



# Private cloud system powers up Bristol's academic AI ambitions

One of the UK's foremost academic AI institutions, the University of Bristol is adapting its IT infrastructure to enhance students' dynamic learning environment, and accelerate the pace of world-changing research.

AI supercomputing resources are now essential tools in the pursuit of academic excellence, and AI education has already been embedded across most of the University courses, from law to social science, mathematics and engineering. To sustain and strengthen its services, the University has now introduced a single, unified platform capable of delivering AI services quickly and securely.

The modern private cloud has been built on VMware Cloud Foundation (VCF), with the aid of the cloud consulting company, Xtravirt. It has been designed to allow Bristol to run its student and researcher-focused AI facility as Infrastructure-as-a-Service (IaaS), which is deployed using VMware Private AI Foundation with NVIDIA. This means that students and researchers can now access GPU-as-a-Service with multi-tenancy support for AI workloads, enhanced GPU and virtual GPU (vGPU) monitoring, and simplified model usage through Model Runtime. For the most sensitive projects Bristol can air-gap AI workloads and operate them disconnected from the internet, bolstering data confidentiality.

## Accelerating global connections

The University is already a hub for much of the UK's most cutting-edge AI collaboration, with several dedicated AI courses, catering to the surging demand for graduates who are skilled in the fundamentals of AI, specifically those who can design and build, secure and trustworthy systems. Bristol is home to the Intelligent Systems Laboratory, a federation of several research units that focus on different aspects of AI.

The new IT infrastructure enables undergraduate and postgraduate students to access high-end virtual environments, fostering responsible AI development and ethical considerations. It has been designed to enable a dynamic 'next-gen stack' that bridges basic desktop compute with AI capabilities.

## A unique facility from which to drive progress

Powered by NVIDIA H100 and L40S GPUs, the platform balances the needs for performance for large-scale training, and efficiency for real-time inference. This private AI environment is also global, supporting the University's overseas students who will all be able to access Bristol's AI toolset.

The platform's ability to share GPU and CPU resources flexibly means that AI teaching environments will be provisioned and reallocated across different courses and departments with more ease. Crucially, it means Bristol is able to provide a clear audit trail, showing exactly where data is held.