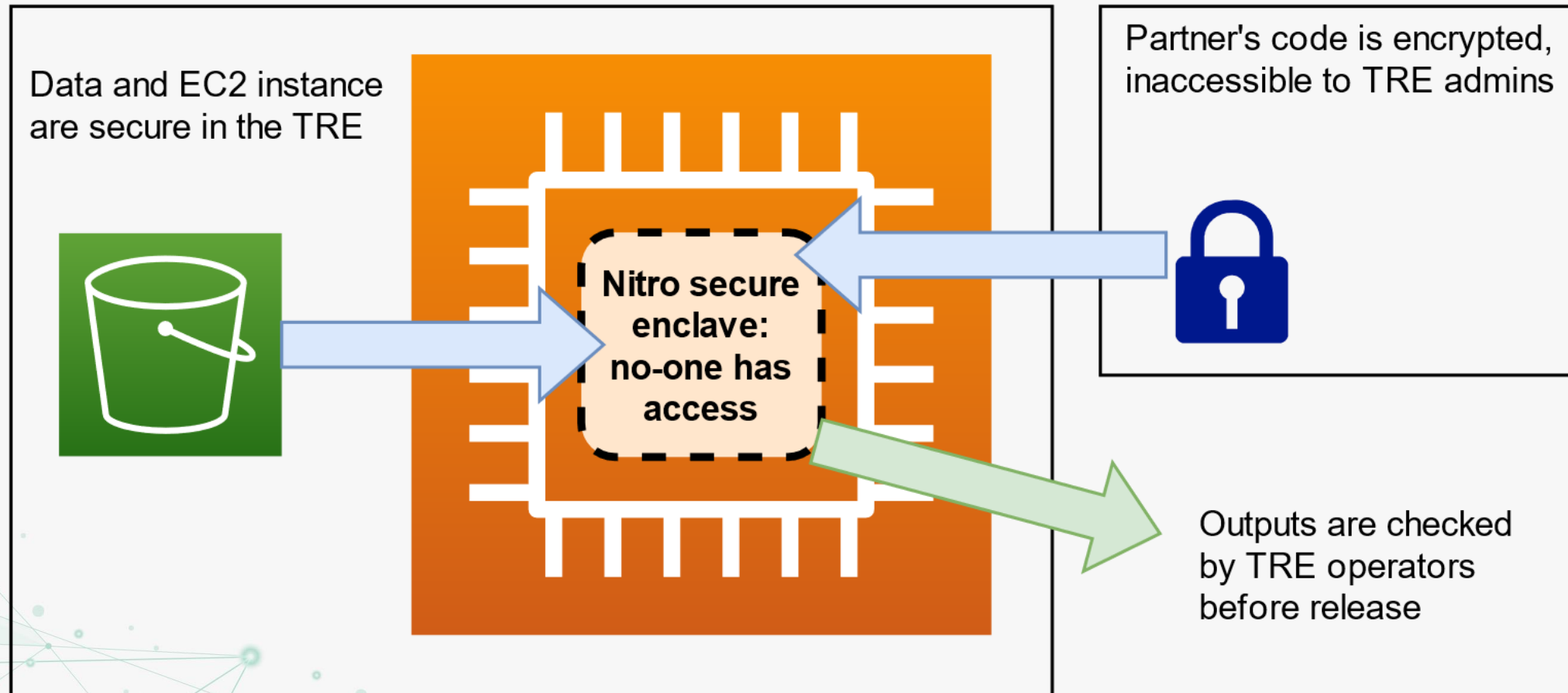
What's Next?



TREEHOUSE

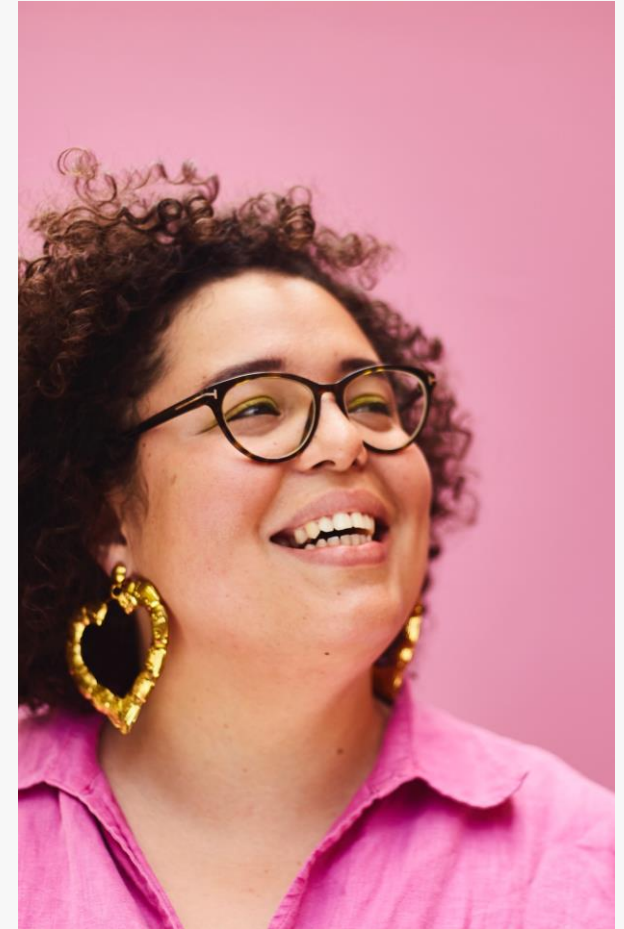
Secure Enclave

- Open source IXI dataset of brain MRIs
- Using open source deep learning classifier developed as part of PICTURES



PPIE

- Have experienced PPIE team in place
- Developing material now
- Workshops early May & early June
- DARE Sprint output late June/early July.





University
of Dundee



UK Research
and Innovation

HDRUK
Health Data Research UK



ADRUK
Data-driven change

Dundee

Dr Simon Li

Dr James Sutherland

Dr Chuang Gao

Prof. Emily Jefferson

AWS Team

Brian Nelson

Dr Phil Edwards

Maedeh Sharif Khodaei

Karan Shah

Syl Taylor

Rebecca Maeds

PPIE

Antony Chuter

Jill Beggs

PICTURES

Susan Krueger

Dr Esma Mansouri-Benssassi

Thank you for listening



TREEHOUSE

Health
Informatics
Centre