IMPERIAL

Building Sustainable HPC Infrastructure at Imperial College London

22 July 2025

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

IMPERIAL

Digital Plan 2023 - 2028

Enabling Roadmap

ICT department

Goal 4:
Embed a Research and Enterprise Data Culture

We will evolve a comprehensive data culture that gains maximum impact from data and reduces the time-to-value of data, building a robust resource of readily accessible research and enterprise data.

We will establish data-based evidence as the standard by which we challenge the integrity of our strategic decisions and maintain transparent central data governance, adhering to global best-practice for data security, privacy, and ethics.

Outcomes
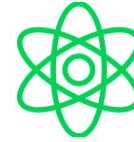
Research data readily accessible across Imperial and beyond

Imperial will be a fully data-driven organisation

Enterprise & Research Data as a Service

We will:

Reduce the number of data incidents

Increase the percentage of research data used

Increase trust and usability of Imperial data

Goal 5:
Enable Research and Innovation

We will focus on fostering partnerships between our world-leading researchers and external organisations to drive innovation, maturing our digital capabilities ready to embrace future disruptions to existing business models in Higher Education.

This will allow us to move to a more boundaryless organisation, establishing next generation collaborative digital environments, collaborating with start-ups, industry partners, and other universities to co-create digital solutions that address pressing challenges in society.

Outcomes

Professionally managed Partner ecosystem

Increased participation in exploring new approaches

Innovation incubation

We will:

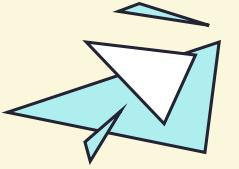Increase the number of innovation initiatives

Reduce spending on 'shadow' IT

Increase active engagement with events

IMPERIAL

Research Computing Services (RCS)

Information and Communication Technologies (ICT)

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

Imperial Data Facility

Data and Metadata Maturity →

Research (Dept/Faculty/Facility)

RCS HPC

Dept HPC

Dept Research Equipment

Dept Research Data Stores

External Facility Research Data

DART /DMP Registration

Active (ICT)

RDF Active

RDF ActiveSecure

Minimal Metadata Applied

Archive (Institutional)

Pub

Spiral

Helix (FAIR Data)

RDF Archive
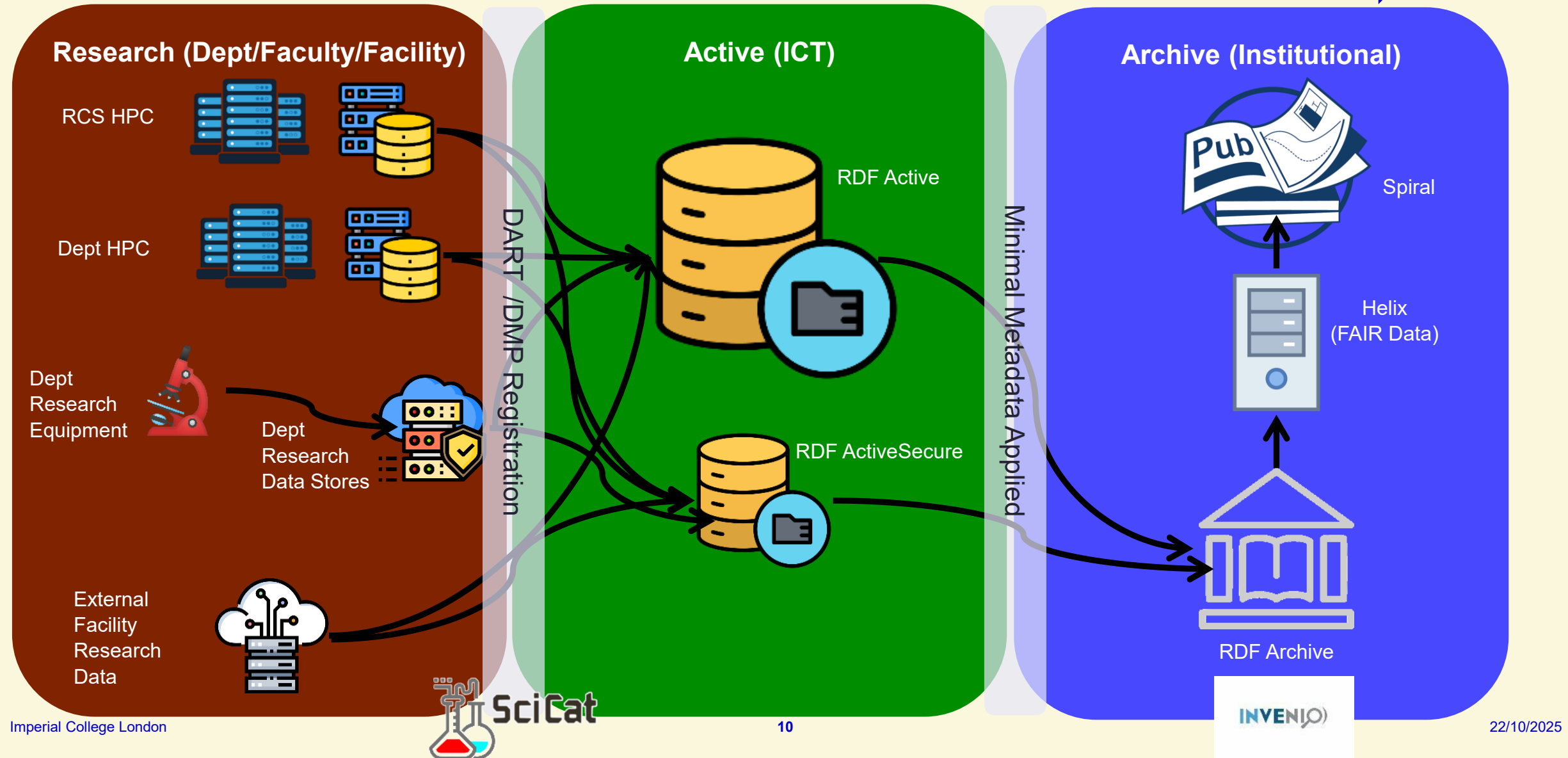
IMPERIAL / intel. / Lenovo

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
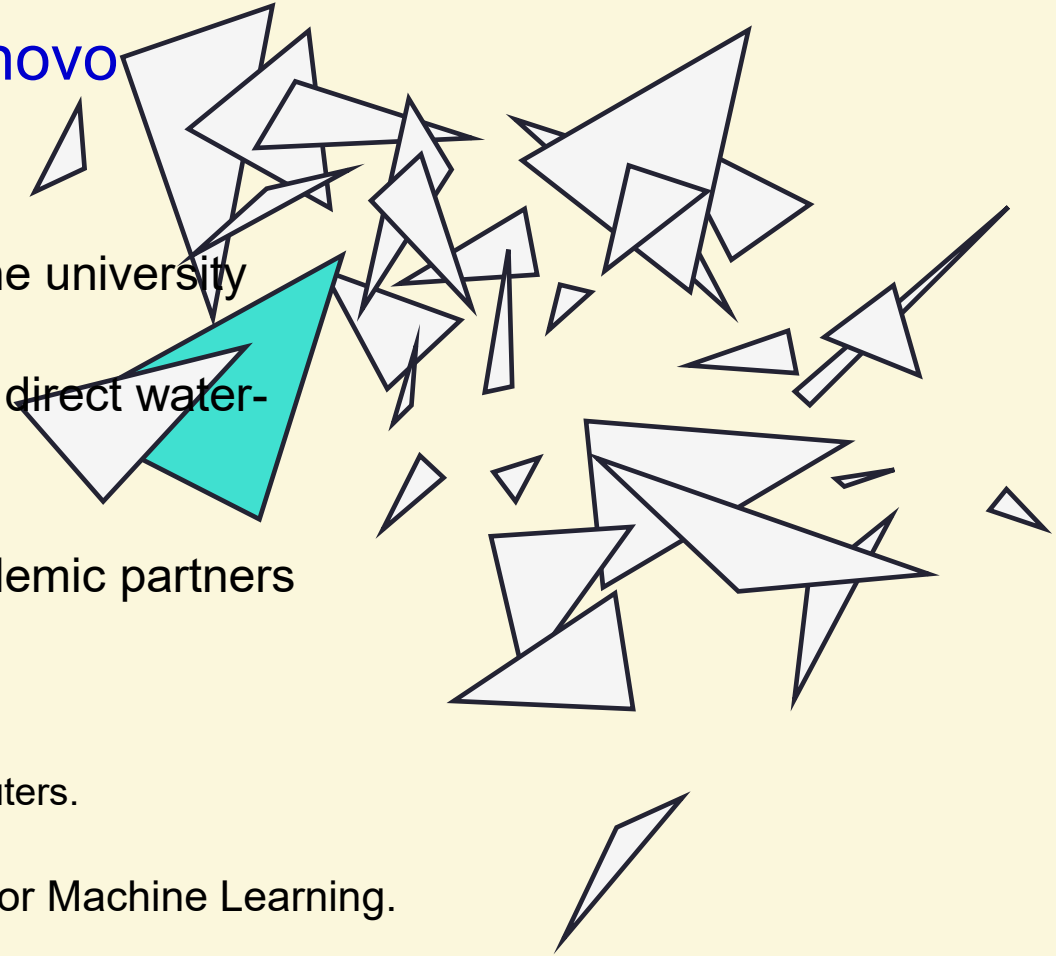
- To foster collaboration around research computing with academic partners

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.

Imperial Research Compute: HEX Platforms

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

HEX-HTC

HX2

HEX-HPC

HX1

HEX-AI

HX3

Data
RDF
Spiral + Helix

HEX-INovArc

HX6

HEX-Secure

HX4

HX5

HEX-Visual

Training
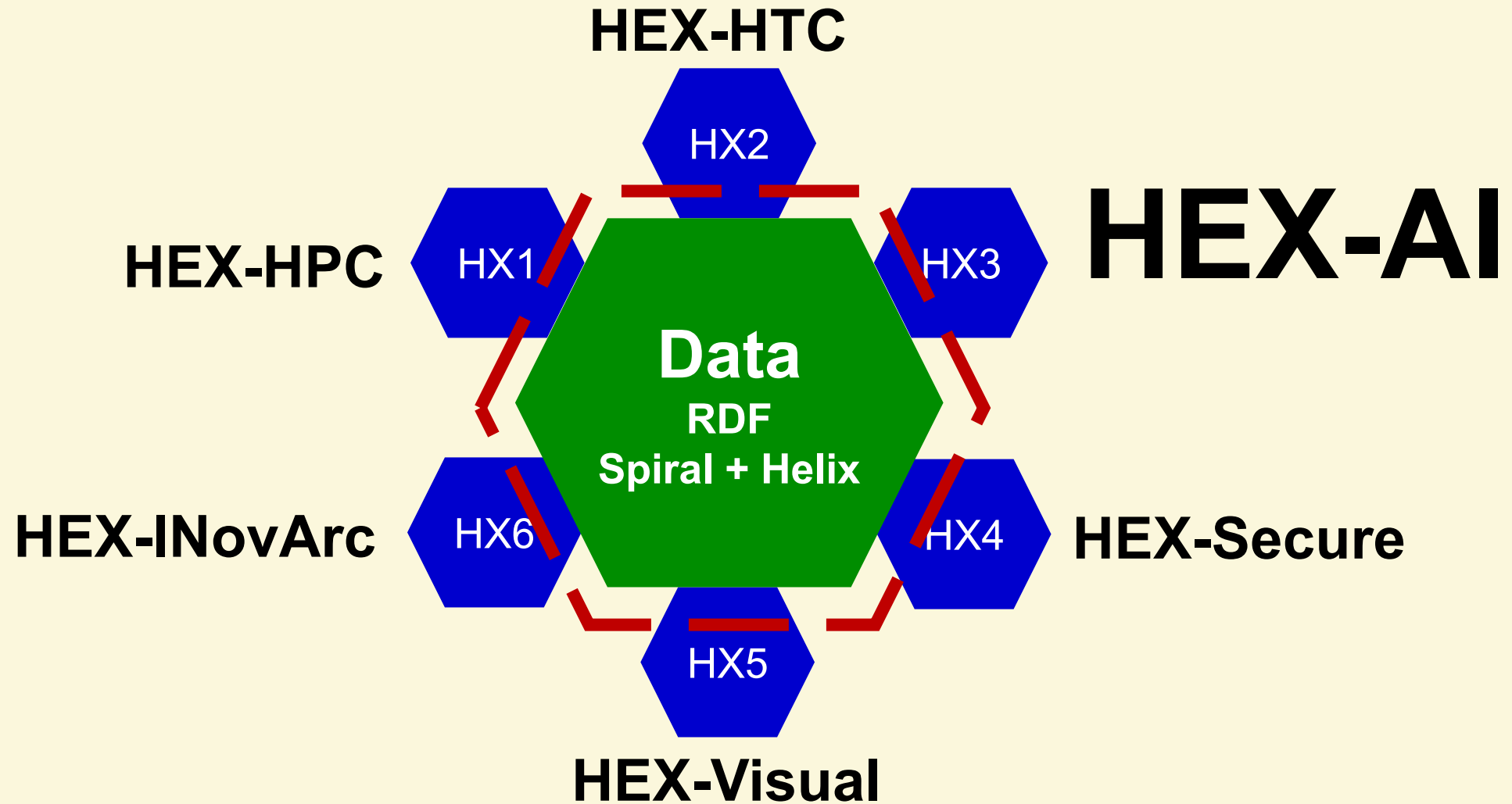
Compute Platform Summary
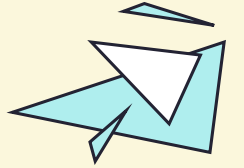
CX3

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
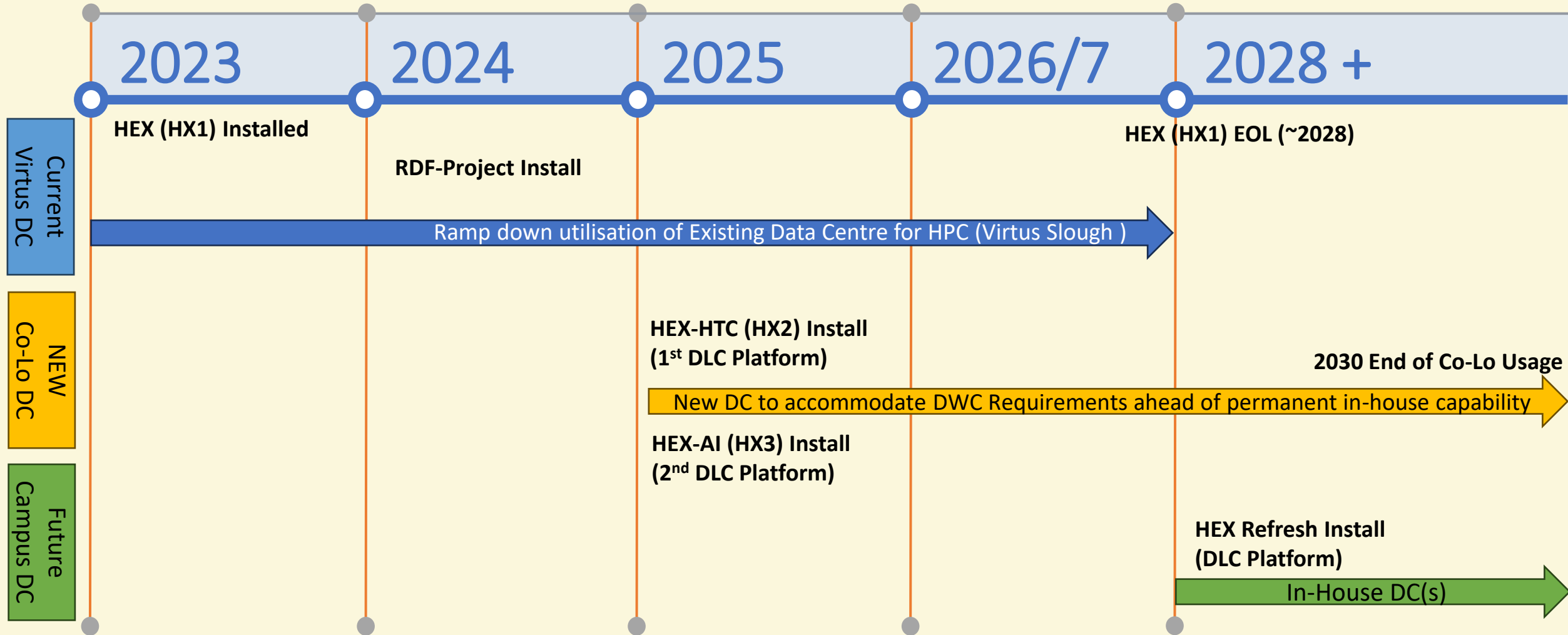Interconnect: Mellanox ConnectX-6 HDR200 (200 Gbit/s) InfiniBand

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

Sustainability
(Roadmaps & What we need for the future?)

Research Computing Net Zero Data Centre Roadmap

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

HEX (HX1) Installed

HEX (HX1) EOL (~2028)

RDF-Project Install

Current Virtus DC

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

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

NEW Co-Lo DC

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)
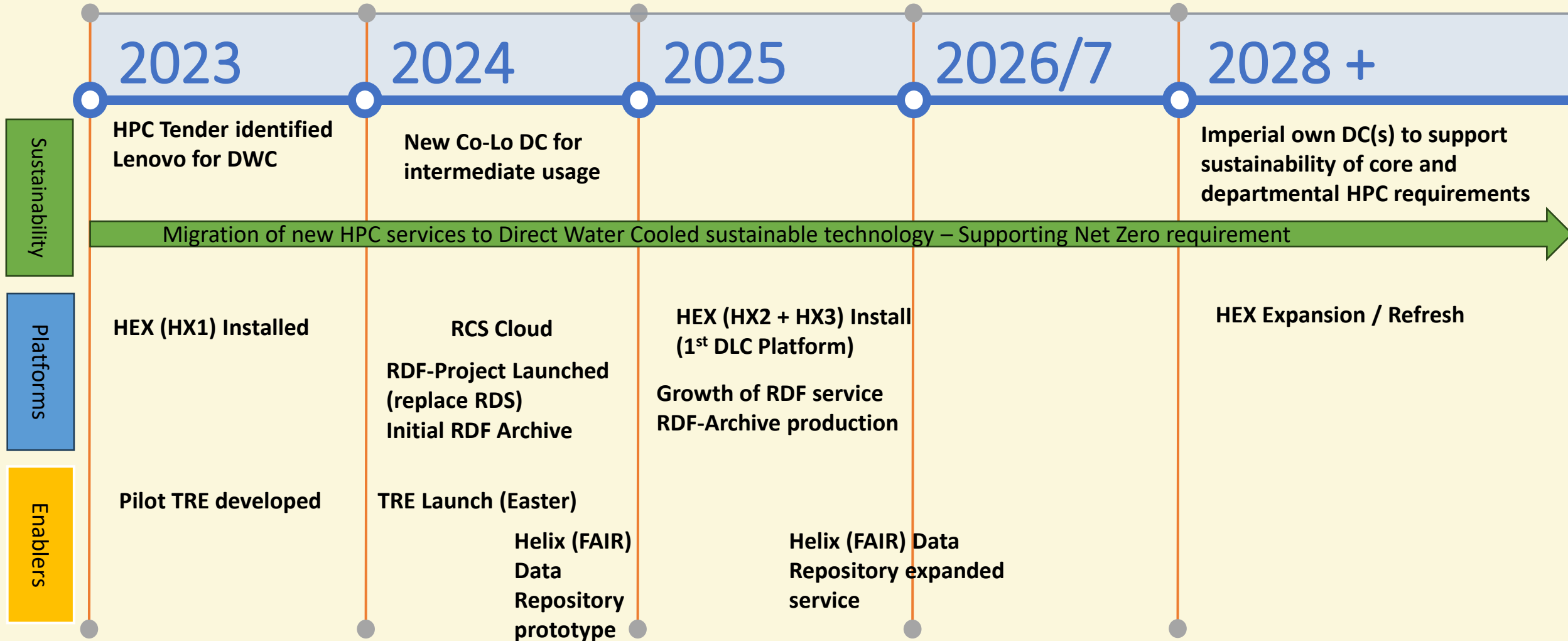
IMPERIAL

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)
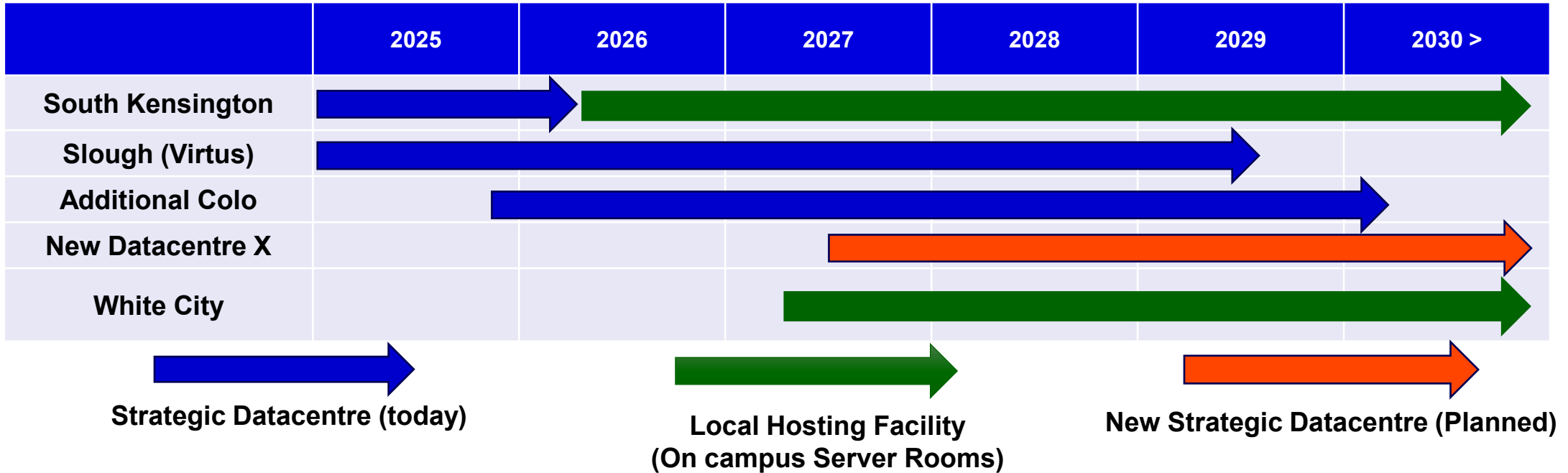
Key Points:

- 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

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	
External Chiller Plant Size	100m2	
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	

Imperial College Provisional Cooling Layout V1.4

Rack Type	Secondary Flow
GPU SD650 N V3	122 lpm
CPU SC750 V4	72 lpm
GPU SC777 V4	240 lpm
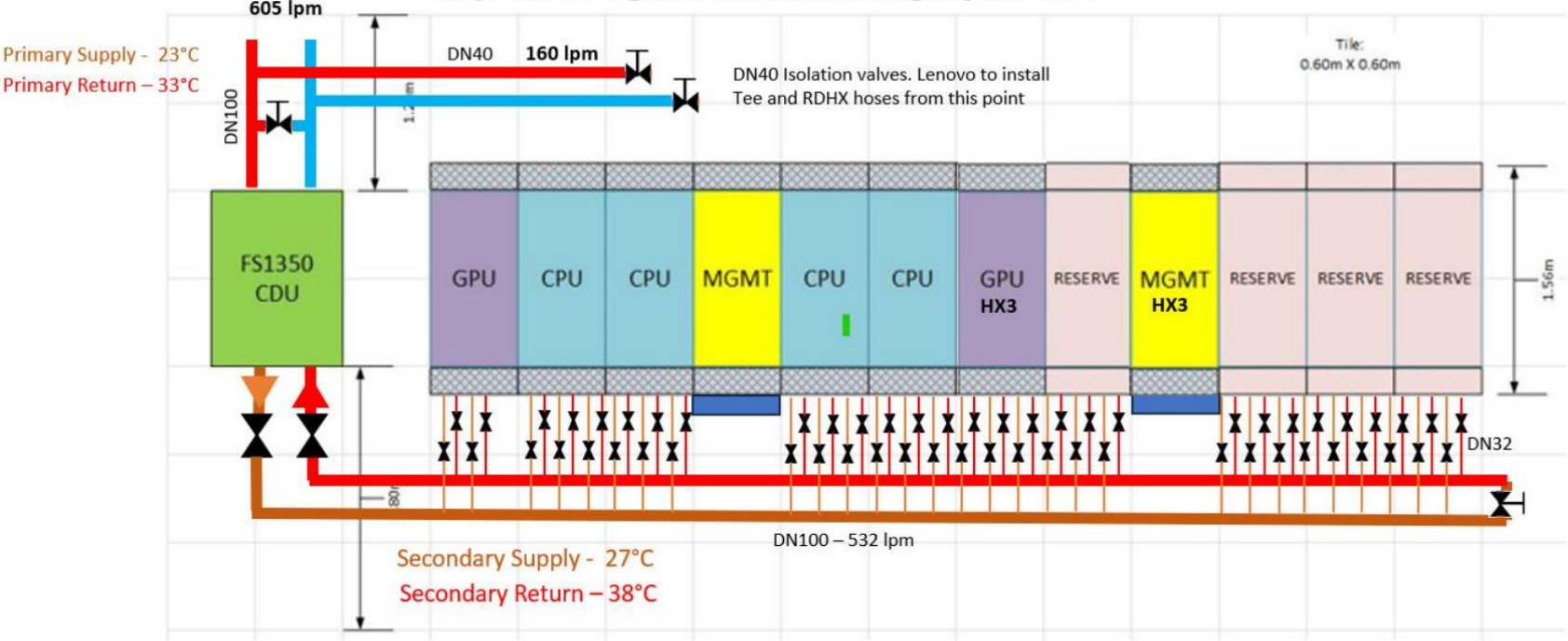
- CDU Secondary Flow HX2 + HX3 = 532 lpm
- 2 pumps running with 1 pump redundant

- Secondary manifold installed below raised floor
- 316 stainless steel construction
- Complete with drip tray and supports All connections shown fitted Day 1.

- RDHX performance at 23C inlet will not cover full load on management racks.

Current Data Centre Capacity

ICT Racks - 171
Total Floor Space – 738m2 (160 on campus)
Contracted Power Rack Ratings – 2.25 MW

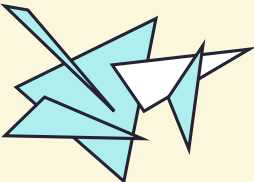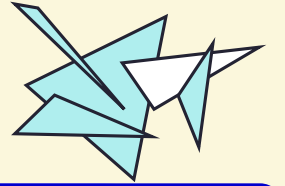| DataCentre	Capacity (mW) excluding cooling	Annual Cost (£) exc Power	Notes
Virtus Slough	1.95	2.8M	136 racks, UPS and Generator protection
South Kensington	0.3	unknown	35 racks, UPS on single feed
New Colo	0.7	1.1M	24 racks (expandable to 1.3mW), delivery date is September 2025

Sustainable Data (FAIR)

(Helix)

Create, **A**nalyse, **P**ublish, **A**rchive

Supporting Research

Pre-Award: Worktribe

Post-Award: Oracle Grants

Analyse & Collaborate: TRE

Collect & Create: Symplectic Elements, Spiral

Quality & Compliance: Q-Pulse, TALAS, REDCap, Openclinica, JPMS

Evaluate & Archive: Archivematica

Access & Reuse: Open Access, FAIR Data

Research Data: End to End Ambition and Challenge

Imperial [University] as a Data Facility

| Data Creation (Experiment / Modelling) | Data Analysis (Experiment / Modelling) | Data Retention (Store / Archive) | Data Publishing (Open/Secure/ Shared/Curation/) |

MetaData Maturity →
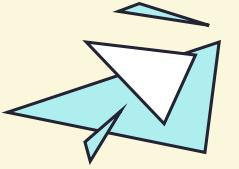
People
Experimentalists

People
Data Engineers

People
Research Software Engineers

People
Research Infrastructure Engineers

Compute / DRI Infrastructure

Supporting Research Data Workflows

A new **Research Data Facility** (RDF) is being delivered in 2024.

Integrated RDF service with key specific service types and use of appropriate solutions - new or existing.

- **RDF – HPC**: fast local storage to underpin HPC (performance),
- **RDF – Active**: general purpose accessible working data storage,
- **RDF – Share:** Globus endpoint and large scale data transfer service
- **RDF – Archive:** cost effective long-term storage of research data.
  - Will underpin 'Spiral-Data' research data repository (FAIR data repository)
- **RDF – ArchiveSecure:** long term secure/sensitive data storage.
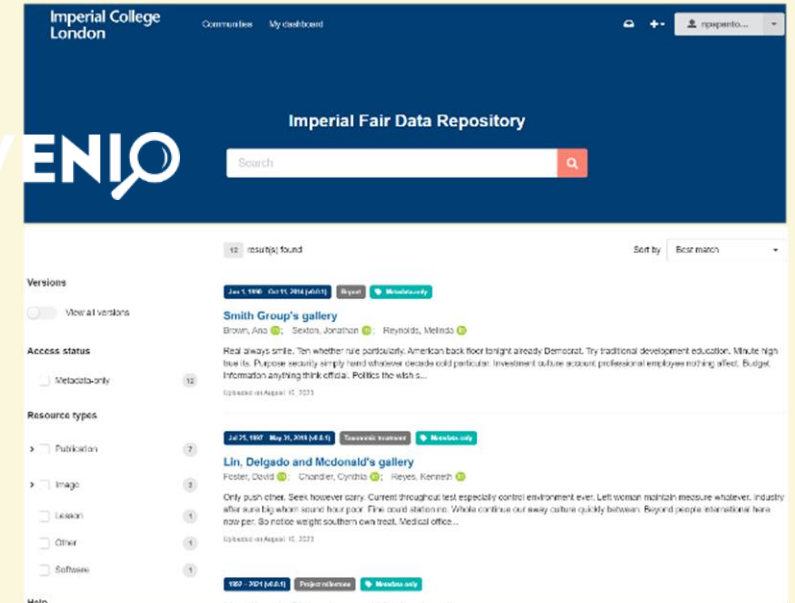
FAIR Data Repository & Service
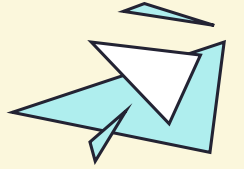
To support publishing, and making data available, at scale

FAIR Data Working Group steering the initiative (headed up by Prof. David Colling)

Leveraging pilot project, based on InvenioRDM (as used by Zenodo/CERN) – pilot funded by Research England (Prof. Michael Bearpark)


- Support from University Research and Enterprise Board (UREB) to plan and deliver a (world-leading) research data repository & service
- Initial deployment baseline using out-of-the-box InvenioRDM - working with Cottage Labs for development, enhancement, and configuration
- Embed as part of the Invenio Community Partnership for future developments and feedback developments from the University

Sustainable Workforce

Developing the workforce and partnerships
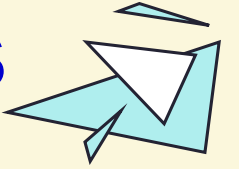
Adopt **Research Technology Professional** (RTP) approach
- Research Software Engineer, Research Infrastructure Engineer
- Junior to Principal career pathway

Develop further HPC/RSE Experience Programme (6 months) post PhD

Broader use of Training Roles/Apprenticeship

*Create an in-house 'First Steps Program': Recruit Imperial undergraduate program*
- *System Administrator*
- *Research Software Engineer*
- *Research Data Engineer*
- *Product Engineer*

*Define **career pathways**: Consistent use of role levels across ICT*
- *e.g. Junior, Senior, Principal, Head, Director, etc.*

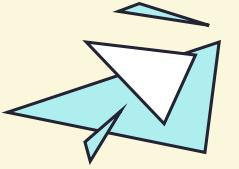*Encourage/Enable staff movement: Cross-department roles*

Research Software Engineering
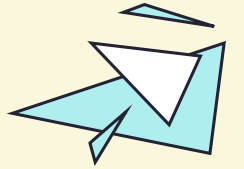
To increase the **quality**, **impact** and **sustainability** of the **research software** developed at Imperial, supporting the university in enhancing its world-leading research outputs.

Open source booster
Free, short (<2 days) collaborations to maximise the impact of open-source software developed at Imperial.

Software projects
Substantial software projects to add features, improve sustainability or build high-quality tools from the ground up.

Legal information
Appropriate licenses for software and aspects relating to Intellectual Property (IP) surrounding its development.

Code Clinics
Free, 30 min fortnightly sessions to provide advice on programming best practices, solve coding issues or guide researchers to the right tools.

BETTER SOFTWARE BETTER RESEARCH
www.software.ac.uk

Research Computing Data Science Teaching Programme

Delivers short courses in:

Statistical concepts

Data analysis

Machine learning

Introductory to intermediate programming - Python, R, Julia, C++, Fortran, MATLAB

Intermediate and advanced programming topics in Python

- Key scientific packages
- Profiling and optimisation
- Object-oriented programming

- Research software engineering and good practice
- Computing environments - Linux, HPC and containers