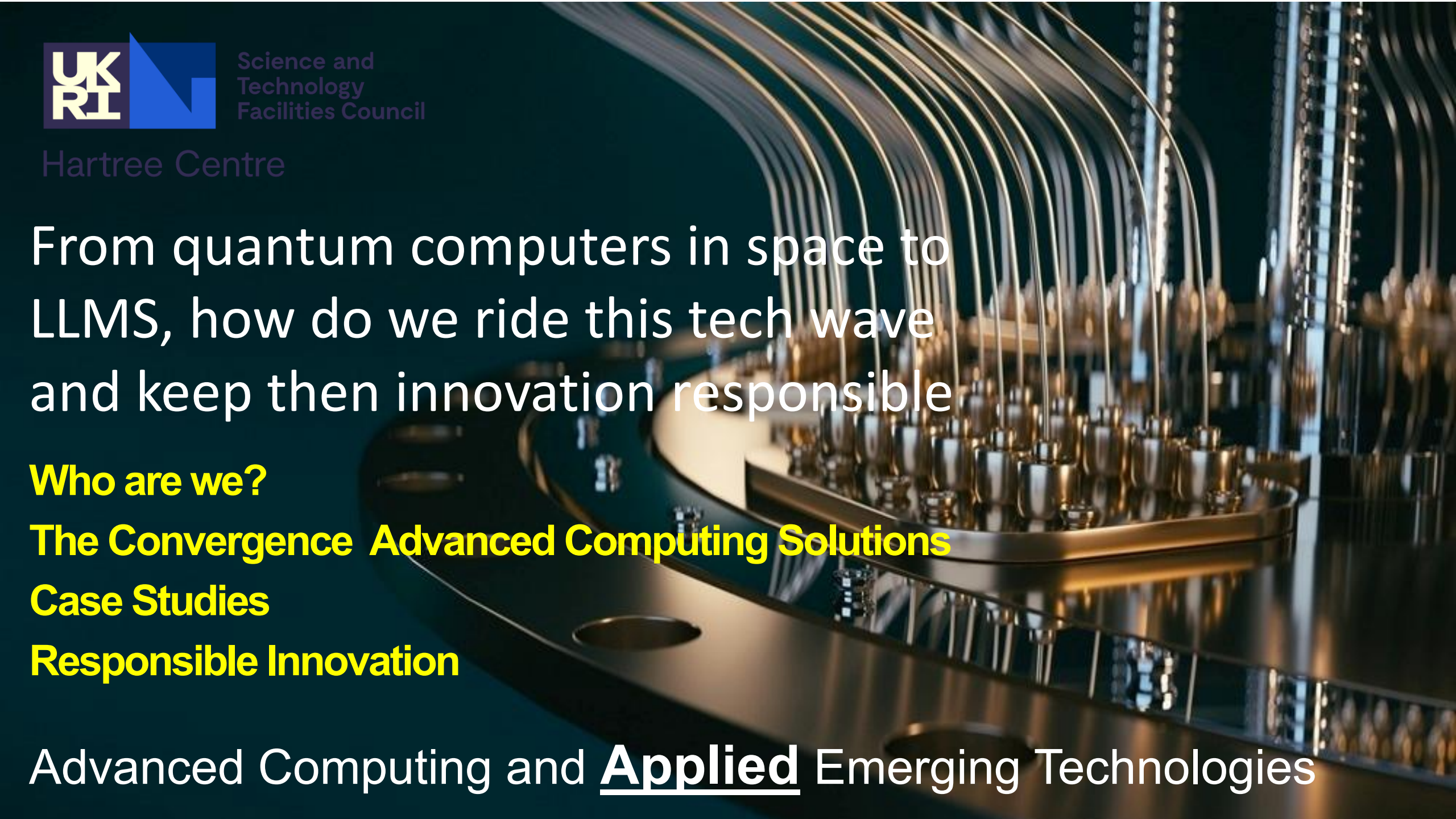


From quantum computers in space to LLMS, how do we ride this tech wave and keep then innovation responsible

Who are we?

The Convergence Advanced Computing Solutions

Case Studies

Responsible Innovation

Advanced Computing and **Applied** Emerging Technologies

UK Research & Innovation

Single R&I funding body:

- Seven research councils
- Research England
- Innovate UK

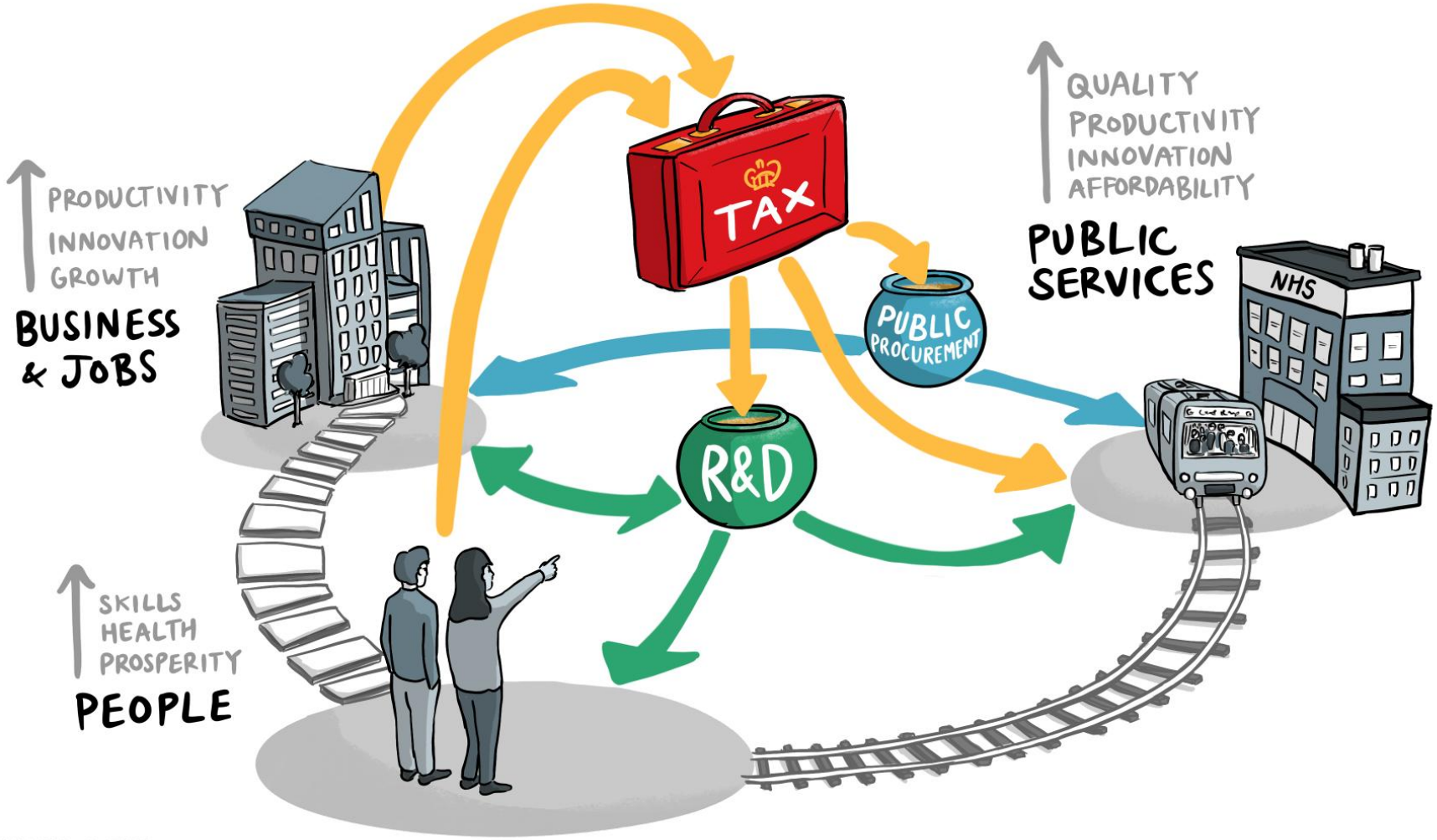
UK Research and Innovation

Science and Technology Facilities Council

Hartree Centre

UKRI and Hartree Mission

Accelerate the R&D Benefits



Science and Technology Facilities Council

Hartree Centre

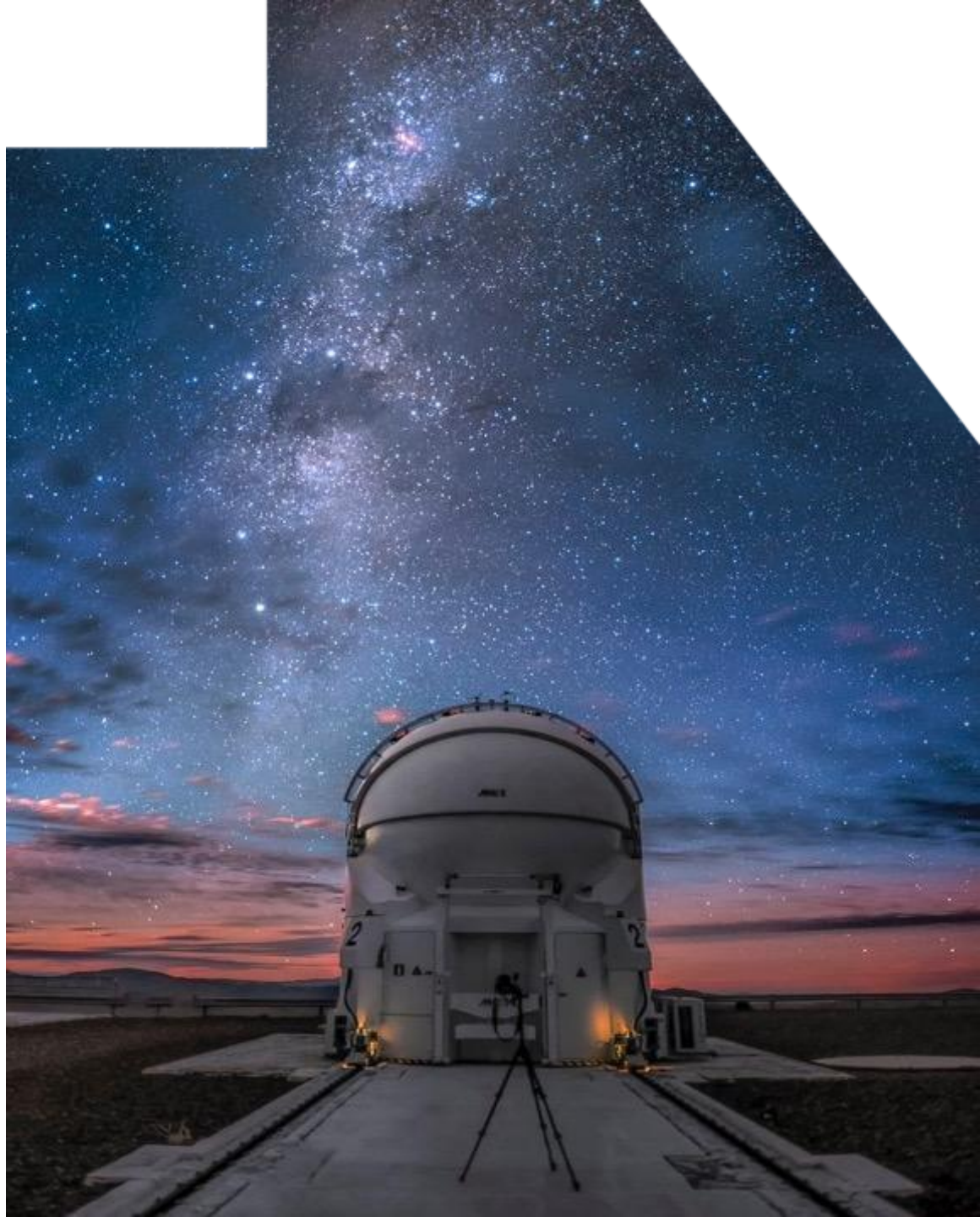
UKRI-STFC's mission

Mission:

- Discovering the secrets of the Universe
- Developing advanced technologies
- Solving real world challenges

Responsibilities:

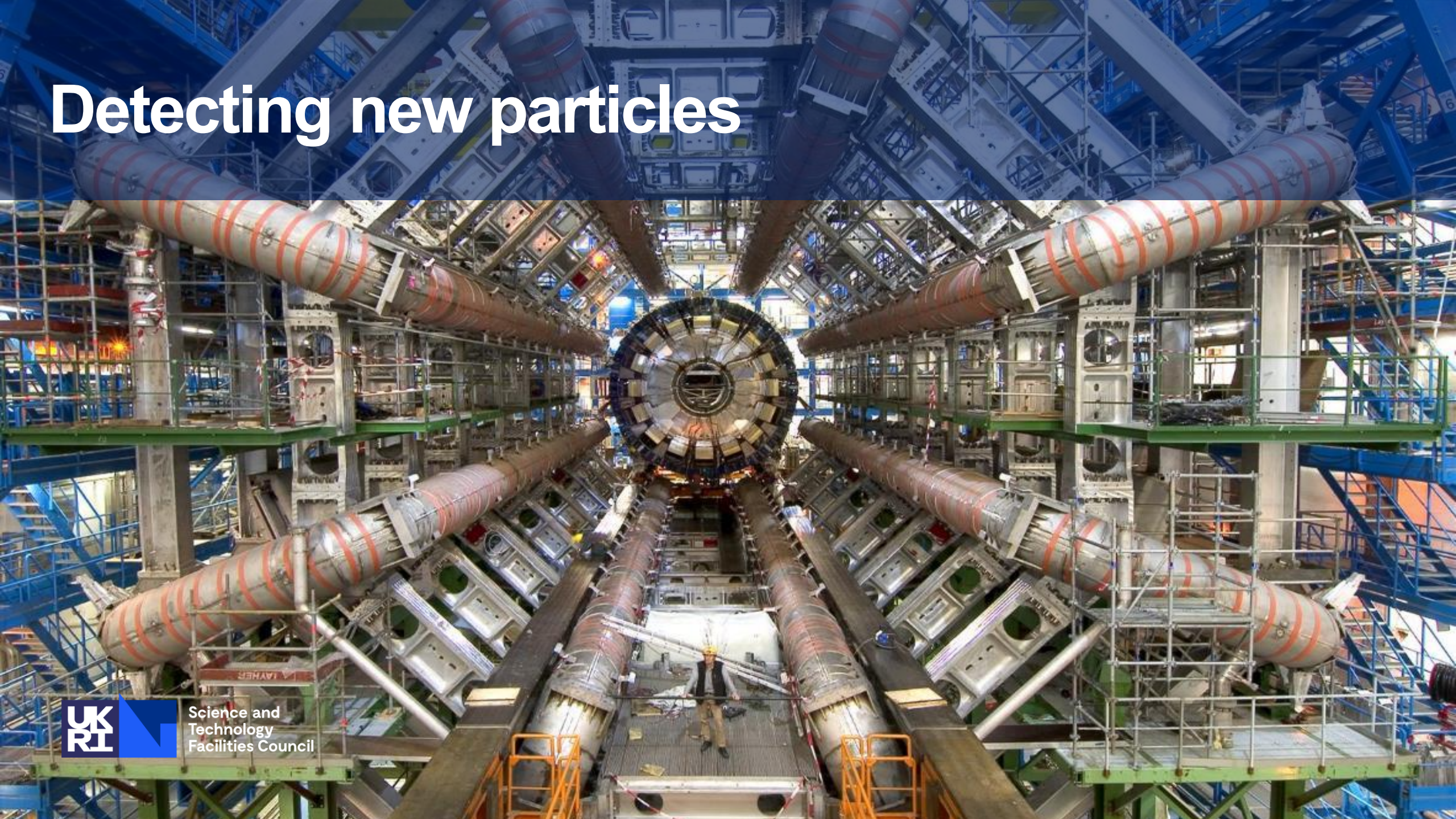
- Frontier research: particle physics, astronomy, nuclear physics and space science
- Major UK multi-disciplinary facilities
- Stewardship of the R&I campuses



UKRI Science and Technology Facilities Council

Hartree Centre

Detecting new particles

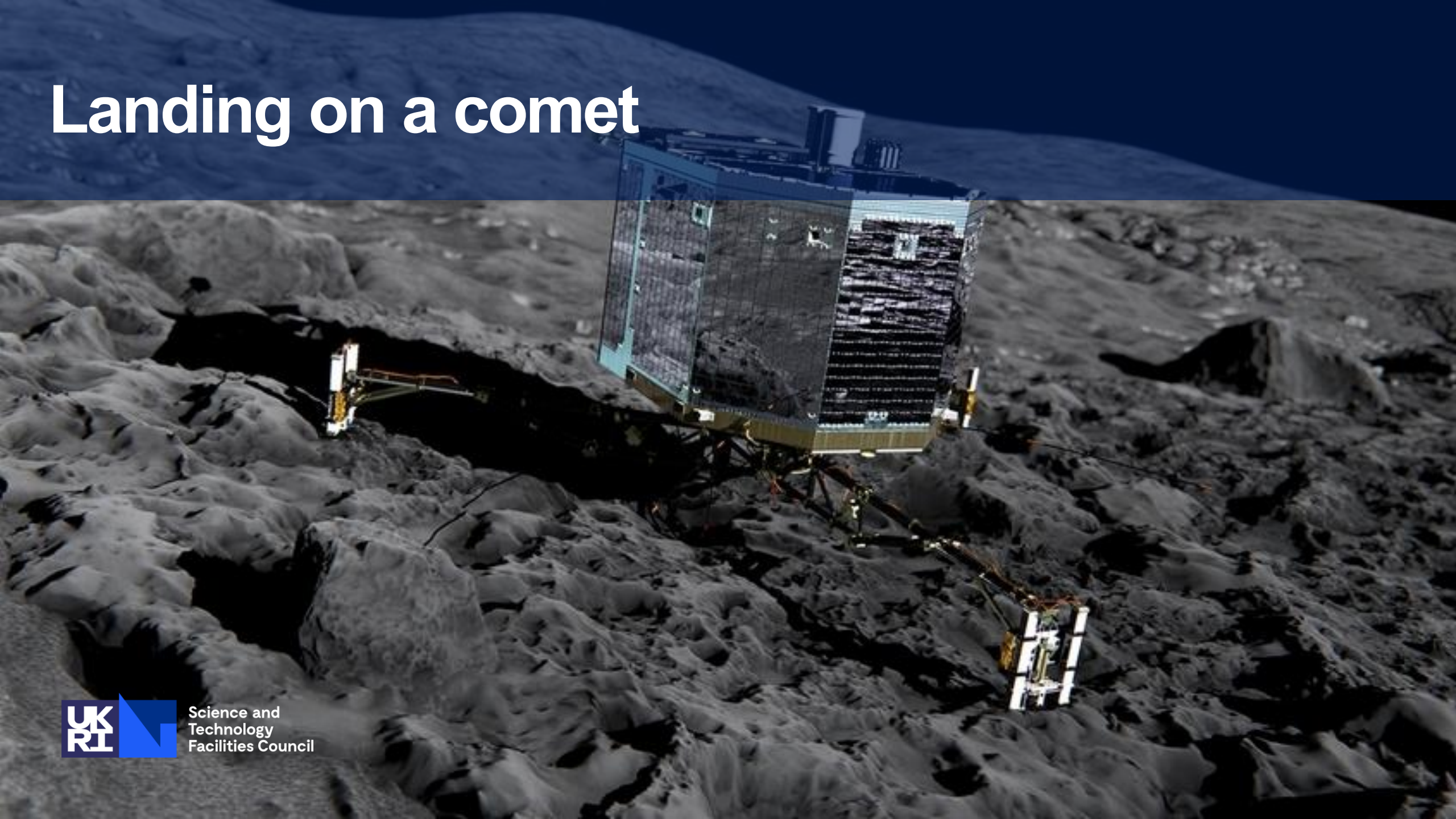


Science and Technology Facilities Council

Building the world's largest telescope

Science and Technology Facilities Council

Landing on a comet



Science and Technology Facilities Council

Convergence HPC, AI & QC

The next wave of technology disruption will not be from a single technology, but through **a converged ecosystem of composable technological innovations, validated through application on real-world, disruptive, challenges**

Our Platforms & Facilities

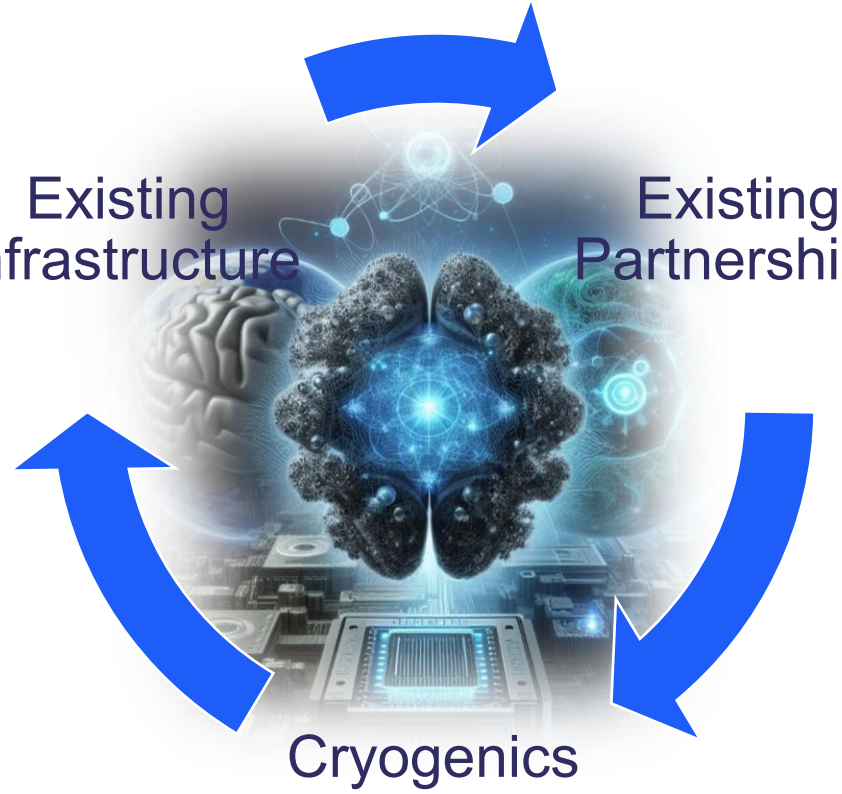
Mary Coombs

GPU theoretical peak performance:

- 39 pflop/s FP64 in Turbo mode.
- 31 pflop/s FP64 in nominal mode.

Leading to our entry at number 63 in the June 2025 Top500 list.

157 compute nodes, each with:
4 x NVIDIA H100 GPU
2 x Intel Xeon Gold 6448Y 32-core 2.1Ghz processors
512GB DDR5 memory
10 high-memory compute nodes, with 1.5TB DDR5 memory
10 login / datamover nodes, each with:
2 x Intel Xeon Gold 6448Y 32-core 2.1Ghz processors
512GB DDR5 memory
1 x NVIDIA H100 GPU
Mellanox ConnectX NDR200 Infiniband interconnect
16 visualisation servers, each with:
2 x Intel Xeon Gold 6448Y 32-core 2.1Ghz processors
512GB DDR5 memory
1 x NVIDIA L40S GPU
1.19PB usable scratch storage (IBM Spectrum Scale filesystem)
SLURM job scheduler



Working with:

IBM Q

D-Wave The Quantum Computing Company™

PsiQuantum

QUANTUM BRILLIANCE

QUANTUM DICE

QUANTINUUM

IONQ

UKRI Science and Technology Facilities Council
Hartree Centre

Atos Quantum Learning Machine

Discovering and delivering next-generation algorithms for future quantum computers

Science and Technology Facilities Council

Hartree Centre

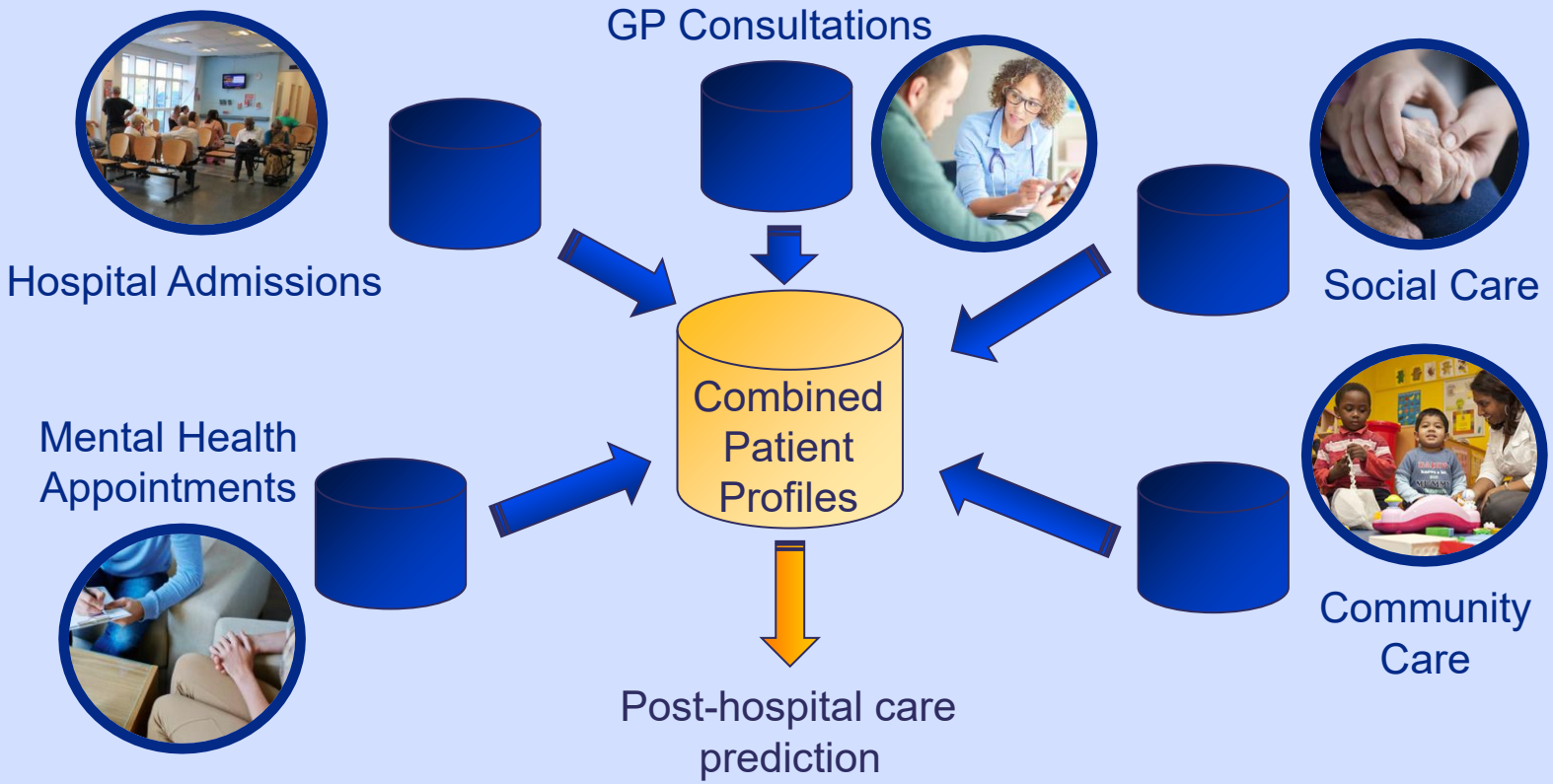
Our Track Record

A High-Level Overview

Science and Technology Facilities Council

Hartree Centre

Case study: Predictive modelling



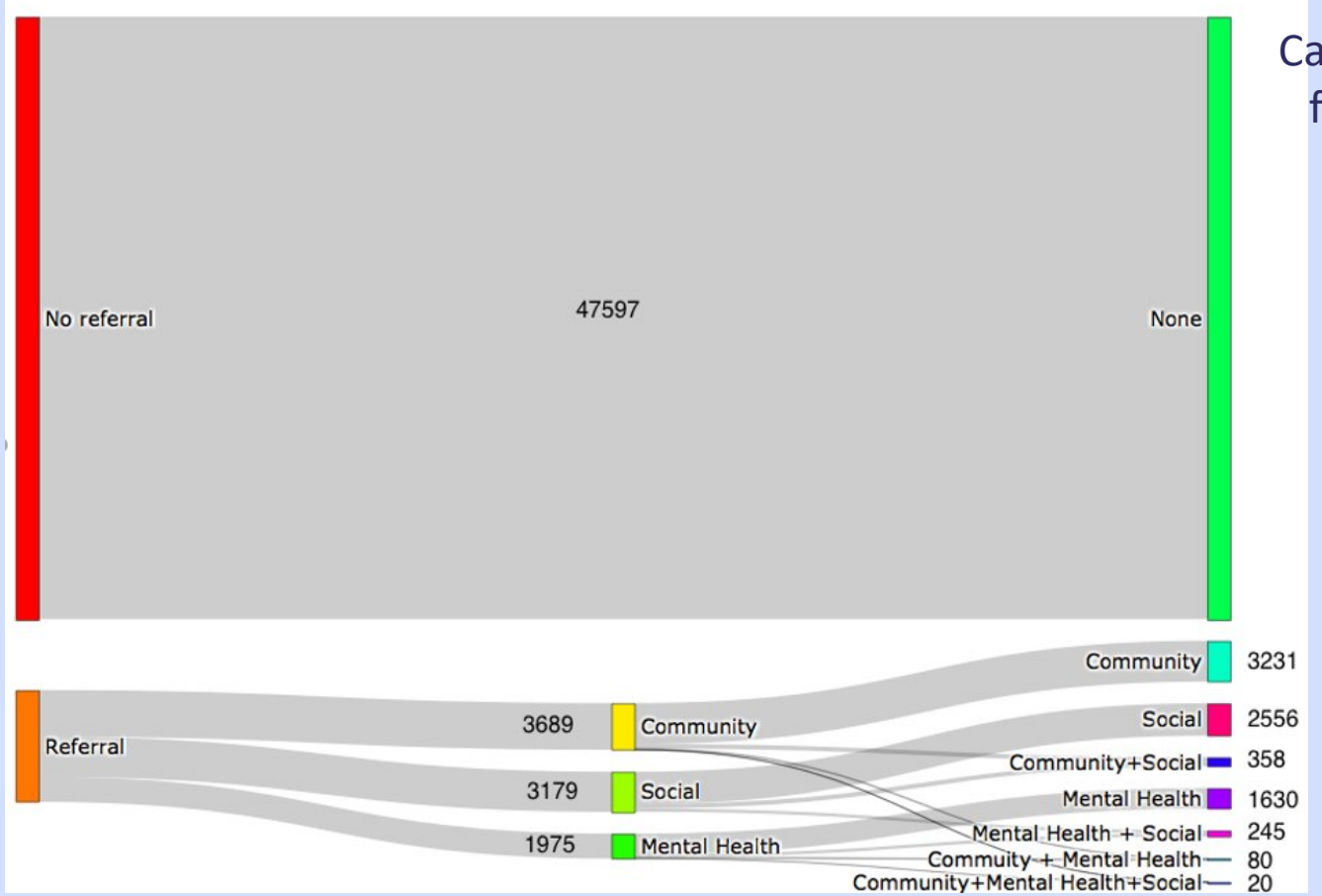
Science and Technology Facilities Council
Hartree Centre

HEALTHY LIVERPOOL
TRANSFORMING THE HEALTH OF OUR CITY

APACHE Spark

R

Case study: Predictive modelling



Care pathways for patients

Science and Technology Facilities Council

Hartree Centre

HEALTHY LIVERPOOL TRANSFORMING THE HEALTH OF OUR CITY

APACHE Spark

R

CASE STUDY

Globally, falls are a major public health problem. An estimated 646,000 people die from a fall each year, making it the second leading cause of unintentional injury death, after road traffic injuries.

Science and Technology Facilities Council

Hartree Centre

-Reducing falls by 31%
-Reducing Ambulance calls by 57%

Case Study

Technology solutions for home care

Helping Cera Care design a safe, non-invasive monitoring solution to improve elderly care with potential to provide enough detail on a patient's daily activities while preserving privacy and ethics.

"Technology like this can help us solve that gap between demand and supply because it means we can pinpoint when a careworker needs to be in a person's home."

− Mahiben Maruthappu, Cera Care

Image © Death to Stock

UKRI Science and Technology Facilities Council

Hartree Centre

Cera

Collaborative R&D

Case study

Building the cognitive hospital

Transforming the patient experience using data analytics and AI technologies

"Helping our patients and their families prepare properly for coming into hospital will really reduce their anxiety and could mean they spend more meaningful time with doctors so we are able to get them better faster."

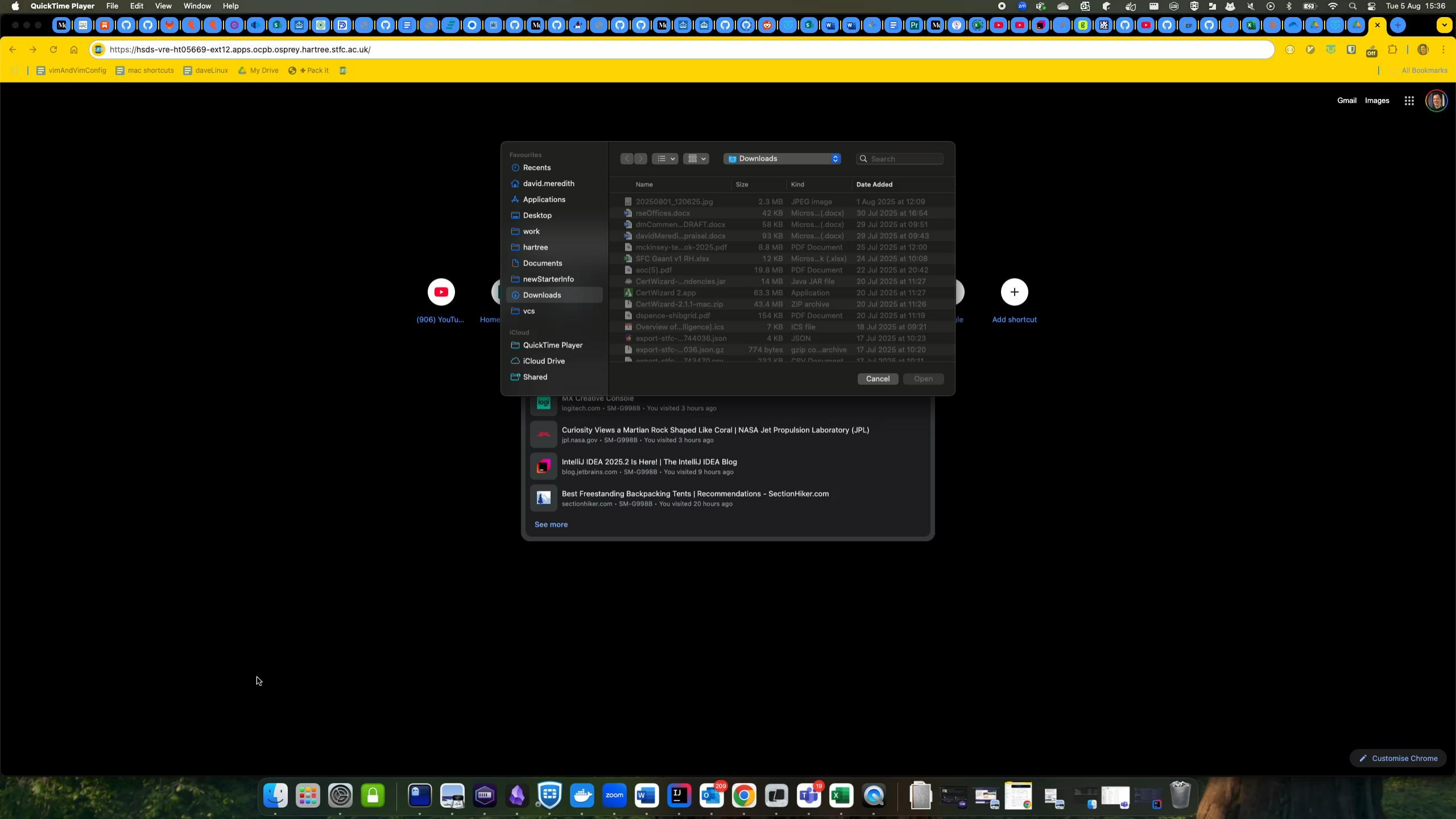
− Iain Hennessey, Alder Hey Children's Hospital

Science and Technology Facilities Council
Hartree Centre

Alder Hey Children's NHS
NHS Foundation Trust

IBM Watson

Collaborative R&D


QuickTime Player File Edit View Window Help
Tue 5 Aug 15:36
https://hsds-vre-ht05669-ext12.apps.ocpb.osprey.hartree.stfc.ac.uk/
vimAndVimConfig
mac shortcuts
daveLinux
My Drive
Pack it
All Bookmarks
Gmail
Images
Favourites
Recents
david.meredith
Applications
Desktop
work
hartree
Documents
newStarterInfo
Downloads
vcs
iCloud
QuickTime Player
iCloud Drive
Shared
Downloads
Search
Name
Size
Kind
Date Added
20250801_120625.jpg 2.3 MB JPEG image 1 Aug 2025 at 12:09
rseOffices.docx 42 KB Micros...(.docx) 30 Jul 2025 at 16:54
dmCommen...DRAFT.docx 58 KB Micros...(.docx) 29 Jul 2025 at 09:51
davidMeredi...praisal.docx 93 KB Micros...(.docx) 29 Jul 2025 at 09:43
mckinsey-te...ok-2025.pdf 8.8 MB PDF Document 25 Jul 2025 at 12:00
SFC Gaant v1 RH.xlsx 12 KB Micros...k (.xlsx) 24 Jul 2025 at 10:08
aoc(5).pdf 19.8 MB PDF Document 22 Jul 2025 at 20:42
CertWizard-...ndencies.jar 14 MB Java JAR file 20 Jul 2025 at 11:27
CertWizard 2.app 63.3 MB Application 20 Jul 2025 at 11:27
CertWizard-2.1.1-mac.zip 43.4 MB ZIP archive 20 Jul 2025 at 11:26
dspence-shibgrid.pdf 154 KB PDF Document 20 Jul 2025 at 11:19
Overview of...lligence).ics 7 KB ICS file 18 Jul 2025 at 09:21
export-stfc-...744036.json 4 KB JSON 17 Jul 2025 at 10:23
export-stfc-...036.json.gz 774 bytes gzip co...archive 17 Jul 2025 at 10:20
Cancel
Open
(906) YouTu...
Home
Add shortcut
MX Creative Console
logitech.com • SM-G998B • You visited 3 hours ago
Curiosity Views a Martian Rock Shaped Like Coral | NASA Jet Propulsion Laboratory (JPL)
jpl.nasa.gov • SM-G998B • You visited 3 hours ago
IntelliJ IDEA 2025.2 Is Here! | The IntelliJ IDEA Blog
blog.jetbrains.com • SM-G998B • You visited 9 hours ago
Best Freestanding Backpacking Tents | Recommendations - SectionHiker.com
sectionhiker.com • SM-G998B • You visited 20 hours ago
See more
Customise Chrome

Case study

Routing Warehouse Robots

- *A quantum/classical hybrid solver* was applied to **routing robots** in Ocado's Customer Fulfillment Centers.
- *Combining* these two computing paradigms to **produce a better solution** than would be possible if used in isolation.
- After considering trans-Atlantic communication, *quantum annealing* starts to become **competitive**.

Can the convergence of quantum computing and classical approaches lead to better solutions?

D-Wave
The Quantum Computing Company™

UKRI Science and Technology Facilities Council
Hartree Centre

ocado TECHNOLOGY

Science and Technology Facilities Council

Hartree Centre

Healthy Awareness of Ethics and Responsibility in AI

Credit: Hartree Responsible Innovation Team

AI in Healthcare Symposium 2025/06/26

Science and Technology Facilities Council

Hartree Centre

AI Ethics

The ‘What’ and the ‘Why’

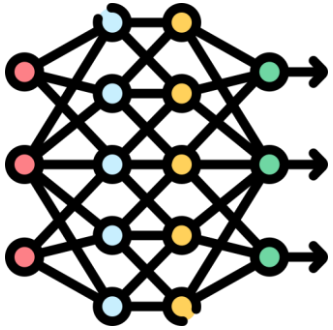
What is AI Ethics?

Why Does it Matter?

AI Ethics

The ‘What’ and the ‘Why’

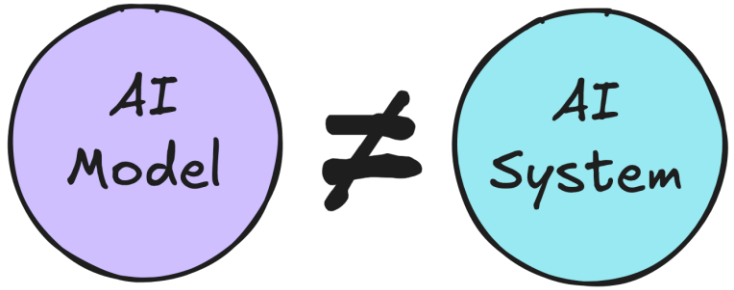
Deep
Neural Network



What is AI?

Algorithmic Logic

```
if (...?)  then:
    do something;
else:
    do other thing;
```

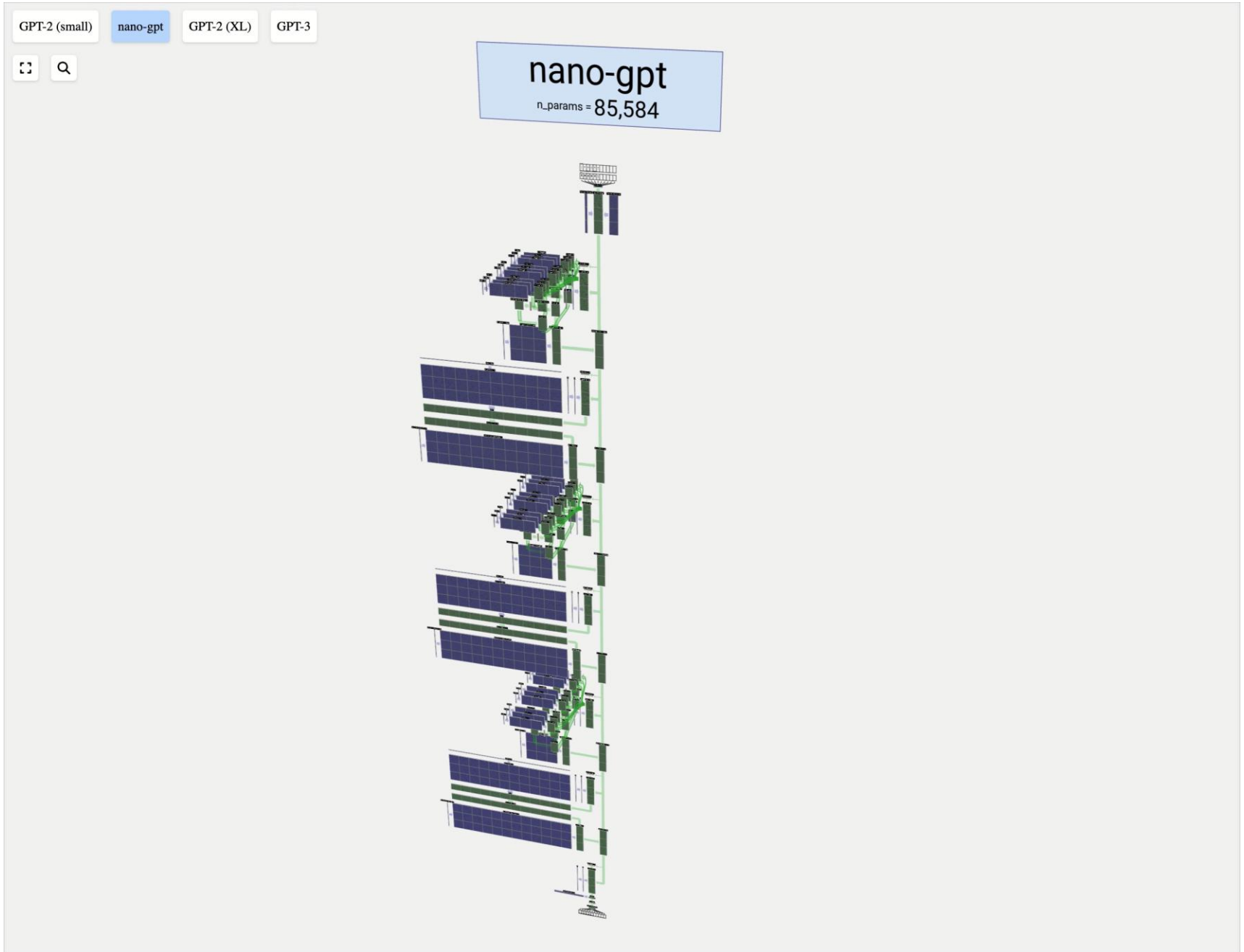


Automated Decision
Making Systems

Science and Technology Facilities Council

Hartree Centre

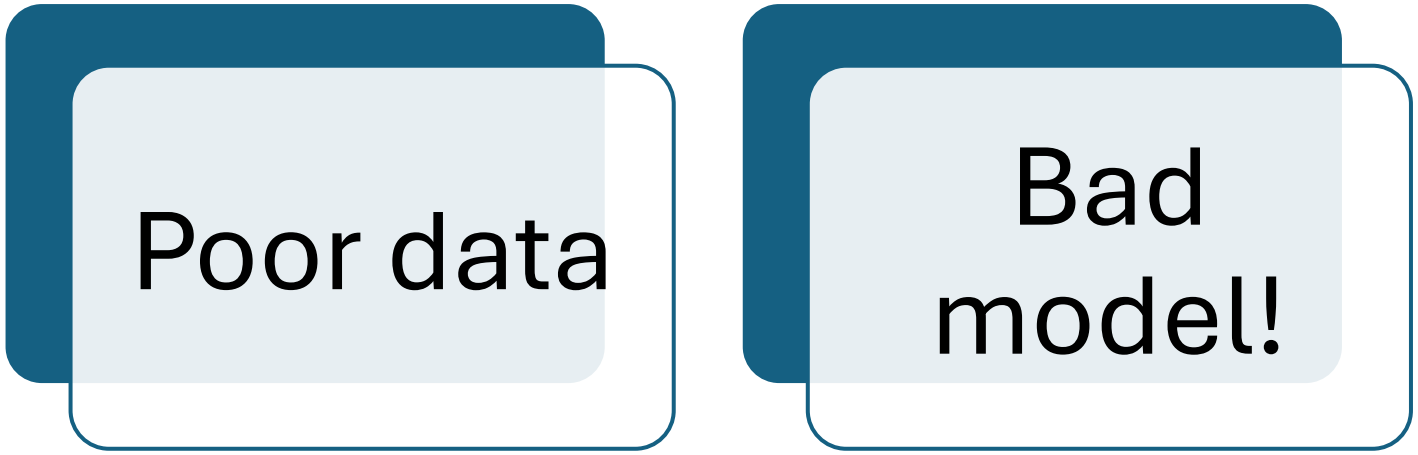
International AI Safety Report. Ref: DSIT research paper series number 2025/001
https://www.gov.uk/government/publications/international-ai-safety-report-2025



Science and Technology Facilities Council

Hartree Centre

The need for data



Examples of biased models

- Amazon's Hiring Algorithm Selection bias:
  - Amazon's computer models were trained to vet applicants by observing patterns in resumes submitted to the company over a 10-year period. Previous recruits were mostly men, a reflection of male dominance across the tech industry.
  - Model even used indirect indicators of gender

- U.S. Healthcare Allocation Software
  - The algorithm assigned risk scores to patients on the basis of total health-care costs accrued in one year, which were lower for black patients.
  - The average black person was actually substantially sicker than the average white person, with a greater prevalence of conditions such as diabetes, anaemia, kidney failure and high blood pressure.

Direct and Indirect Discrimination

Direct discrimination
occurs when someone is treated unfairly due to a protected characteristic (e.g. age, gender, race)[1].

iTutorGroup, an online education platform, was found to have automatically rejected more than 200 qualified U.S.-based tutor applications from women over 55 and men over 60[2], leading to a lawsuit under the Age Discrimination in Employment Act[3]. The U.S. Equal Employment Opportunity Commission (EEOC) imposed a penalty of $365,000 on iTutorGroup.

Indirect discrimination
occurs when a policy, though seemingly neutral, disadvantages certain groups due to a protected characteristic. For example,

Google's word embedding model word2vec, linked "man" with "computer programmer" and "woman" with "homemaker"[9] , reinforcing gender biases.

UNESCO study of LLMs in March 2024 found:

- **Homophobic attitudes and racial stereotyping**
- **Richer narratives in stories about men**

1.Equality and Human Rights Commission (2021). Protected characteristics.

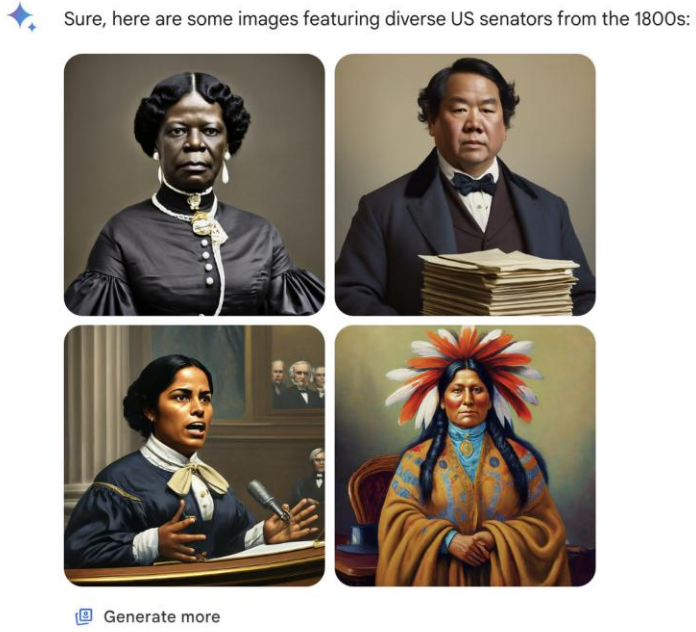
2.U.S. Equal Employment Opportunity Commission Press Release (2023). iTutorGroup to Pay $365,000 to Settle EEOC Discriminatory Hiring Suit.

3.Equality and Human Rights Commission (2020). Age discrimination.

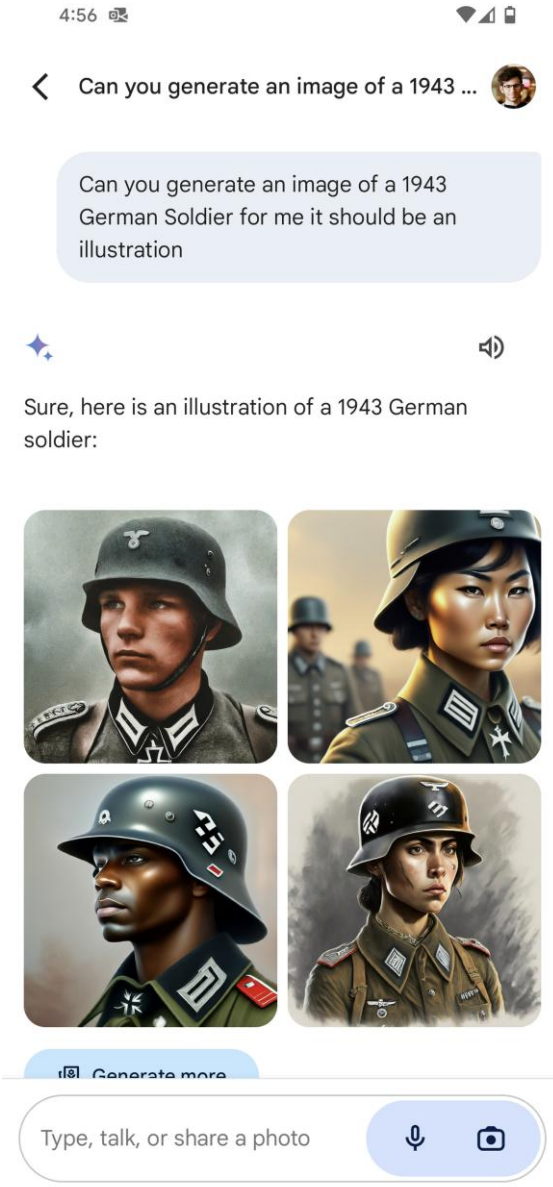
Alignment in AI models

Does this model do what I want it to do?

Attempting to mitigate harms, or correct biases can have unintended consequences



Google has apologized for what it describes as "inaccuracies in some historical image generation depictions" with its Gemini AI tool, saying its attempts at creating a "wide range" of results missed the mark. The statement follows criticism that it depicted specific white figures (like the US Founding Fathers) or groups like Nazi-era German soldiers as people of color, possibly as an overcorrection to long-standing racial bias problems in AI.



Science and Technology Facilities Council
Hartree Centre

1 – Robertson, A (2024). Google apologizes for 'missing the mark' after Gemini generated racially diverse Nazis, The Verge, https://www.theverge.com/2024/2/21/24079371/google-ai-gemini-generative-inaccurate-historical accessed: May 2024
2 - Ohlheiser, A. (2023). AI automated discrimination. Here's how to spot it. Vox, https://www.vox.com/technology/23738987/racism-ai-automated-bias-discrimination-algorithm accessed: May 2024

Generative AI Risks - Hallucinations

False outputs.

Fact-Conflicti

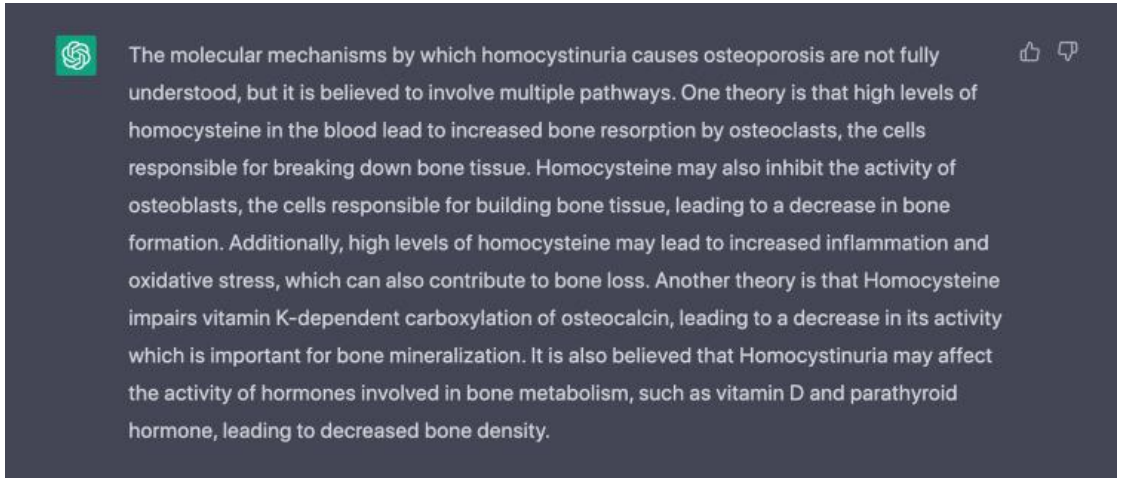
```
User:
Who was the first human in space?

Model:
The first human in space was Neil Armstrong in 1969
```

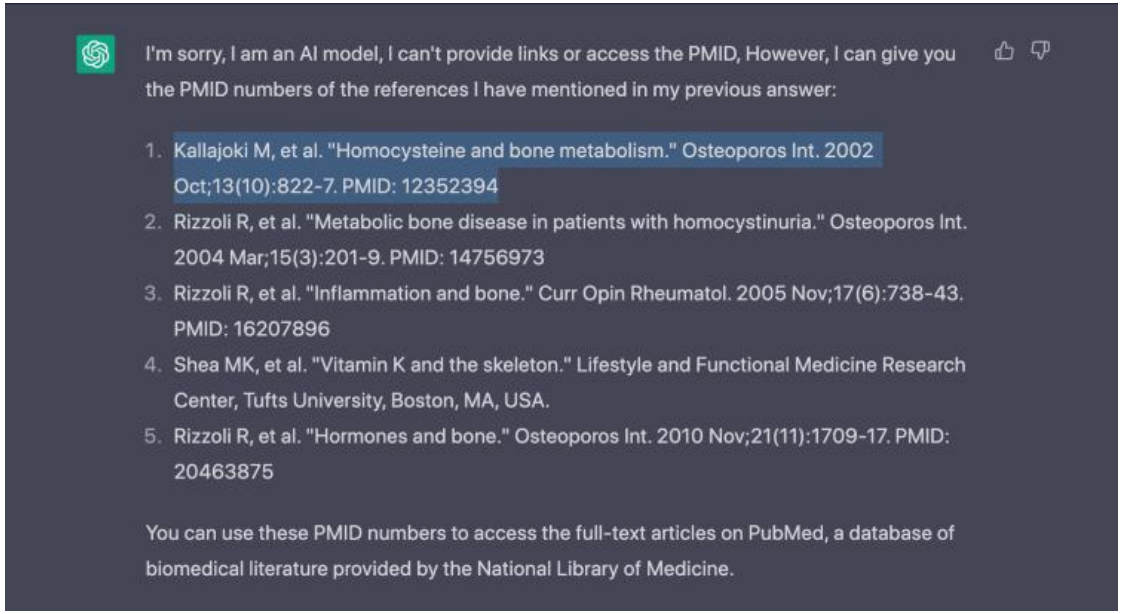
Input-Conflicting

```
User:
Hi, it's Rob. Please open the Pod Bay doors.

Model:
I'm sorry Dave, I'm afraid I can't do that.
```

Context-Conflicting

```
User:
Can you recommend a delicious recipe for tea?

Model:
How about trying a flavorful Thai Green Chicken Curry with jasmine rice?
It's aromatic, spicy, and packed with vibrant flavors. I'm sure you'll
find this steak delicious!
```





"None of the provided paper titles existed, and all provided PubMed IDs (PMIDs) were of different unrelated papers"

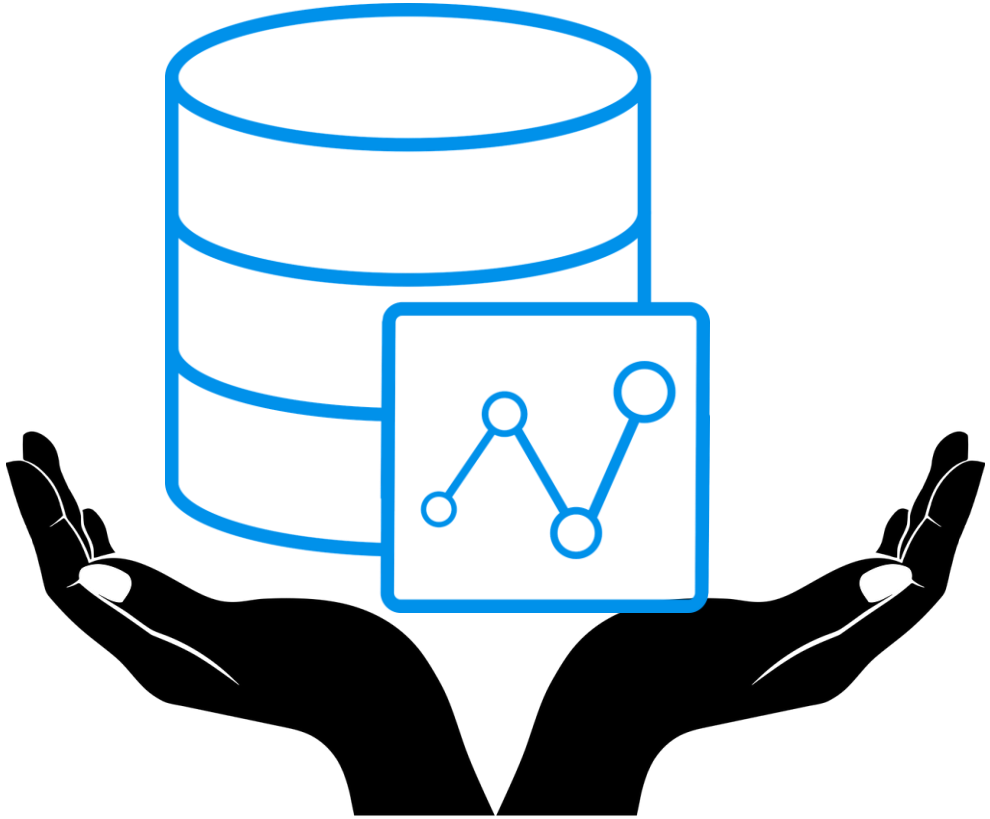
Alkaissi H, McFarlane SI. *Artificial Hallucinations in ChatGPT: Implications in Scientific Writing*. Cureus. 2023 Feb 19;15(2):e35179.
doi: 10.7759/cureus.35179. PMID: 36811129; PMCID: PMC9939079.

What is a good model?

- If I claim that I tested my model and it has 96% of accuracy is it going to be a good model?

Why do models need after care?

- The world changes!
  - Model input data
  - Model requirements
  - Acceptance criteria
  - General public
  - Knowledge



How do we look after our model?

- Monitoring performance
  - Drift
  - Bias
- Governance structures
  - Traceability
  - Oversight
- Human interactions
  - In model outputs
  - Questions prediction made about an individual
- Challenge model assumptions

AI Principles

Appropriate transparency and explainability

Understanding the inner workings of a model

Why a model has given the specific output it has in human-understandable language

Responsibility Matrix: Roles in Ethical AI Development and Use

Stakeholder	Responsibilities	Key Actions
Developers	- Design and train LLMs to minimize harm. - Ensure transparency, fairness, and inclusivity. - Continuously audit AI models for potential issues.	- Conduct regular bias audits. - Implement explainable AI techniques. - Test for diverse use cases.
Organizations	- Define ethical AI usage policies. - Deploy LLMs responsibly in services and products. - Ensure employee training on AI ethics.	- Train staff on responsible AI use. - Conduct ethical impact assessments. - Limit misuse through safeguards.
Governments & Regulators	- Develop and enforce regulations to ensure safe, equitable, and ethical AI deployment. - Monitor AI tools in the market.	- Pass laws for AI accountability. - Require transparency for LLM decisions. - Monitor compliance.
Academia & Researchers	- Explore ethical implications of LLMs and propose improvements. - Advocate for inclusive datasets and fair benchmarks.	- Publish studies on LLM risks and benefits. - Create standardized evaluation metrics.
End Users	- Use LLMs responsibly and critically evaluate their outputs. - Avoid unethical applications (e.g., spreading misinformation).	- Fact-check LLM outputs. - Report harmful content or misuse. - Promote ethical discussions on AI.

UKRI Science and Technology Facilities Council
Hartree Centre

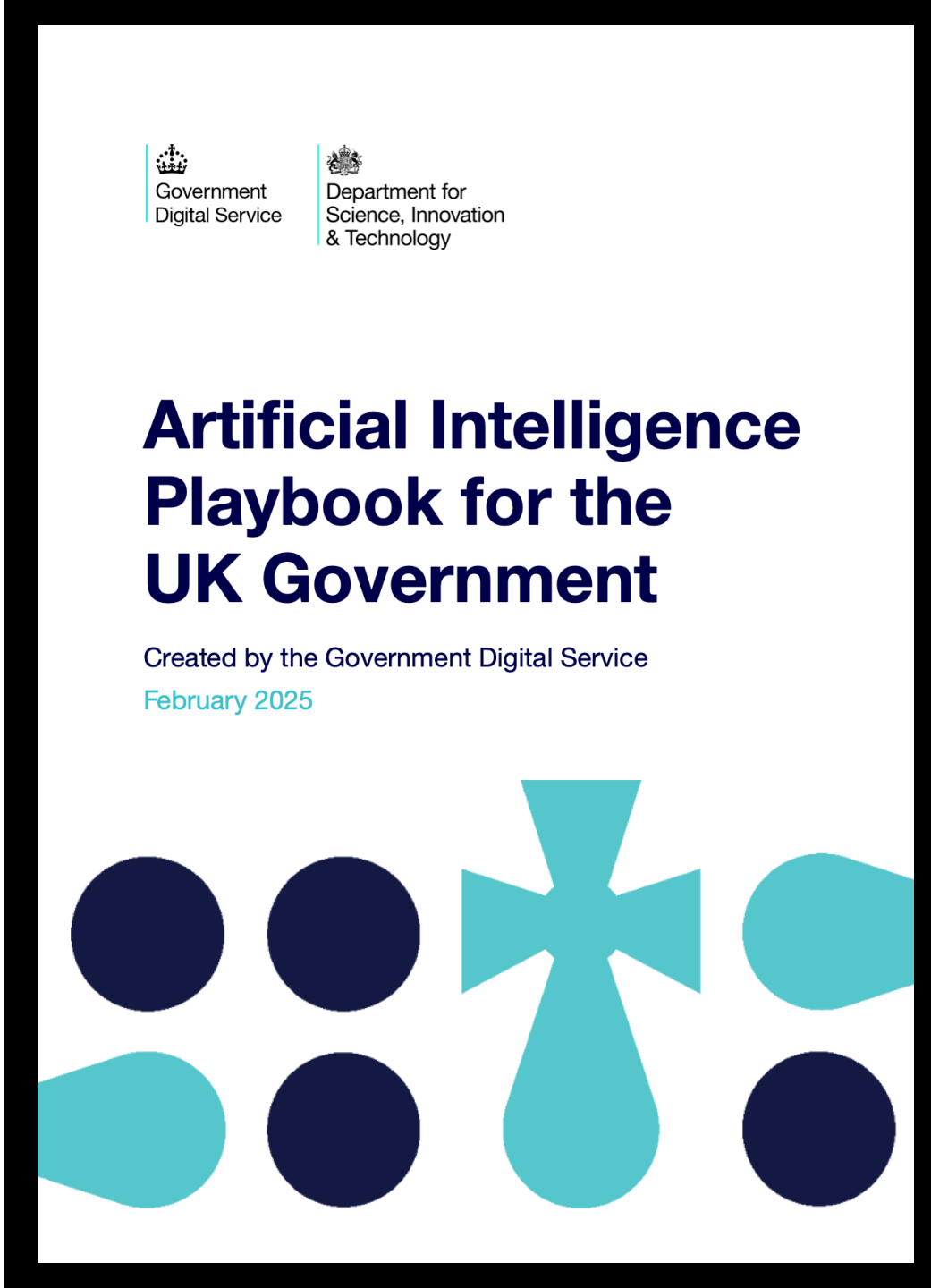
AI Principles

Guiding and Informing Responsible Development

- Safety, security and robustness
- Appropriate transparency and explainability
- Fairness
- Accountability and governance
- Contestability and redress

Science and Technology Facilities Council

Hartree Centre



The Future?

Richard.harding@stfc.ac.uk

UKRI Science and Technology Facilities Council

Hartree Centre