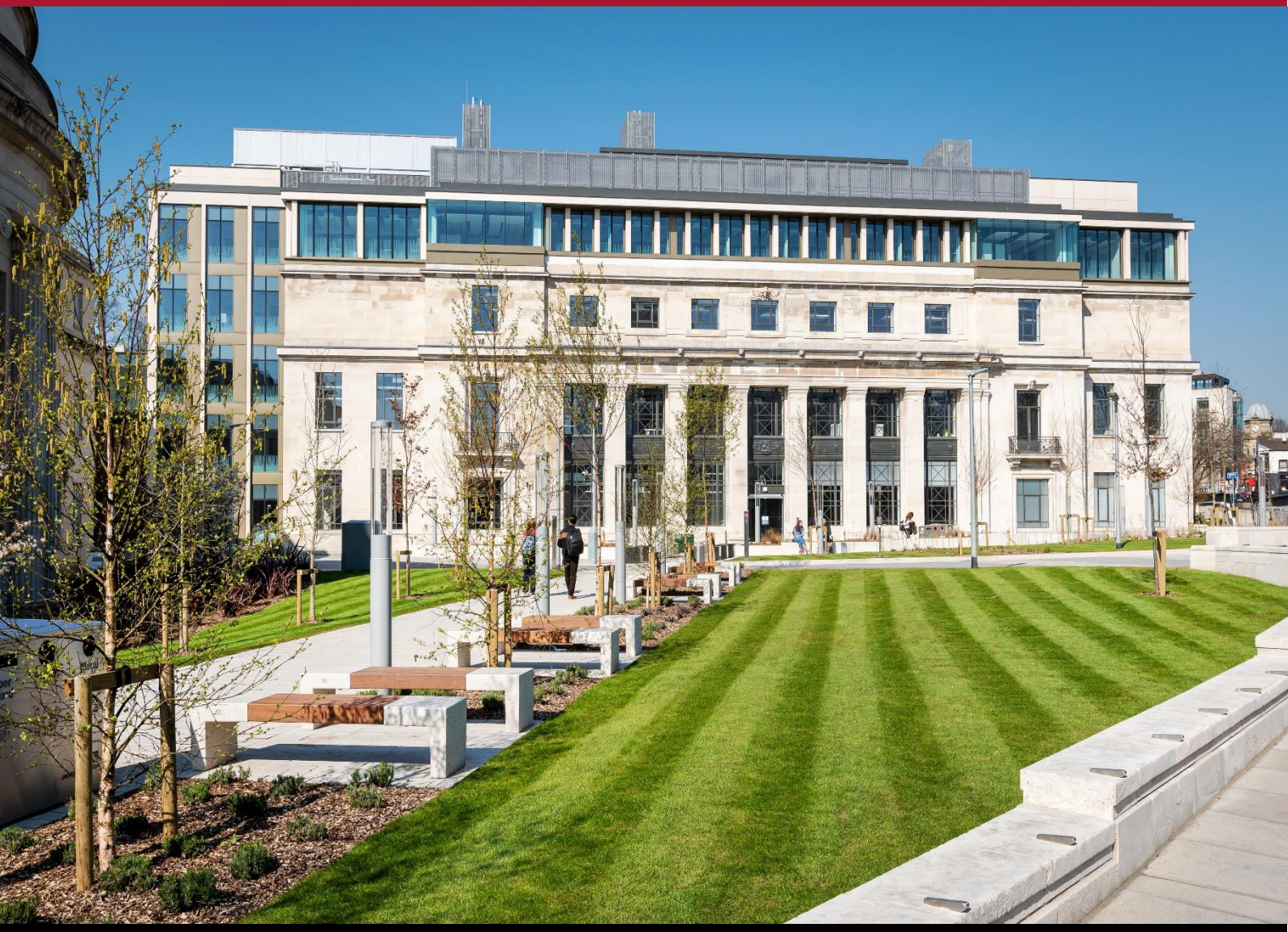
UNIVERSITY OF LEEDS

CANDIDATE BRIEF

Research Fellow in Program Optimization for Large Models

Faculty of Engineering and Physical Sciences



Salary: Grade 7 (£38,205 – £45,585 pro rata p.a.)

Reference: EPSCP1157

Closing date: Sunday 18 August 2024

Part time, 28 hours per week

Fixed term until 31 December 2024

We are open to discussing flexible working arrangements

Research Fellow in Program Optimization for Large Models, School of Computing.

Do you have in-depth technical knowledge of program optimization and distributed systems? Are you interested in working with industry to develop methods to better accelerate LLM workloads in a large environment? Would you like to participate in an LLM-related competition?

In order to reduce the burden for developing LLM-based applications, a novel system is being developed and implemented in this project. This study expands upon a doctoral study conducted at the University of Leeds, which was successfully validated. In an industrial setting, we developed prototypes and methods to identify bottlenecks in the cloud-based environment and to understand the relationship between resources and workloads of LLMs. In light of heterogeneity and cost effectiveness, the proposed solution suppresses current methods.

The purpose of this study is to leverage current libraries and tools to gain a better understanding of the behaviour of the latest LLMs and the requirements of upcoming LLM-based applications. The aims of the project are to develop a lean system with a user-friendly interface that is specifically aimed at algorithm experts. The project will involve the program optimization techniques, for single machine to distributed environment, from heterogeneous memory as well as computing units.

What does the role entail?

As a Research Fellow, your main duties will include:

- Implementing a prototype or testing methodologies to improve an acceleration performance on general LLMs workloads;
- Documenting the methods and results in a clear format for peer review;
- Generating and pursuing independent and original research ideas in the appropriate subject area;
- Developing research objectives and proposals and contributing to setting the direction of the research project and team including preparing proposals for funding in collaboration with colleagues;
- Making a significant contribution to the dissemination of research results by publication in leading peer-reviewed journals and by presentation at national and international meetings;



- Working independently and as part of a larger team of researchers, both internally and externally, to develop new research links and collaborations and engage in knowledge transfer activities where appropriate;
- Evaluating methods and techniques used and results obtained by other researchers and to relate such evaluations appropriately to your own research;
- Maintaining your own continuing professional development and acting as a mentor to less experienced colleagues as appropriate;
- Contributing to the training of both undergraduate and postgraduate students, including assisting with the supervision of projects in areas relevant to the project.

These duties provide a framework for the role and should not be regarded as a definitive list. Other reasonable duties may be required consistent with the grade of the post.

What will you bring to the role?

As a Research Fellow, you will have:

- A PhD (or have submitted your thesis before taking up the role) in Computer Science or a closely allied discipline;
- A strong background in program optimization and distributed system;
- Strong experience in the Large Language Models (LLMs) field;
- Good time management and planning skills, with the ability to meet tight deadlines and manage competing demands effectively without close support;
- Academic publication(s) as first author in top-tier system conferences (i.e. SuperComputing (SC), ASPLOS, OSDI, ATC, SOSP);
- Excellent communication skills both written and verbal, and the ability to communicate your research at national and international conferences;
- A proven ability to work well both independently and in a team;
- A strong commitment to your own continuous professional development.

You may also have:

- Experience of pursuing external funding to support research;
- Experience of using programming languages such as Python, C/C++.



How to apply

You can apply for this role online; more guidance can be found on our How to Apply information page. Applications should be submitted by **23.59** (UK time) on the advertised closing date.

Contact information

To explore the post further or for any queries you may have, please contact:

Prof. Jie Xu, Professor in School of Computing
Tel: +44 (0)113 343 5193
Email: J.Xu@leeds.ac.uk

Additional information

Faculty and School Information
Further information is available on the research and teaching activities of the Faculty of Engineering & Physical Sciences, and the School of Computing.

A diverse workforce
As an international research-intensive university, we welcome students and staff from all walks of life and from across the world. We foster an inclusive environment where all can flourish and prosper, and we are proud of our strong commitment to student education. Within the Faculty of Engineering and Physical Sciences we are dedicated to diversifying our community and we welcome the unique contributions that individuals can bring, and particularly encourage applications from, but not limited to Black, Asian and ethnically diverse people; people who identify as LGBT+; and people with disabilities. Candidates will always be selected based on merit and ability.

The Faculty of Engineering and Physical Sciences are proud to have been awarded the Athena SWAN Silver Award from the Equality Challenge Unit, the national body that promotes equality in the higher education sector. Our equality and inclusion webpage provides more information.



Working at Leeds
We are a campus-based community and regular interaction with campus is an expectation of all roles in line with academic and service needs and the requirements of the role. We are also open to discussing flexible working arrangements. To find out more about the benefits of working at the University and what it is like to live and work in the Leeds area visit our Working at Leeds information page.

Information for disabled candidates
Information for disabled candidates, impairments or health conditions, including requesting alternative formats, can be found on our Accessibility information page or by getting in touch with us at hr@leeds.ac.uk

Criminal record information

Rehabilitation of Offenders Act 1974
A criminal record check is not required for this position. However, all applicants will be required to declare if they have any 'unspent' criminal offences, including those pending.

Any offer of appointment will be in accordance with our Criminal Records policy. You can find out more about required checks and declarations in our Criminal Records information page.

Salary Requirements of the Skilled Worker Visa Route

Please note that this post may be suitable for sponsorship under the Skilled Worker visa route but first-time applicants might need to qualify for salary concessions. For more information, please visit: www.gov.uk/skilled-worker-visa.

For research and academic posts, we will consider eligibility under the Global Talent visa. For more information, please visit: https://www.gov.uk/global-talent.