The Rosalind Franklin Institute

Research Software Engineering

Dr Laura Shemilt
10/06/25
University Sao Paulo

Me, My Work and The Rosalind Franklin Institute

The Rosalind Franklin Institute

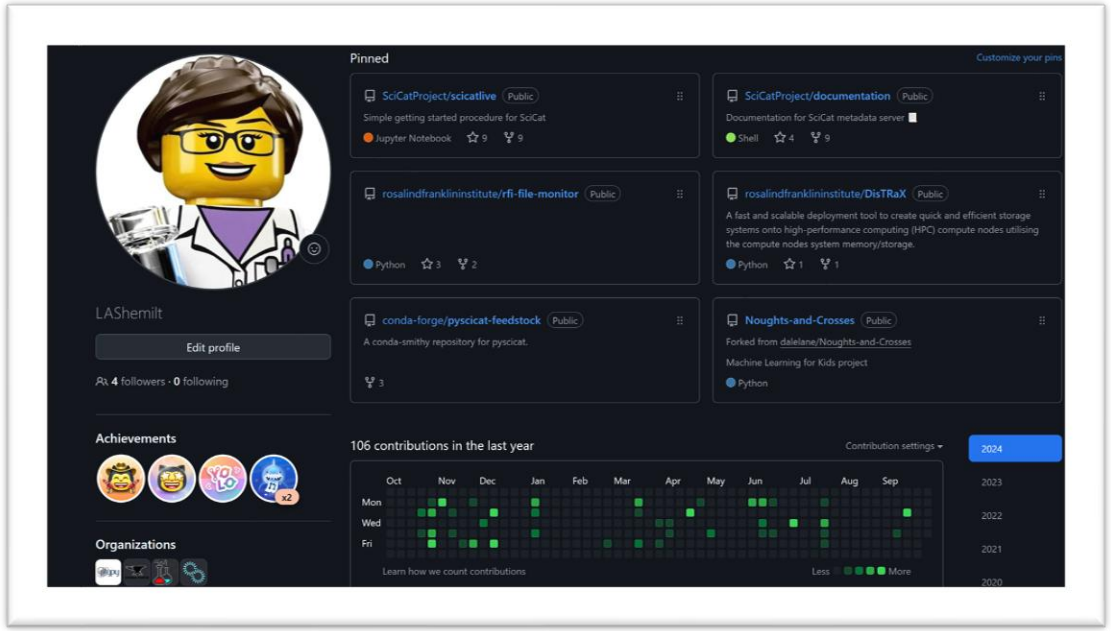
About Me

SciCat
advance your data to advance your science

SOFTWARE SUSTAINABILITY INSTITUTE

The Rosalind Franklin Institute

University of Southampton



kubernetes

ANSIBLE

openstack.

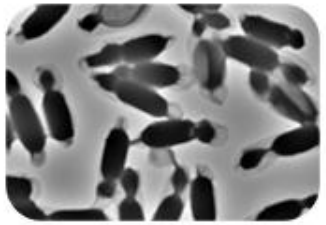
The Rosalind Franklin Institute

technology innovation
transforming life science
improving human health

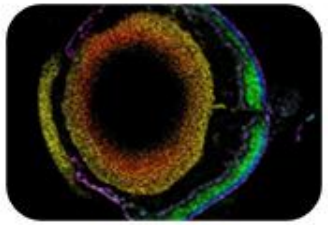
The Rosalind Franklin Institute

About the Rosalind Franklin Institute

Technology Innovation Challenges

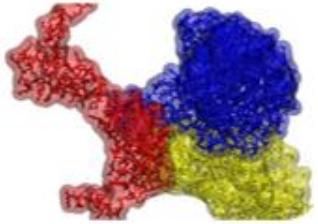


Multidimensional Molecular Imaging



Integrated Chemical Imaging in Cells and Tissues

Molecular Perturbations: Chemistry Engineering Biology

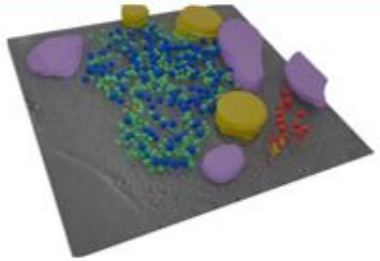


AI and Informatics for Predictive Biology

Application *pull*

Technology *push*

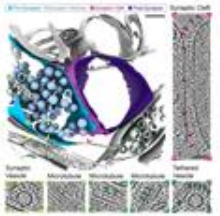
Life Science Challenges



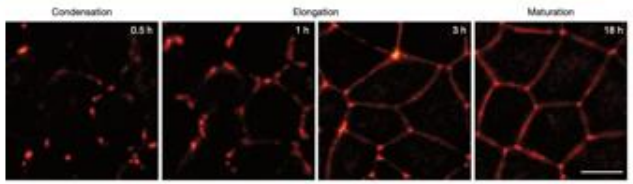
Biology across scales

Host-pathogen interactions

Emerging interest areas



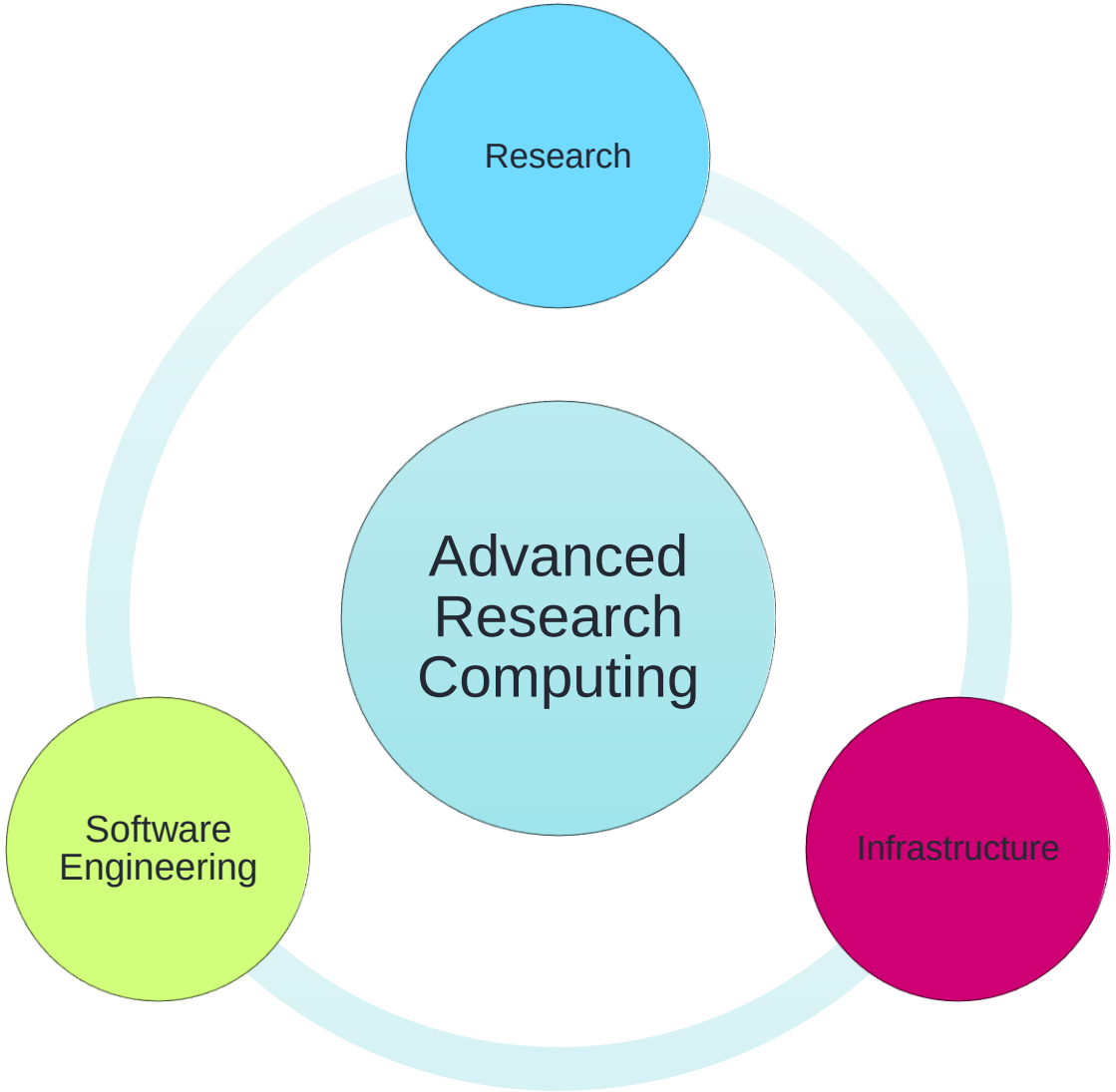
Cell compartmentation

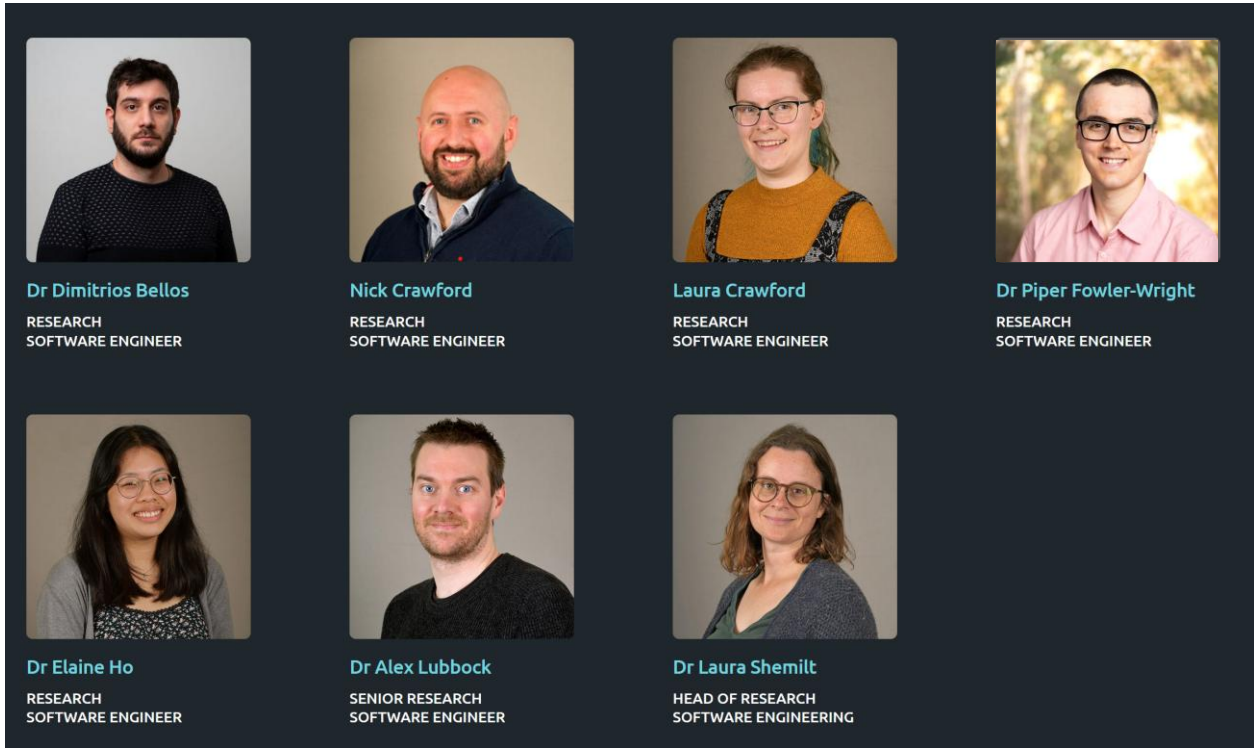


Cell-cell interaction

The Rosalind Franklin Institute

Advanced Research Computing





The Rosalind Franklin Institute

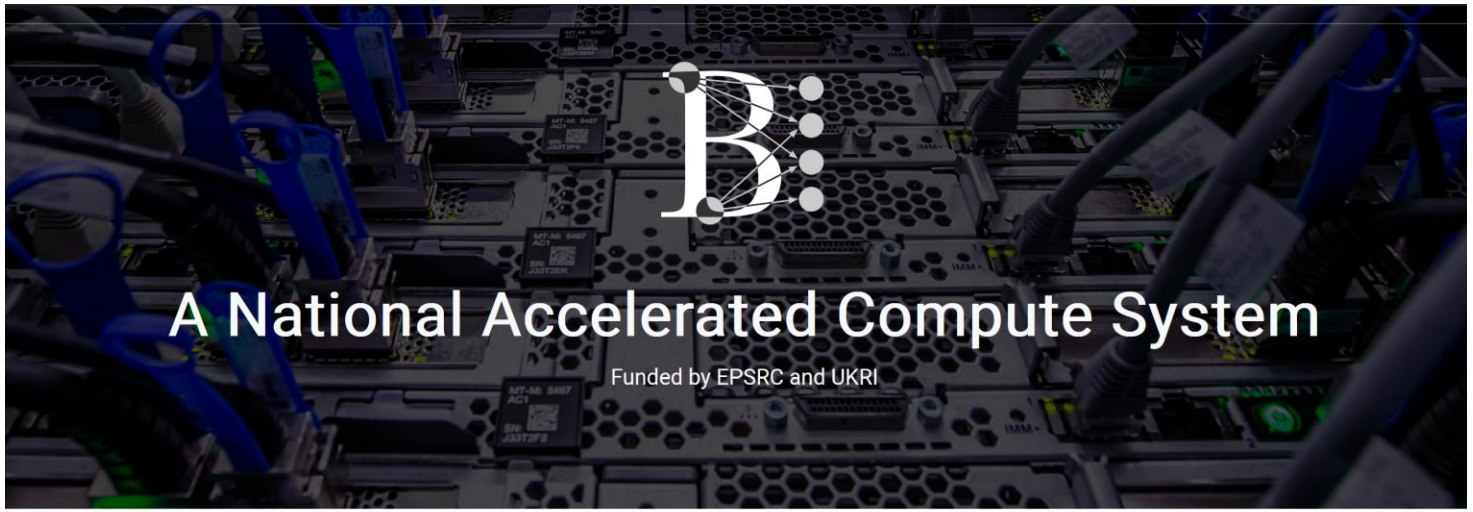
Our Partners: STFC – Cloud Computing

Scientific Computing Division

- Private IaaS: GPU and CPU Virtual Machines
- Storage provided by CEPH

Science and Technology Facilities Council

Baskerville– High Performance Computing



228
GPUs

57
Compute Nodes

5400 TB
Storage

UNIVERSITY OF BIRMINGHAM

diamond

The Rosalind Franklin Institute

UKRI Engineering and Physical Sciences Research Council

The Alan Turing Institute

AIRR – AI Research Resource

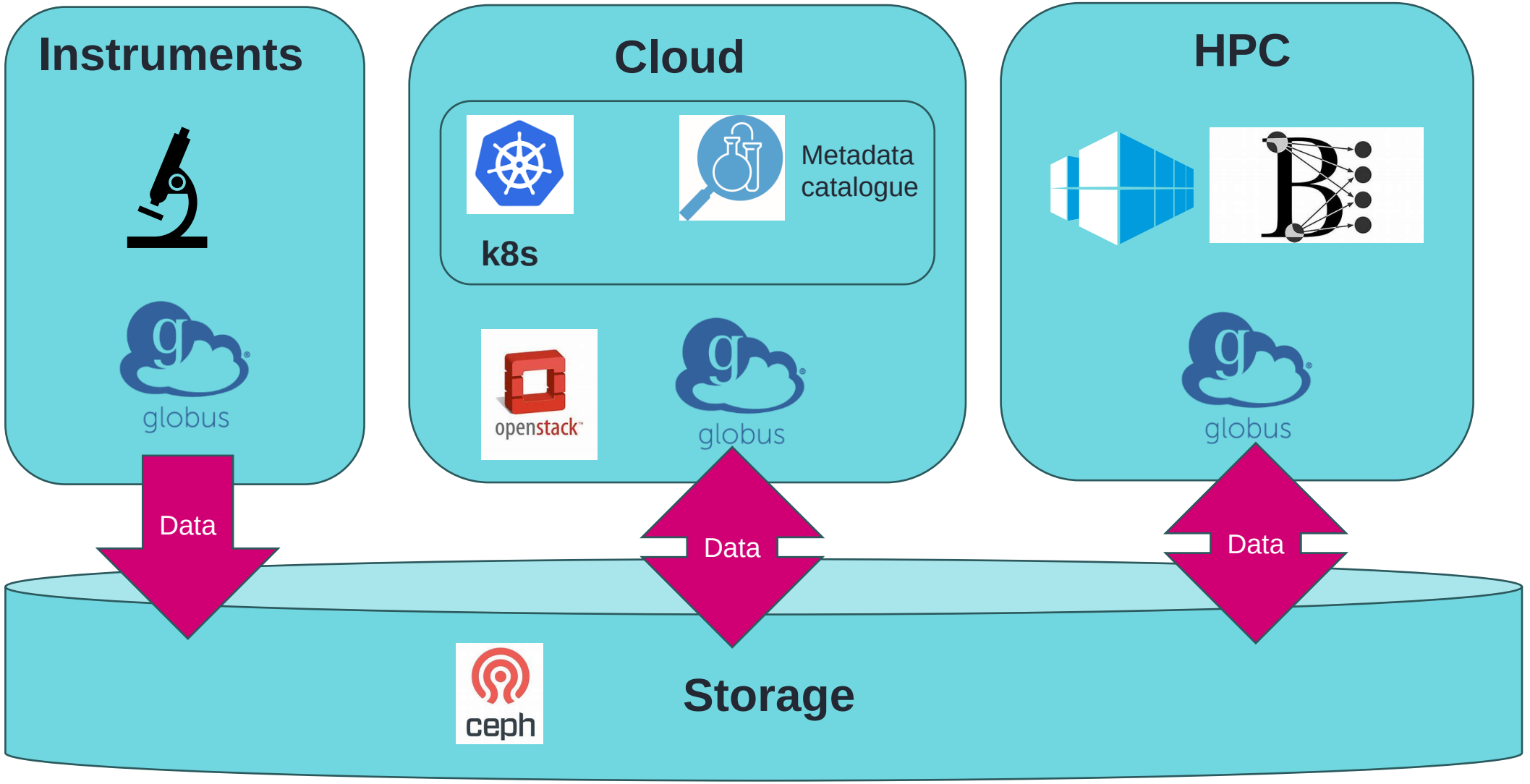
5280 NVIDIA GH200 Grace Hopper Superchips
Storage 29 PB

> 1000 Intel® Data Center GPU Max Series

https://www.bristol.ac.uk/news/2025/january/isambard-ai-govt-action-plan.html

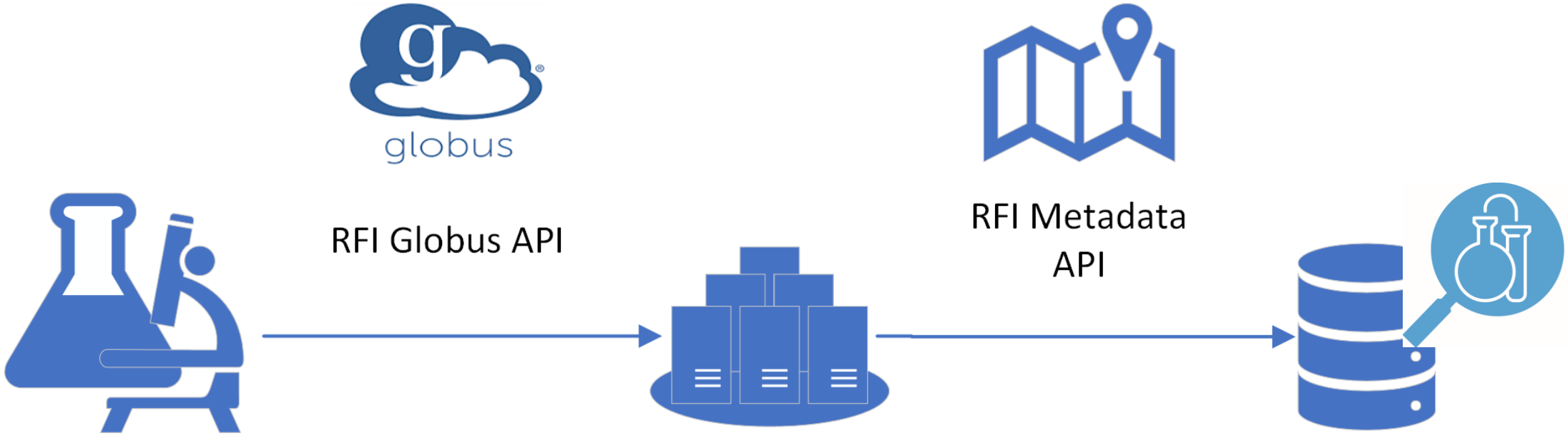
https://www.hpc.cam.ac.uk/d-w-n

Our platform



The Rosalind Franklin Institute

Data Capture and Management



Working with us

CAPES Post Doctoral Sandwich Exchange (PDSE)

Alisson G. Chiquitto

PhD student at PPGAB

Advisors:

Dr Alexandre Rossi Paschoal

Program Lead Biological Data Science, The Rosalind Franklin Institute, Associate Professor UTFPR

Dr Mark Basham

Scientific Director of AI for Predictive Biology, The Rosalind Franklin Institute

Open Source Software

International metadata catalogue for Science

Website: https://scicatproject.org

Open Source Software

The **rosalindfranklininstitute** organisation on Github

Antigen App	Web application for storing Nanobody data
FlowCron	ease of use data transfer and slurm script execution
Ot2Rec	Automated pipeline for Cryo-ET reconstruction
Parakeet	TEM digital twin
RFI-Globus-API	easy programmatic interaction with Globus

Research Software Engineering

The Rosalind Franklin Institute

Software is a great starting point for **collaboration**…

… but software collaborations can be a complete **nightmare**.

Why is it so hard?

Domain Experts

- Limited/ no knowledge of the software engineering process.
- Not thinking to share “obvious information”.
- Want to retain lots of freedom
- It’s “invisible”, it should be just a click of a button.

Software Engineers

- Often have limited/no domain expertise
- Need to abstract everything as much as possible
- Need to translate human requirements into computer requirements
- Are not magicians

RSE community

Website: https://society-rse.org/

Conferences:

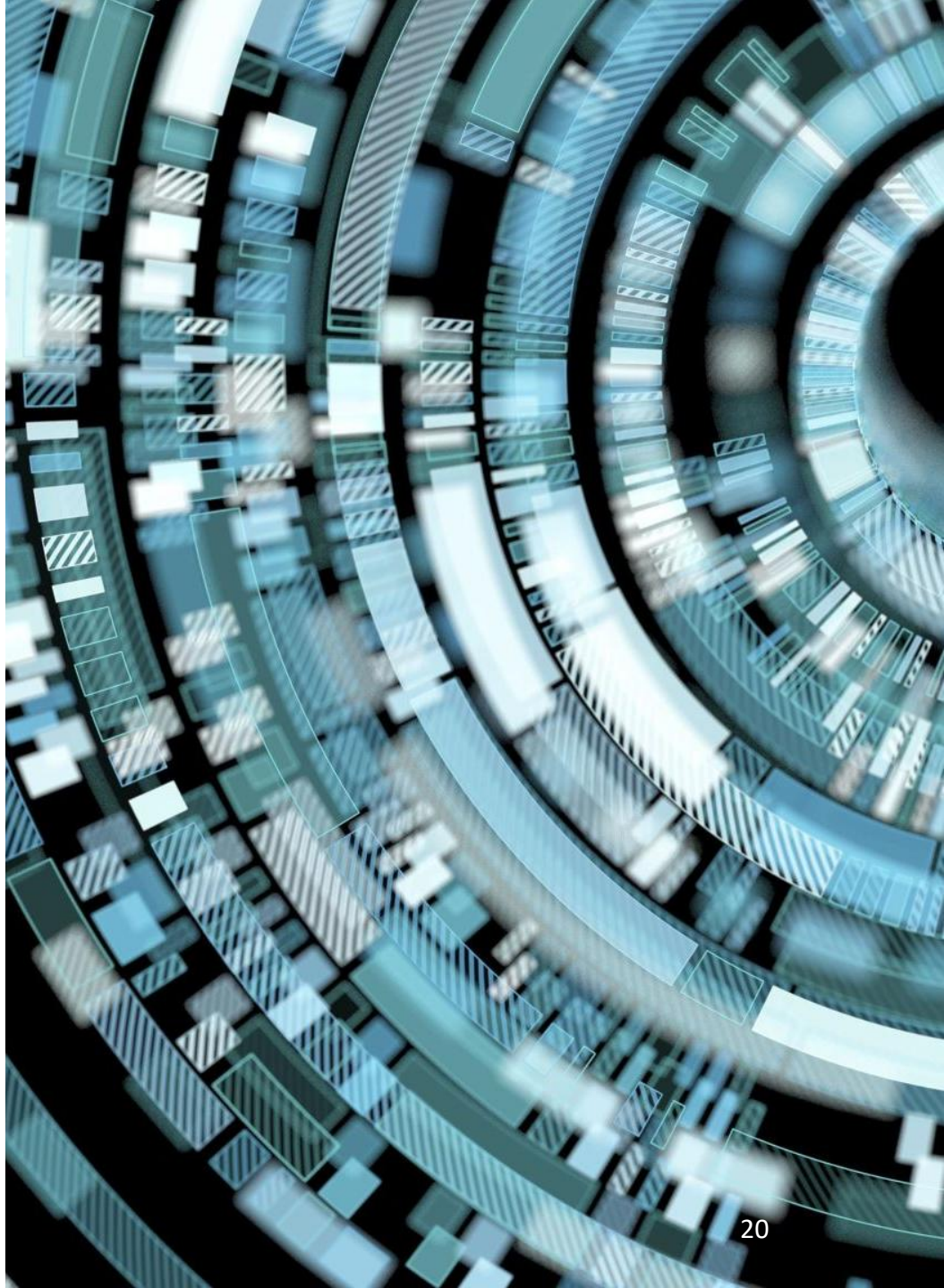
RSECon25 9th-11th September Warwick, UK

USRSE'25 6th-8th October, Philadelphia US

Enter the Research Software Engineer

"Research Software Engineers (RSEs) **combine knowledge of a research domain and the research process with software development skills and practices to create software for research**. They regularly apply expertise in programming to advance research."

Northwestern Research Computing Services



Why should you bother?

Speeds up process

Provides a standard

Moves towards better repeatability/reproducibility

Solve a research problem

What skills do you need?

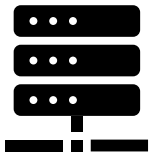
Digital Research Competencies Framework

https://direct-framework.github.io/digital-research-competencies-framework/

- A PhD or significant experience working in a research-focused setting (3-4 years)
- Demonstrate good software engineering practicesUse version control such as git
- Working with containers
- Comfortable using Linux
- Excellent organisational and communication skills
- Ability to co-ordinate multiple aspects of work to meet deadlines
- Ability to work as part of a team, as well as independently
- Ability to solve problems using innovative and flexible thinking
- Passionate about learning new technologies
- Version Control

Green RSE

The need for green computing



Estimated Carbon Emissions of Data Centres worldwide

~126 MT CO2e per year

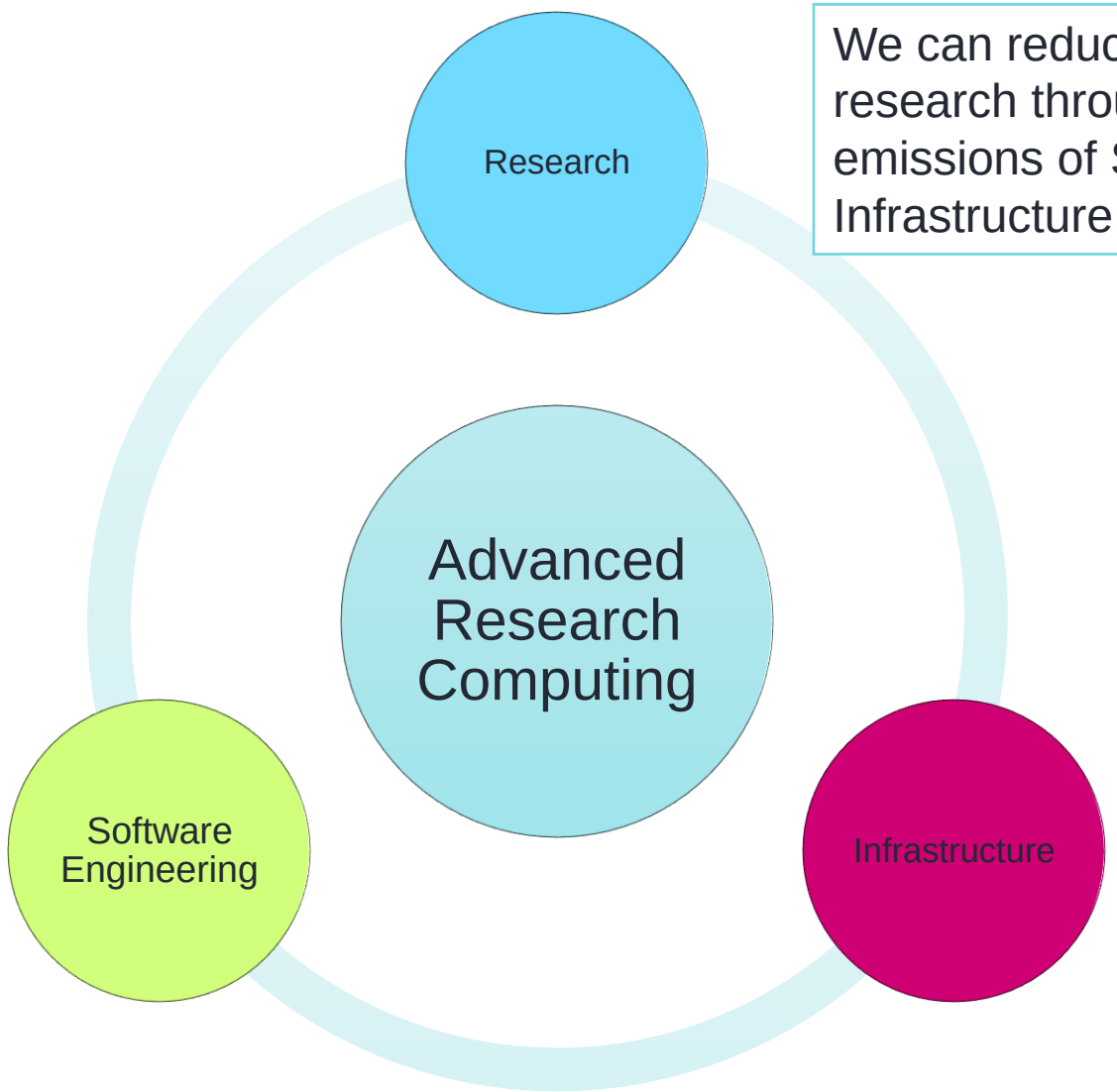
Estimated Carbon Emissions of LLMs

~500T CO2e per year

Estimated Carbon Emissions of the UK

~303 MT CO2e per year

Green Advanced Research Computing



We can reduce the carbon emissions of research through reducing the emissions of Software and Infrastructure

Sustainable Software = Green Software

Reductions in Carbon = Reduction in Cost

Green Software Principles

Energy Efficiency

Consume the least amount of electricity possible

Hardware Efficiency

Use the least amount of embodied carbon possible

Carbon Awareness

Do more when the electricity is clean and less when it's dirty

Green Software Foundation

greensoftware.org

The Rosalind Franklin Institute

GREENER principles

- **G**overnance
- **R**esponsibility
- **E**stimation
- **E**nergy and **E**mbodied Impacts
- **N**ew Collaborations
- **E**ducation
- **R**esearch

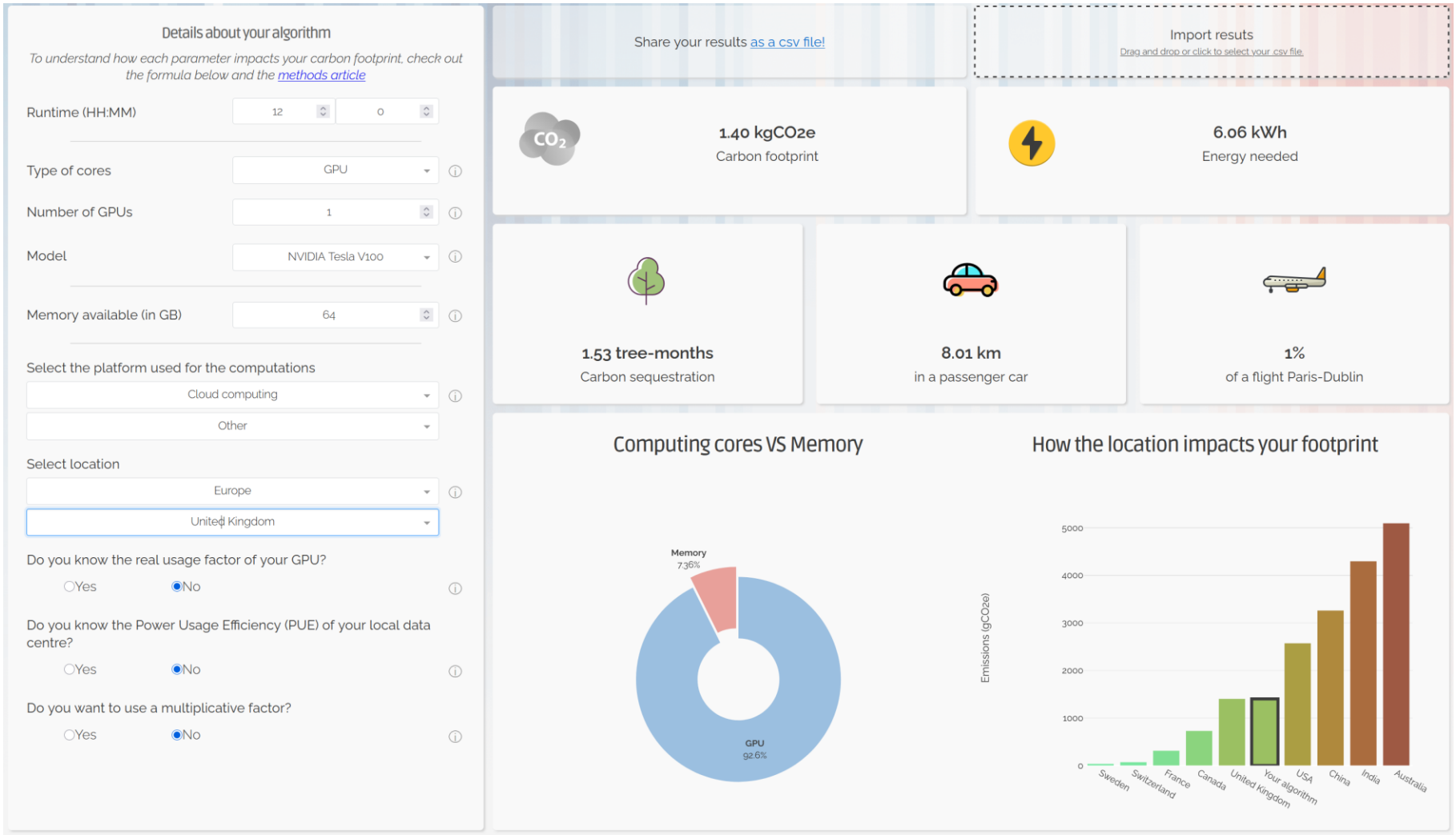
Lannelongue et al. Nature Comp. Sci.;3(6):514-521.
doi: 10.1038/s43588-023-00461-y

Green DiSC

- Accreditation of Green Software practices
- Roadmap for researchers on how to make their software greener
- Open source and can be used by anyone

https://www.software.ac.uk/GreenDiSC

Green Algorithms



www.rfi.ac.uk

https://calculator.green-algorithms.org/

The Rosalind Franklin Institute

Green RSE

Green RSE special Interest Group (SIG) aims to build a **supportive, collaborative community** of research software engineers and researchers who code, to work together on **understanding how to reduce the environmental impact of research software.**

- Remote events to reduce emissions from travel

https://socrse.github.io/green-sig/

Green Infrastructure

Top-Down Approach
Those who run infrastructure have the ability to make their platform greener.

- What can you do as a user?
- What can you do as a someone who runs infrastructure?

Successes
- Isambard-AI #4 in Green500 - https://top500.org/lists/green500/2024/11/
- NetDrive: UK Funding for Net Zero Infrastructure
- Training Course on Green Software on ARCHER https://www.archer2.ac.uk/training/courses/250513-green-computing/

Thanks to

The Rosalind Franklin Institute
Mark Basham
Michelle Darrow
Alexandre Paschoal
Alex Lubbock
Avery Pennigton
Dimitrios Bellos
Laura Crawford
Elaine Ho
Nick Crawford
Piper Fowler-Wright
Neville Yee
James Parkhurst
Fabiana Roderigues de Goes
Sam Kersley
Tibor Auer (Former)
Silvia Ramos (Former)
Joss Whittle (Former)

STFC
Martin Summers
Tom Byrne
Jacob Ward
John Good
Aidan Mc Coomb

SciCat
Max Novelli (ESS)
Dylan McReynolds (ALS)
Carlo Minotti (PSI)
Daphne Van Dijke (DECTRIS)

BriCS
Matthew Williams
Christopher Woods
Sadaf Allam
Wahab Kawafi
James Allam

Baskerville
Andrew Morris
Gavin Yearwood
James Allsop (Former)

Green RSE
Kirsty Pringle
Loic Lannelongue
Andy Turner