

DIRAC & VAST DATA INNOVATION PLACEMENT: BUILDING THE NEXT GENERATION OF INTELLIGENT DIGITAL RESEARCH INFRASTRUCTURE

APPLICATION PACK

www.vastdata.com
www.dirac.ac.uk



DiRAC
High Performance
Computing Facility



DIRAC & VAST DATA INNOVATION
PLACEMENT:
BUILDING THE NEXT GENERATION
OF INTELLIGENT DIGITAL
RESEARCH INFRASTRUCTURE

Application Pack

Contents

In collaboration with VAST Data, DiRAC is pleased to invite applications for a 6-month Innovation Placement in building the next generation of intelligent digital research infrastructure.

02 VAST Data company profile

04 About DiRAC

06 Placement details, applicant profile

08 How to apply!



VAST Data

Company background

VAST Data is an AI-native “Thinking Machine” architecture that consolidates the disparate silos of high-performance computing (HPC)—including parallel file systems, object storage, and structured databases—into a single, unified global namespace.

By leveraging a Disaggregated Shared-Everything (DASE) architecture, it eliminates the traditional trade-offs between the massive throughput required for checkpointing and the high-concurrency IOPS needed for random-access inference or multi-agent simulations. For researchers, this translates to a “zero-tier” storage environment where all data resides on high-density flash, effectively removing the need for complex data staging or manual tuning.



VAST delivers the first AI Operating System, natively unifying and orchestrating storage, database, and compute to unleash the true power of agentic computing and data-intensive applications.

With integrated services like the VAST DataEngine and VAST DataBase, the platform not only stores exascale datasets but also provides the computational triggers and metadata indexing necessary to automate the entire research lifecycle, from initial data ingestion to long-context agentic reasoning.

VAST Data's fundamental architecture makes it the only platform able to provide an end-to-end AI ecosystem with the governance and security needed to maintain public trust.

DiRAC HPC

About DiRAC

Established in 2009, DiRAC provides HPC services to the UK's Scientific Research Communities in theoretical cosmology, nuclear physics, astrophysics, particle physics, and solar and planetary science. DiRAC is funded by the Science and Technology Facilities Council (STFC), part of UK Research and Innovation (UKRI). To date, capital funding for DiRAC systems has been provided by the Department for Business Innovation and Skills (BIS), the Department for Business, Energy and Industrial Strategy (BEIS), STFC and UKRI. UKRI is now part of the Department of Science Innovation and Technology (DSIT). DiRAC operations are funded by STFC.



DiRAC hosts three HPC services: the Extreme Scaling Service, the Memory Intensive Service, and the Data Intensive Service, with each tailored to the specific types of computational workflows needed to deliver our Science Programme. Innovation is a key part of DiRAC's activities and all our services are co-designed in collaboration with our research community, our technical and software engineering teams, and our vendor partners.



62% of employers deem relevant work experience of significant or critical importance when searching for new recruits (UK Department for Education, Employer Skills Survey 2022)

Innovation Placements

DiRAC Innovation Placements provide a superb opportunity for doctoral students to collaborate with industry leaders on cutting-edge research projects of mutual benefit. Projects take place over six months and address current challenges with innovative, state-of-the-art solutions, ensuring that the research is relevant and impactful for both parties.

Innovation Placements offer a unique opportunity for researchers to bridge the gap between academia and industry, contribute to impactful research, and gain practical experience in the commercial sector.

Benefits to student

- > **Practical Experience:** A placement allows you to apply theoretical knowledge in a real-world setting, enhancing your understanding of how your research can impact industry.
- > **Skill Development:** You can develop valuable skills critical to your professional development, such as project management, teamwork, and communication in a corporate environment.
- > **Enhance your Research:** Exposure to industry challenges can inspire new thinking, reasoning, and innovative approaches to your research, potentially leading to impactful findings.

Benefits to Career

- > **Improved Employability:** Experience in an industrial setting will enhance your CV, making you more attractive to employers in both academic and non-academic fields.
- > **Confidence Building:** Navigating a professional environment will build your confidence in your abilities and prepare you for future workplace challenges.
- > **Career Exploration:** It provides insight into different career paths outside academia, helping you make informed decisions about your future.
- > **Networking Opportunities:** Placements facilitate connections with industry professionals, which can lead to collaborations, mentorship, and future job opportunities.

Placement Details

Proposed start date: January 2027

The placement invites candidates to build the next generation of a Cognitive Open Research Environment by developing an AI-enabled Federated Digital Research Environment.

The successful applicant will work with cutting-edge infrastructure, including serverless data pipelines and the VAST Data AI Operating System, to transform passive document repositories into interactive knowledge services using Retrieval Augmented Generation (RAG) and agentic interaction.

As a participant in this project, you will play a key role in developing and orchestrating LLM-powered conversational agents using the Common agent framework and Python. Your work will involve managing complex data pipelines, including automated document ingestion, chunking, and vector embedding, through the VAST DataEngine and VastDB. Furthermore, you will integrate state-of-the-art NVIDIA NIM endpoints for inference and reranking to refine the system's Retrieval-Augmented Generation (RAG) capabilities. By the conclusion of the placement, you will have contributed to a significant reference architecture and co-authored a white paper that documents these research outcomes for the global academic and technical community.

Responsibilities

Agent Development & Integration: Design, build, and test LLM-powered conversational agents using Python, FastAPI, and the Common agent framework.

Data Pipeline Management: Configure and optimize serverless data ingestion pipelines, document chunking, and vector embedding generation using the VAST DataEngine and VastDB.

AI/ML Orchestration: Integrate and evaluate NVIDIA NIM endpoints for LLM inference and reranking within the VAST InsightEngine RAG architecture.

System Deployment: Assist in deploying and maintaining the application stack within a containerized environment (Docker/Kubernetes).

Documentation & Output Generation: Author a comprehensive reference architecture, contribute to a white paper/conference paper on the project's findings, and package open-source tools for the academic community.

Collaboration: Work closely with the host organization, DiRAC representatives, and the wider development team to share technology and ensure project milestones are met.

Skills & Experience

We are looking for candidates who have:

- > Minimum of a BSc or MSc in Computer Science, Software Engineering, Artificial Intelligence, Data Science, or a related technical field.
- > Strong programming skills, particularly in **Python** (experience with web frameworks like FastAPI is highly desirable).
- > Foundational understanding of, or practical experience with, Large Language Models (LLMs), **Retrieval-Augmented Generation (RAG)**, **operational AI inference**, **model tuning**, and vector databases.
- > Familiarity with modern deployment stacks, including containerization (Docker) and orchestration (Kubernetes).
- > Strong grasp of software design, execution, automation, and testing metrics.
- > Ability to work effectively as a member of a geographically distributed development team and collaborate across complex organizations.
- > Excellent written and verbal communication skills, with the ability to translate complex technical architectures into clear reference documentation and academic papers.
- > Excellent time management skills, capable of independently prioritizing tasks and meeting deadlines in a fast-paced, innovative environment.
- > Highly effective analytical, problem-solving, and decision-making capabilities.

Equality, Inclusion & Diversity

We welcome applications from all, regardless of background.

How to apply

Complete the application form on our website at dirac.ac.uk/innovation-placements by 9.00am on 28th October 2026.



DiRAC_placements@leicester.ac.uk

Placement Stipulations

Placements are fully funded but you must get your supervisor or PI's permission before applying – under UKRI rules participation in the scheme is only allowed with their consent. Students must be affiliated with a UK based university.

The successful candidate will remain based at their home university. We do our best to offer flexibility; part-time working can be arranged as long as the placement does not exceed 1 year.



DiRAC