IMPERIAL

Information and Communication Technologies (ICT)

Building Sustainable Digital Research Infrastructure: Data Centre(s)

9 September 2025

Dr Andrew Richards
a.j.richards@imperial.ac.uk

Our Goals
Bringing the Digital
Plan to life

Digital Plan

1 Create World-class Digital Education Experience

2 Enable a Digital Campus

3 Digitally Enhanced Professional Services

4 Embed a Research and Enterprise Data Culture

5 Enable Research & Innovation

Imperial College London

Research Computing Themes

- **Hybrid research computing platforms**
  - On-premise, cloud, interactive, responsive

- **FAIR Research Data platforms**
  - Store, process, share, archive, curate data

- **Research Software Engineering Hub**
  - Develop, advise, define standards, accelerate research

- **Underpinning research management applications**
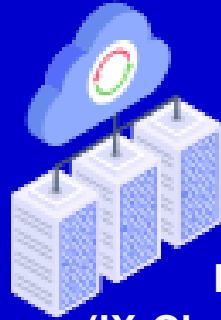 - Support end to end research lifecycle applications
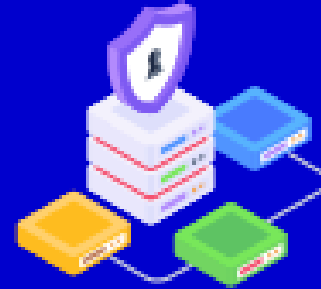
Research Computing Solutions

(*in development)

Compute/ *Analysis*
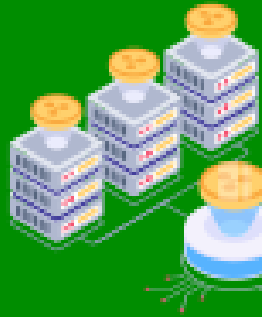
HPC/HTC
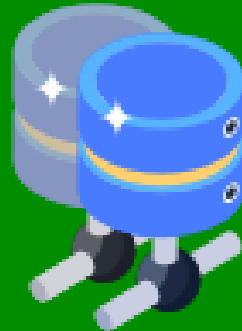(CX3, HEX) (HX1/HX2)

Interactive
(IX-Cloud, *HEX-AI (HX3)*)

TRE
(Sensitive/Secure (*HX4*))

Storage

RDF – HPC
(2PB/System)

RDF – Active
(10PB Dual site)

RDF – ActiveSecure*

Research **Data Library** *(Archive and Publish)*

RDF – FAIR
(SPIRAL) (*HELIX**)

RDF – Archive

RDF – Archive (Secure)*

Research Technical Professionals

Imperial Data Facility

Research (Dept/Faculty/Facility)

Active (ICT)

Archive (Institutional)

RCS HPC

Dept HPC

Dept Research Equipment

Dept Research Data Stores

External Facility Research Data

RDF Active (10PB Dual site)

RDF ActiveSecure

SPIRAL

HELIX (FAIR)

RDF Archive

DART / DMP Registration

Minimal Metadata Applied

INVENIO RDM
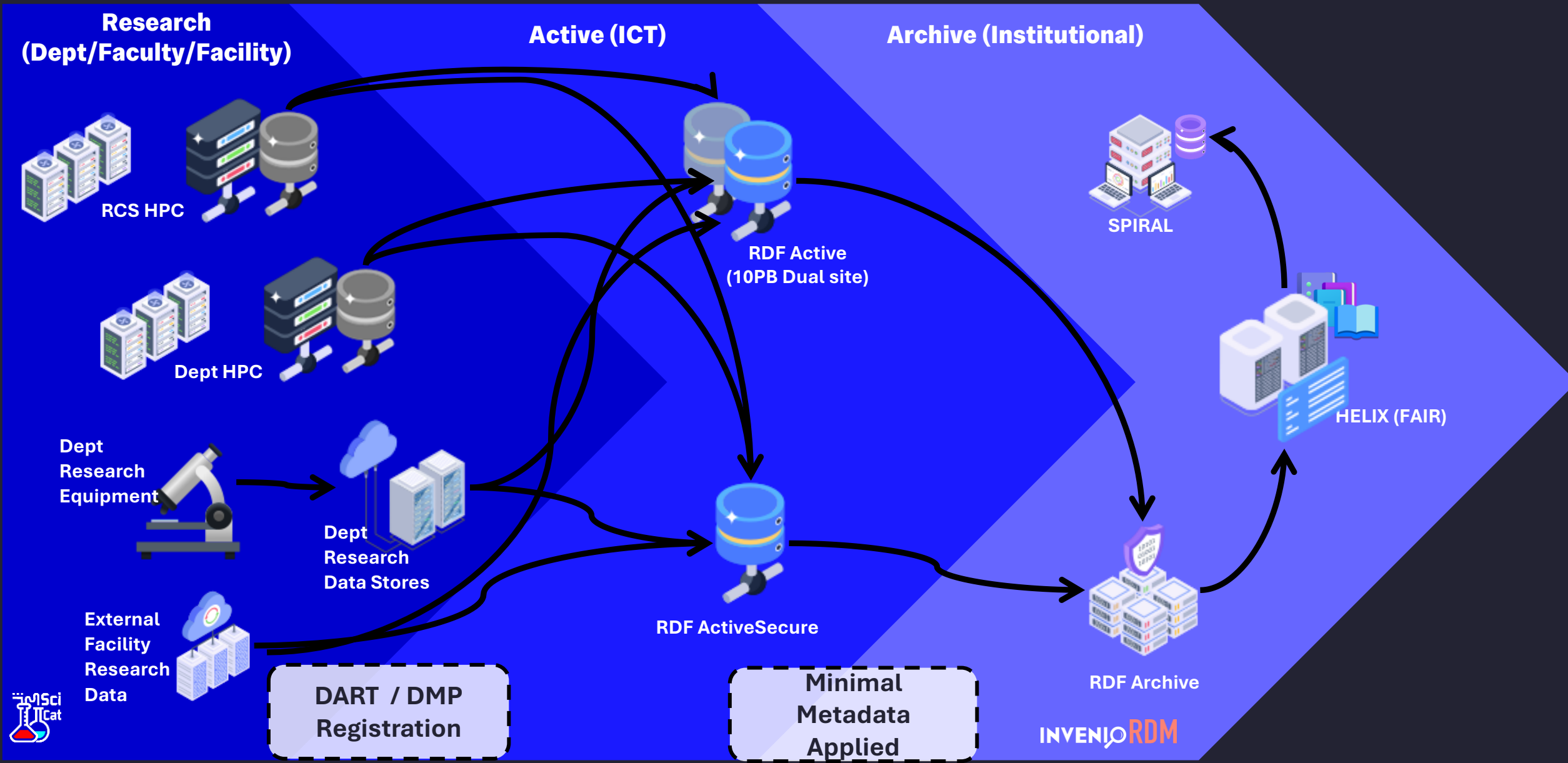
Imperial Research Compute: HEX Platforms

HEX-HTC

HEX-HPC

HEX-AI

HEX-INovArc

HEX-Secure

HEX-Visual

HX1 HX2 HX3 HX4 HX5 HX6

Data
RDF
Spiral +
Helix

Here Comes The Cheese!*

+++ Please Reinstall Universe And Reboot*

ICICLE stands for Imperial College (IC) Intel Corporation (IC) Lenovo (LE)
*Quotes by Terry Pratchett, Discworld

Training

01/10/2025
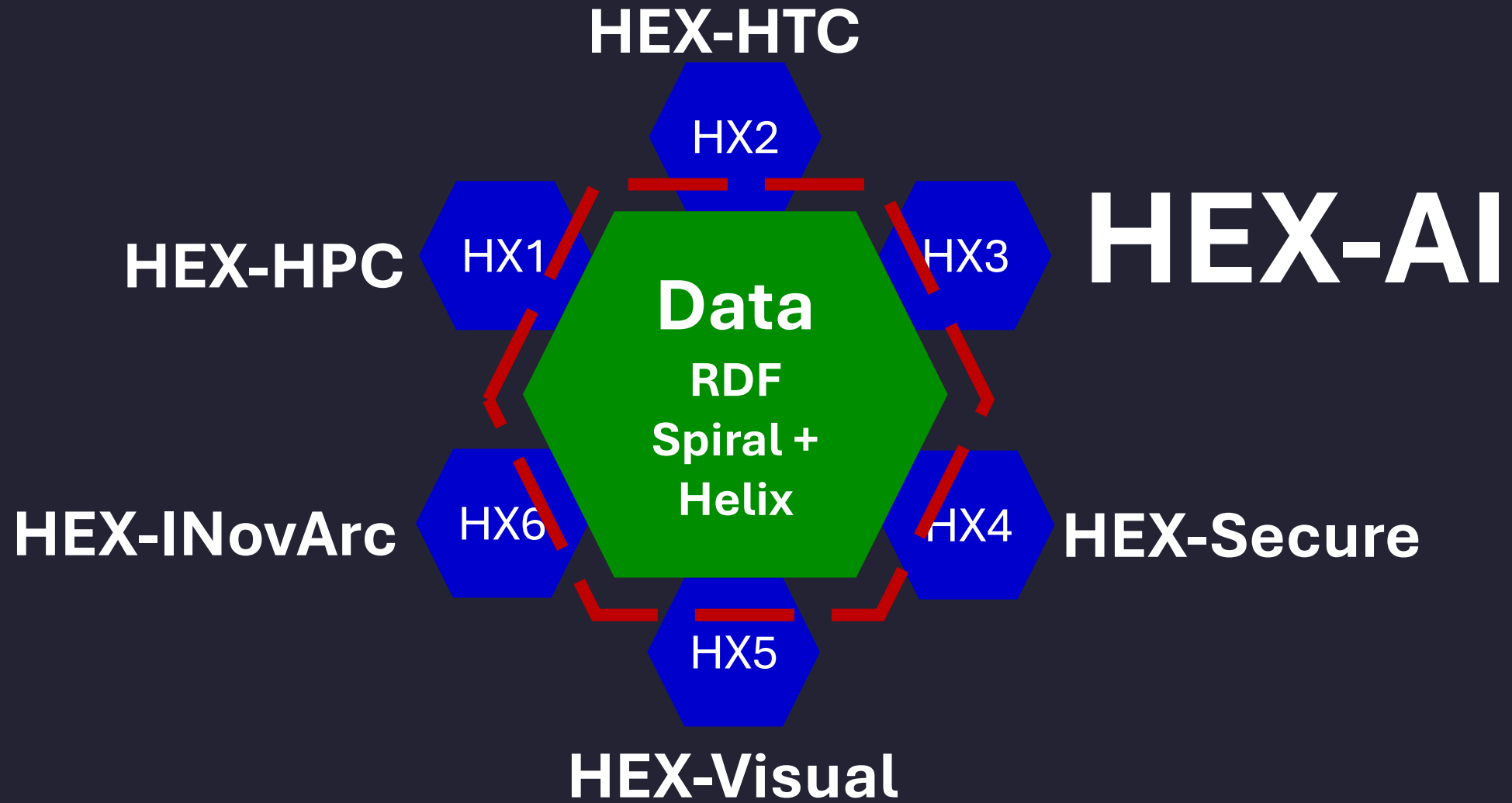
Scale of facilities across Higher Education sector

Institution	Total Cores	Total GPUs
Massachusetts Institute of Technology	118,780	1,844
University of Cambridge	120,744	320
Harvard University	91,488	409
Imperial College of Science, Technology and Medicine	84,944	100
University of Oxford	23,000	258
Stanford University	52,260	704
University College London	41,388	114
University of Southampton	31,200	60
University of Leeds	12,496	40
University of Manchester	17,244	100
California Institute of Technology	16,840	200
University of Bristol	16,072	65
University of California, Berkeley	11,620	60
University of Birmingham	8,192	6
University of Edinburgh	7,000	124
King's College London	6,088	0

Table 1: Publicly available information for centrally provided compute resources.

- As a university facility the services provided are at a scale comparable to typical Tier-2 HPC centres in the UK.
- Publicly available information from a 2023 report produced by RedOak Consulting
- It is noted that from other, direct, sources known resources at some locations such as Edinburgh and Oxford are greater than listed publicly. However, the figures give some indication of comparable resources at other institutes.
- Planned investment would firmly place Imperial with sector leading facilities.

Compute Platform Summary

CX3

325 compute nodes with 2x AMD EPYC 7742
34 computes nodes with 2x Intel Icelake Xeon Platinum 8358 (64 cores, 500GB RAM).
11 GPU nodes with 2x AMD EPYC 7742 (128 cores, 1TB RAM, 8 Quadro RTX 6000 per node)
10 GPU nodes with 2x Intel Skylake 4110 (16 cores, 192GB RAM, 8 Quadro RTX 6000 per node)
12 large mem compute nodes with 2x AMD EPYC 7742 (128 cores, 4TB RAM per node).
Interconnect: 100GbE

Total: 397 compute nodes, 46,880 cores, 403TB RAM, 168 RTX 6000

HX1

303 compute nodes:
Lenovo SD630v2 servers each with 2 x Intel Xeon Platinum 8358 (Ice Lake) 64 cores per node;
288 nodes; 18,432 compute cores; 512 GB RAM per node

GPU nodes:
Lenovo servers each with 4 x NVIDIA A100 80 GB RAM GPUs;
2 x Intel Xeon Platinum 8360Y (Ice Lake) 2.40GHz 36-core processors; 1 TB RAM per node; 15 nodes;
60 GPUs in total
Interconnect: Mellanox ConnectX-6 HDR200 (200 Gbit/s) InfiniBand

Total: 303 compute nodes, 19,512 cores, 159 TB of RAM, 60 A100 GPUs

IMPERIAL / intel. / Lenovo

Imperial College London, Intel Corporation, Lenovo

Technical partnership to:

- Provide next generation(s) of HPC & storage platforms for the university

- To support the sustainability journey through the adoption of direct water-cooled solutions

- To foster collaboration around research computing with academic partners

PhD Studentships

- Quantum Computing Algorithms.
  - Developing and optimising algorithms to use existing Quantum Computers.
  - Fast ML: Using accelerators for Machine Learning.
  - FPGAs and other accelerators.

RTP Experience Program Support.
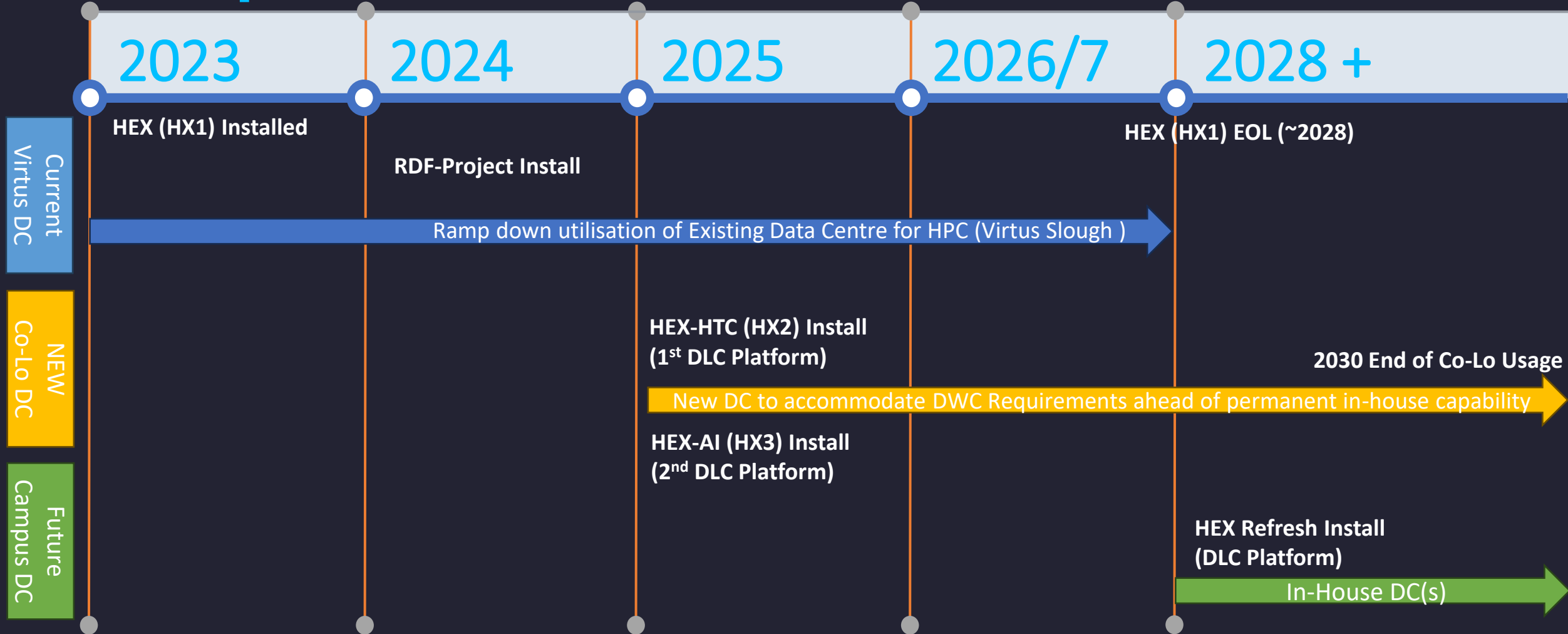
HPC 2022 Refresh

Research Computing Net Zero Data Centre Roadmap

| 2023 | 2024 | 2025 | 2026/7 | 2028 + |

HEX (HX1) Installed

RDF-Project Install

HEX (HX1) EOL (~2028)

Current Virtus DC

Ramp down utilisation of Existing Data Centre for HPC (Virtus Slough)

NEW Co-Lo DC

HEX-HTC (HX2) Install (1st DLC Platform)

2030 End of Co-Lo Usage

New DC to accommodate DWC Requirements ahead of permanent in-house capability

HEX-AI (HX3) Install (2nd DLC Platform)

Future Campus DC

HEX Refresh Install (DLC Platform)

In-House DC(s)

Imperial College London
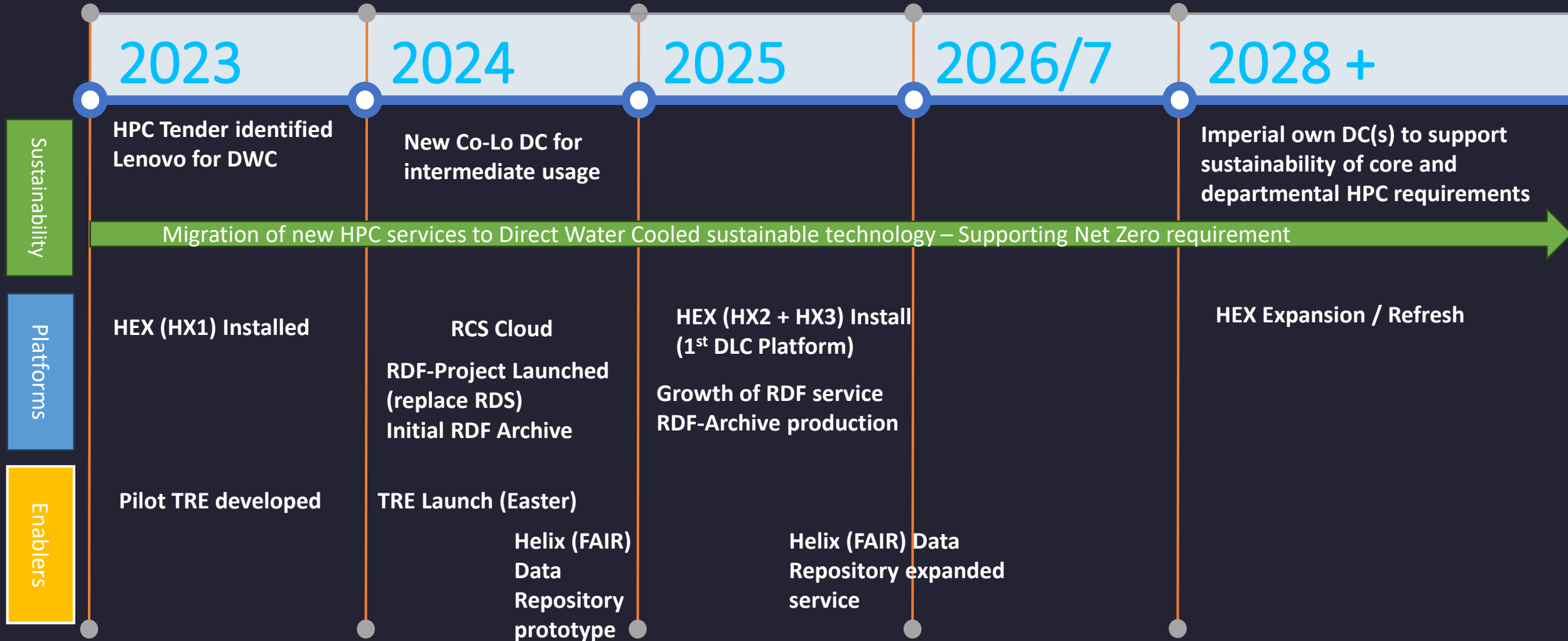
IMPERIAL

Information and Communication
Technologies (ICT)

DataCentre Planning

2025

Iain Darvell

Datacentre Roadmap

	2025	2026	2027	2028	2029	2030 >
South Kensington						
Slough (Virtus)						
Additional Colo						
New Datacentre X						
White City						

Strategic Datacentre (today)

Local Hosting Facility (On campus Server Rooms)

New Strategic Datacentre (Planned)
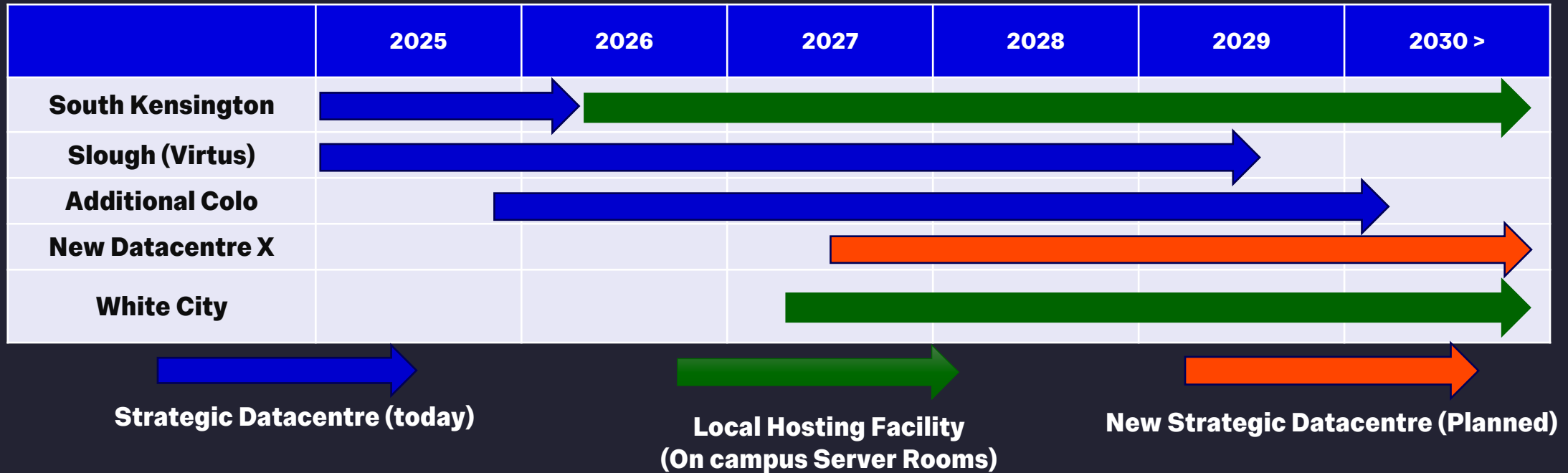
Key Points:

- Current Co-Lo datacentre MSA/contracts up for renewal in 2029 (Slough)
- Annual price increase by CPI currently – expecting material step up in cost upon renewal in 2029/2030

- Datacentre X - future requirement for one strategic datacentre
    - Location and potential costs to be identified and compared with future Colo options

- Requirement also for Local Hosting Facilities at both South Kensington and White City
    - Property/Strategic Change Team to confirm White City can accommodate Local Hosting Facility
    - South Kensington City & Guilds datacentre will continue

Imperial College London

Strategic Data Centre – Ready for 2029

- Research intensive HE institutions continue to have long term datacentre demand given the computing infrastructure required to enable high-performance research computing
- Datacentre and electrical grid capacity are in short supply resulting in high costs for 3rd party (Colo) datacentre capacity.
- Ask of Property is to identify any suitable location within Imperial's property portfolio

Requirements	
Location	TBD
Power	4MW then 6MW
Racks	30 Water Cooled (110kW ea) 70 Air/Water rear cooled (30kw ea)
Data Hall Room Size	400m2
Supporting Plant Size (UPS, Switchgear etc)	200m2
External Chiller Plant Size	300m2
Build type	Modular or Traditional
Physical Security	Access controls, 24x 7 security
Benefits	- Potential reuse of heat - Reduce the reliance, risks and constraints of external datacentre providers - Avoiding uptick in hosting costs 2029 onwards

Local Hosting Facility (On Campus Server Rooms)

To meet computing and hosting needs with a focus on equipment in need of regular physical interaction with researchers/students.

Details	White City (To be built)	South Kensington
Location on Campus	TBD	City & Guilds (as is now) Sir Alex Fleming building (as in now)
Power	500kW	350kW
Racks	24	34(CG) & 9(SAF) total/8 free from Jan '26
Room Size	175m2	As is current provision
Supporting Plant Size	100m2	As is current provision
External Chiller Plant Size	100m2	As is current provision
Target Departments	North Campus	SK based Faculties/Departments
Benefits		• Centralised ICT facility at each campus avoids fragmented and inefficient departmental hosting rooms. • Ensures adequate power and cooling for research and experimental computing. • Lays the foundation for a more sustainable and cost-effective hosting

Current Data Centre Capacity

- ICT Racks - 171
- Total Floor Space – 738m2 (160 on campus)
- Contracted Power Rack Ratings – 2.25 MW
- Contracted PUE at Slough (1.35)
- New DC ~20% reduction in energy costs
- Increasing power density per rack requirements: 100kW -> 300kw +

| DataCentre	Capacity (mW) excluding cooling	Annual Cost (£) exc Power	Notes
Virtus Slough	1.95	2.8M	136 racks, UPS and Generator protection
South Kensington	0.3	unknown	35 racks, UPS on single feed
Digital Realty Woking	0.7	1.1M	24 racks (expandable to 1.3mW), delivery date is September 2025

100 % Renewable Energy ?
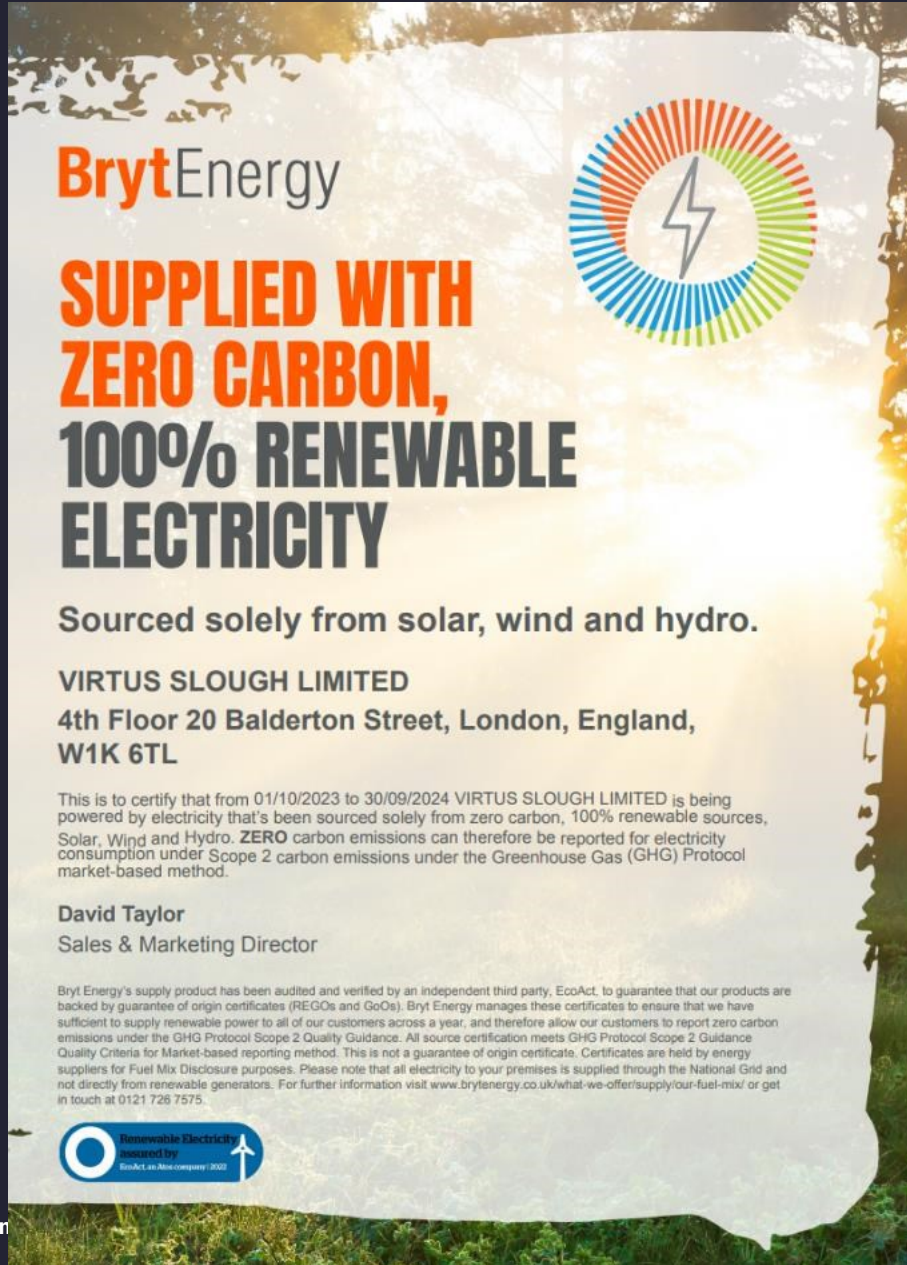
This is to certify that from 01/10/2023 to 30/09/2024 VIRTUS SLOUGH LIMITED is being powered by electricity that's been sourced solely from zero carbon, 100% renewable sources, Solar, Wind and Hydro. ZERO carbon emissions can therefore be reported for electricity consumption under Scope 2 carbon emissions under the Greenhouse Gas (GHG) Protocol market-based method.
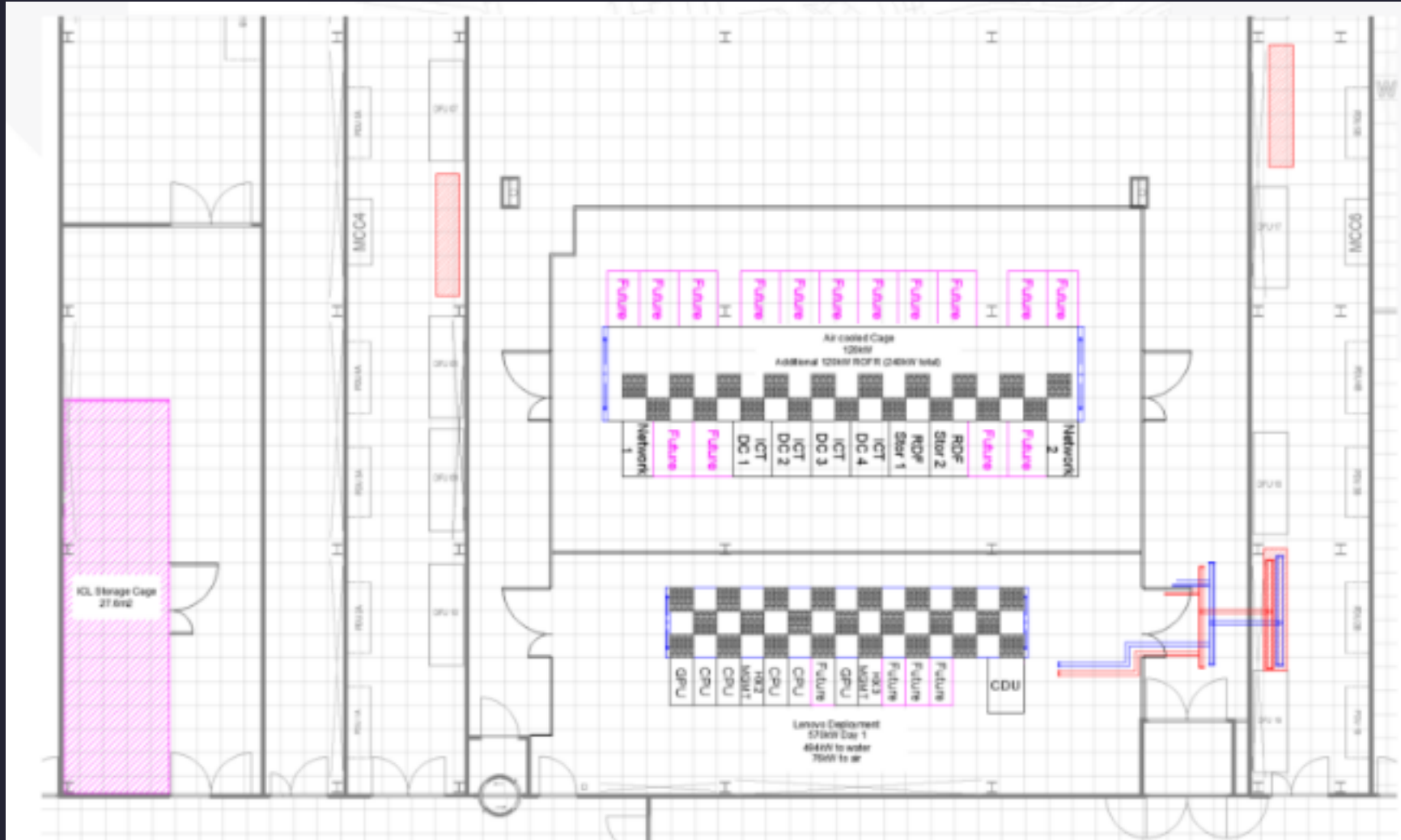
DLC DC Schematic Layout
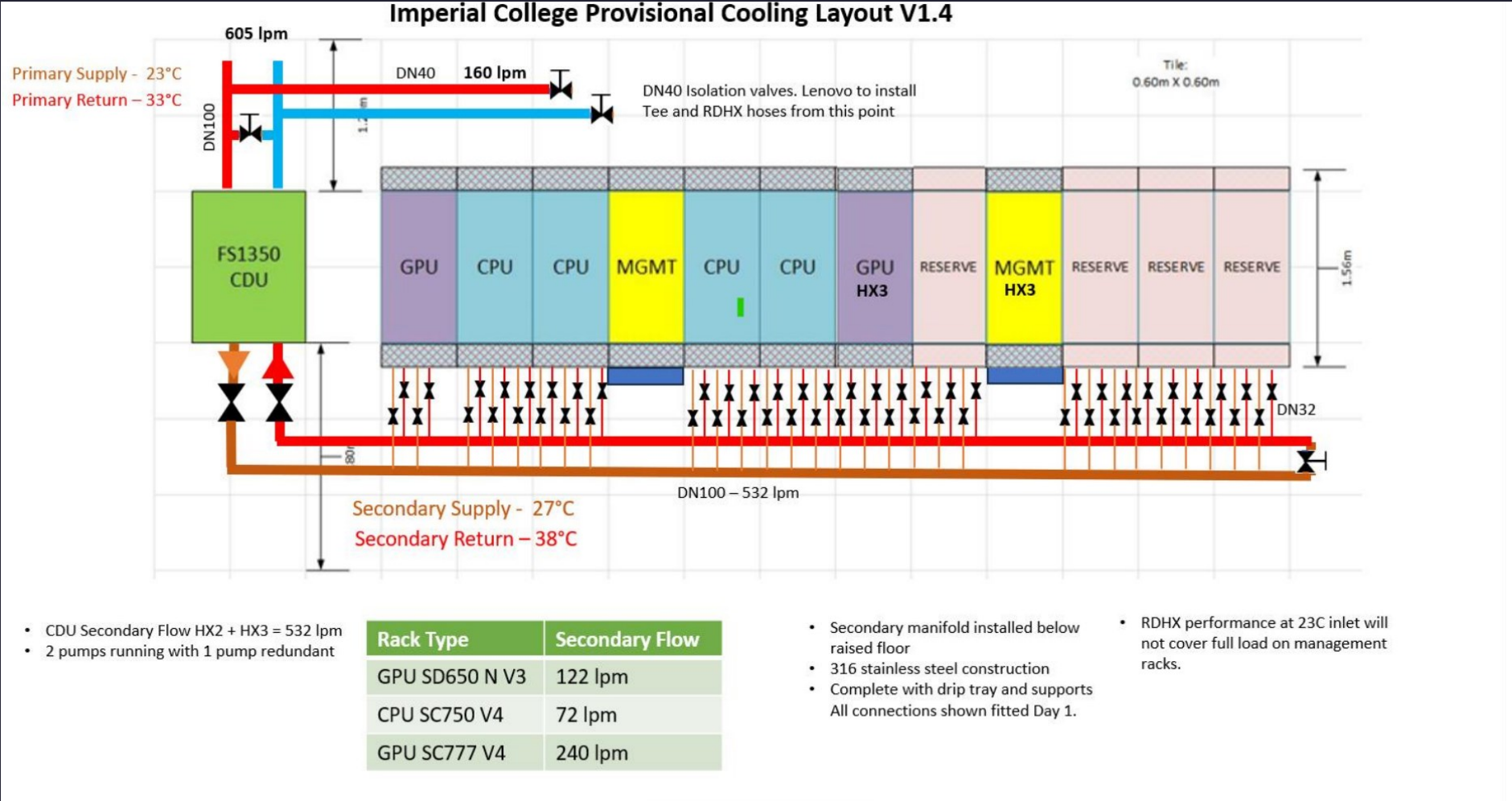
Imperial DLC Data Centre (Woking)
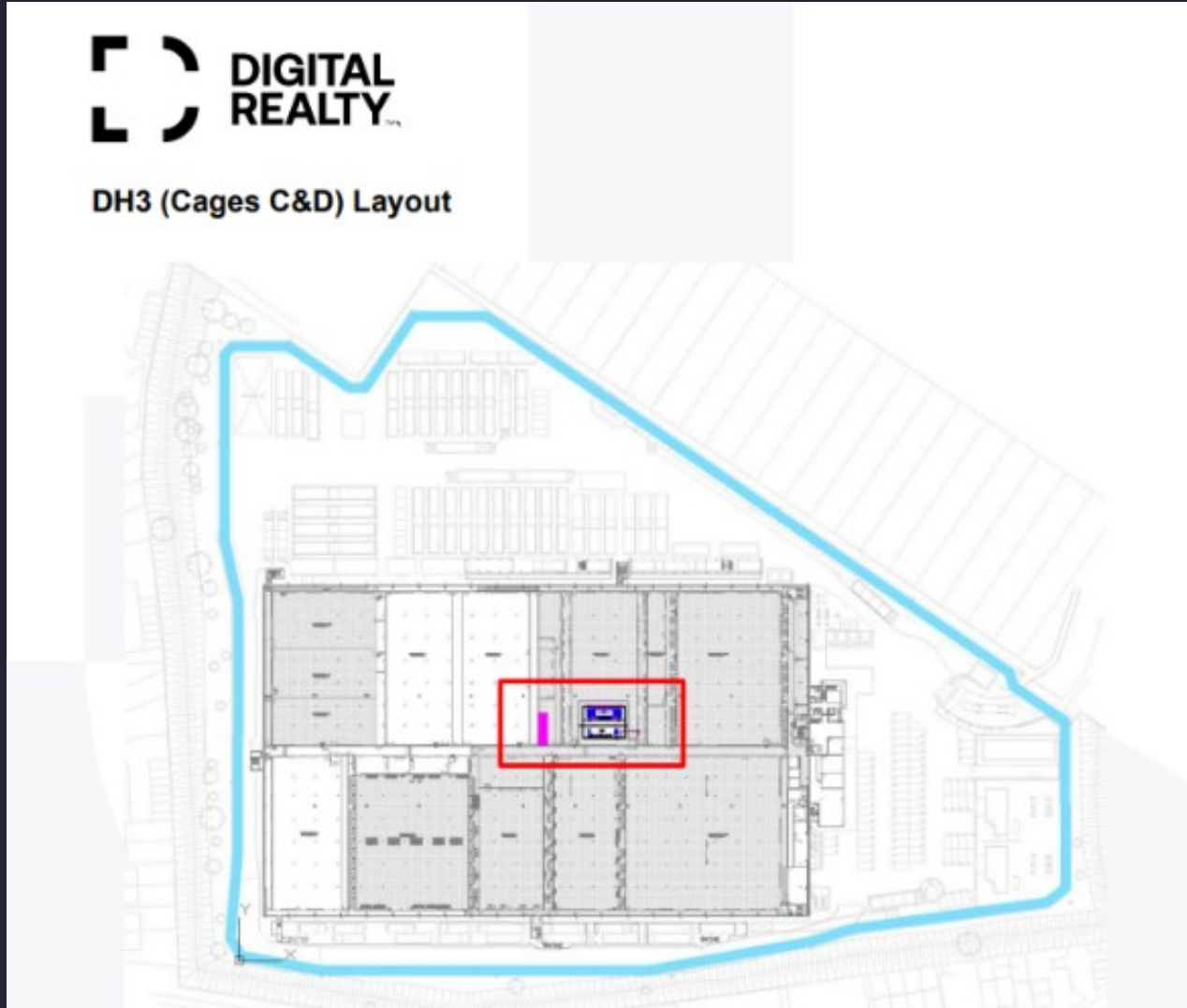
Imperial DLC Data Centre (Woking)

User Awareness
(Jack Coker – Imperial)

A **command line tool** that calculations greenhouse gas **emissions from a given HPC job**

Inspiration from online calculators which can make similar estimates

Inspiration from a similar tool available on the ARCHER2 system

The `jobemissions` tool is available by default to all ARCHER2 users from the command line. You supply a Slurm job ID for a completed job and the tool provides an estimate of the GHG emissions associated with that job (based on the estimation methodologies described above). For example, to provide an estimate for the completed job with Job ID 7654321, you would use:

```
jobemissions 7654321
```

Typical output from the tool would look like:

```
Job details:
                        Job ID:      7654321
                         Start: 2024-11-11T20:51:25
                        Budget:          t01
                         Nodes:           20
                       Runtime:      324000 s
                            CU:     1800.000
        Compute node energy use:      448.973 kWh
        Other hardware energy use:      67.346 kWh (estimated)
             Overhead energy use:       51.632 kWh (estimated)
              Total energy use:      567.951 kWh (estimated)

Emissions estimates:
        Scope 2:       0.000 kgCO2e (ARCHER2 is on 100% certified
                  renewable energy contract so scope 2 emissions are zero)
        Scope 3:      41.400 kgCO2e (23.0 gCO2e/CU)
          Total:      41.400 kgCO2e

        Indicative emissions estimates for UK national grid energy mix
        in S. Scotland at start of job if ARCHER2 was not using
        renewable energy
            Scope 2:       9.655 kgCO2e (567.951 kWh, 17.0 gCO2e/kWh)
            Scope 3:      41.400 kgCO2e (23.0 gCO2e/CU)
              Total:      51.055 kgCO2e

        Scope 2 carbon intensity values from carbonintensity.org.uk
```

Green Algorithms calculator
What's the carbon footprint of your computations?

Details about your algorithm

To understand how each parameter impacts your carbon footprint, check out the formula below and the methods article

Runtime (HH:MM)	12	0
Type of cores	CPU	
Number of cores	12	
Model	Xeon E5-2683 v4	

IMPERIAL
Information and Communication Technologies (ICT)

Thank you

a.j.richards@imperial.ac.uk