RESEARCH DATA SCIENTIST / RESEARCH SOFTWARE ENGINEER

THE ALAN TURING INSTITUTE

There has never been a more significant time to work in data science and AI. There is recognition of the importance of these technologies to our economic and social future: the so-called fourth industrial revolution. The technical challenge of keeping our data secure and private has grown in its urgency and importance. At the same time, voices from academia, industry, and government are coming together to debate how these technologies should be governed and managed.

The Alan Turing Institute, as the UK's national institute for data science and artificial intelligence, plays an important part in driving forward advances in these technologies in order to change the world for the better.

The Institute is named in honor of Alan Turing, whose pioneering work in theoretical and applied mathematics, engineering and computing is considered to have laid the foundations for modern-day data science and artificial intelligence. The Institute's goals are to undertake world-class research, apply its research to real-world problems, driving economic impact and societal good, lead the training of a new generation of scientists, and shape the public conversation around data and algorithms.

After launching in 2015 with government funding from EPSRC and five founding universities, the Institute has grown an extensive network of university partners from across the UK and launched a number of major partnerships with industry, public and third sector. Today it is home to more than 400 researchers and a talented business team.

THE RESEARCH ENGINEERING GROUP

The permanent research staff of the institute's Research Engineering Group work to realise cutting edge research as professionally usable software tools and to apply these to address real-world data science and modelling challenges.

The group's staff are research software engineers and data scientists. We note the considerable overlap between these emerging roles and embrace the breadth of interdisciplinary skills and diversity of approaches entailed in these fields. Staff can choose either job title and change their choice as their career progresses.

In contrast to traditional research careers, we are committed expert collaborators, joining research teams to further the Institute's challenges. We collaborate with scholars across the institute's research community to enhance the applicability of research for particular problems. We work with clients in industry, government and the third sector to turn their data challenges into research questions. We value expertise across many domains and rely on this diversity to design tools, practices and systems to harness the power of data science around the world.

We create software and scripts that implement research and apply it to client data in a readable, reliable and reproducible fashion. We present conclusions of research and analysis to the research community and clients through presentations, research papers, and interactive data visualisations. We work with state-of-the-art, high performance computing and cloud platforms to realise collaborators' data science and artificial intelligence research at scale.

We support the dissemination of research outputs through the publication and maintenance of open source research software packages. We contribute to the sustainability of the open source ecosystem by adding features, fixing bugs, maintaining tools, and supporting community management in new and existing packages.

THE ROLE

Successful candidates will:

1. Apply state-of-the-art and novel data science and artificial intelligence techniques emerging from the Institute and elsewhere to problems faced by the Turing's clients
   - Understand the problems of clients in the public, private and third sectors, and develop appropriate approaches to solving these problems.
   - Understand which data are, or might be, available; and collect and manage this data.
   - Perform analyses, which might include: building statistical models; applying machine learning techniques; building models and simulations; or applying optimisation techniques.
   - Document processes for effective and efficient reuse across multiple domains.
2. Collaborate with research colleagues to develop and maintain software embodying research outputs
   - Develop a good understanding of the relevant theory and the needs of potential users of the software
   - Be responsible for the programming effort, including design and planning
   - Test and validate the software to a high quality standard
3. Present, disseminate and explain our work
   - Feedback the outcomes of analyses to clients and customers in the public, private, and third sectors in written form and in presentations.
   - Share research in the practice of data science and artificial intelligence with the scholarly community through research papers and conferences.
   - Publish, distribute, document and maintain research software packages.
4. Contribute to the life of the Institute and support its community
   - Deliver teaching and training to colleagues and students, including within the team in our regular skills sessions.
   - Support research colleagues to make the most of the institute's secure high performance computing environments for advanced research.

PERSON SPECIFICATION

Essential

Candidates must be able to demonstrate, through examples, all of the capabilities listed below:

- Experience managing, structuring, and analysing research data.
- Experience managing and organising the parameters and results of computational experiments.
- Fluency in one or more modern programming languages used in research in data science and artificial intelligence. (We particularly work in R, Python, and modern C++, but demonstrable use of other programming languages for research, together with a facility for learning new languages, is most welcome.)
- An understanding of the importance of good practices for producing reliable software and reproducible analyses (e.g., version control, issue tracking, automated testing, package management, literate analysis tools such as Jupyter and Rmarkdown)
- Demonstrated enthusiasm and ability to rapidly assimilate new computational and mathematical ideas and techniques on the job and apply them successfully.
- Excellent written and verbal communication skills, including experience in the visual representation of quantitative data, documentation of software packages or data resources, the authoring of research papers or technical reports, and giving presentations or classes on technical subjects.
- Ability to lead one's own work independently, including planning and execution, and to collaborate productively as part of a team.

In addition, for junior staff only:

- An MSc degree or equivalent professional experience in a field with significant use of both computer programming and advanced statistical or numerical methods.
- Any **one** of:
  - attendance at courses (eg. summer schools); **or**
  - significant contributions to open sourced community software; **or**
  - short term work experience/placements contributing predominantly to software development or data analysis; **or**
  - equivalent alternative experience.

In addition, for appointments above junior level only:

- A PhD degree or equivalent professional experience in a field with significant use of both computer programming and advanced statistical or numerical methods.

Desirable

We are a learning team and combine many techniques and approaches across our projects. Therefore a commitment to developing new expertise is very desirable to us. In addition to the skills you already have, we will look for evidence of a commitment to learn by looking at new skills you have developed in previous roles, and your interest and plans for acquiring future skillsets.

In our team, different team members develop different areas of expertise according to their interests and we look to combine these across the team to support the full range of projects we work on. Listed below are examples of some of the skills we are looking to have within the team. If you have experience of one or more of these skills in addition to the essential skills listed above, we strongly encourage you to apply.

- Machine learning, including experience with one or more established software libraries.
- Computational statistics, particularly Bayesian modelling.
- Visualisation for presenting large, complex, or high-dimensional data
- Knowledge management and ontology engineering, semantic web.
- Mathematical and computational modelling of complex systems.
- Logic, planning, verification, and automated reasoning.
- Programming language and API design. Domain specific languages.
- Exposure to mixed or qualitative research methods
- User interface design and development with web technologies, especially for data visualisation and knowledge representation.
- Writing technical documentation.
- Advanced numerical simulation (e.g. FEM, CFD…)
- Experience with public cloud platforms.
- Experience working with confidential and sensitive data for research.
- Developing for high-performance computing hardware (CUDA, MPI, OpenMP).
- Experience contributing to, maintaining and/or leading open source research software projects.
- Experience building open source communities.
- Working with databases and APIs for the acquisition of parameter information for models.
- Experience working with legacy code, especially in traditional scientific programming languages (eg, Fortran, MATLAB, C).
- Developing and/or delivering teaching and training in computational or mathematical methods for research.
- Developing and/or delivering teaching and training in applications of data science methods for non-programming experts.
- Automated testing, software quality assurance and continuous integration.
- Code review in a distributed team.

Our list is not exhaustive. If you have skills and experience you feel are transferable to the role we are offering, we are very interested to hear from you. We welcome any informal inquiries in regards to this.

TERMS & CONDITIONS

The salary offered for this position is: £30-34,000 Junior Research Software Engineer / Junior Research Data Scientist; £35-45,000 Research Software Engineer / Research Data Scientist.

This is a full time, permanent post, to be held at the Institute's site at the British Library, Euston Rd, London.

Although this is offered as a full time role based in the London offices, we are extremely supportive of other working models compatible with candidates' lives. Requests to work flexibly, in location or in time, or other reasonable adjustments, will be given positive consideration.

A generous benefits package includes flexible working, 30 days' holiday excluding bank holidays, Cycle2Work, childcare vouchers, contributory pension, health and life assurance and range of other benefits that you would expect from a good employer.

Secondments from University Partners will be considered for a minimum two-year period of secondment.

HOW TO APPLY

Please submit a CV (maximum 3 pages, no photos) and a cover letter telling us why you would like to work at the Turing. It is common for people to be shortlisted for interview solely on the basis of evidence contained in their CV and cover letter. However, if you are happy to share links to research papers, blog posts or Github repositories containing work that you have made significant contribution to, please submit them along with your application, we will also include these in our evidence base for shortlisting.

Interview

As part of their interview candidates will be expected:

- To prepare a 10 min presentation on code the candidate has written that **either** demonstrates some important aspects of research software engineering **or** demonstrates an important data science algorithm. Existing examples of work are encouraged as long as the code was substantially written by the candidate and the candidate should be prepared to answer questions about both the code and the research challenge it addresses or the algorithm it demonstrates. Any source code shared for the interview will be treated in the strictest of confidence.
- To describe to us their previous experience and competencies for the role. There will also be a problem solving discussion with the interview panel, using a whiteboard and pen to arrive to an understanding of a proposed data analysis challenge.

Further information

If you have queries or would like to discuss the role further, please contact Dr James Geddes, Principal Research Data Scientist, at jgeddes@turing.ac.uk.

EQUALITY DIVERSITY AND INCLUSION

The Alan Turing Institute is committed to creating an environment where diversity is valued and everyone is treated fairly. In accordance with the Equality Act, we welcome applications from anyone who meets the specific criteria of the post regardless of age, disability, ethnicity, gender, gender reassignment, marital and civil partnership status, pregnancy, religion or belief or sexual orientation. Reasonable adjustments to the interview process can also be made for any candidates with a disability.
Happy to talk flexible working.

Please note all offers of employment are subject to continuous eligibility to work in the UK and satisfactory pre-employment security screening which includes a DBS Check.

Full details on the pre-employment screening process can be requested from HR@turing.ac.uk**.**

The Alan Turing Institute is committed to creating an environment where diversity is valued and everyone is treated fairly. In accordance with the Equality Act, we welcome applications from anyone who meets the specific criteria of the post regardless of age, disability, ethnicity, gender, gender reassignment, marital status, pregnancy, religion or belief or sexual orientation. Reasonable adjustments to the interview process can also be made for any candidates with a disability.